# Table of Contents

*Photographic Life Logging and the Exploration of Autobiographical Memory* .......................................................... 3-8  
Brittney Medina, Faculty Mentor: Jesse Rissman, Ph.D.

*The Role of The Medial Temporal Lobe In Non-Mnemonic Processes* ................................................................. 9-14  
Edward Owens, Faculty Mentor: Jesse Rissman, Ph.D.

*Categorical Generalization in Toddlers versus Adults* ................................................................. 15-22  
Elham Zargarkalimi, Graduate Mentor: KP Thai

*Holding All the Cards: The Effect of Prior Learning on Study Self-Scheduling* .......................................................... 23-29  
Gayan Seneviratna, Faculty Mentor: Robert Bjork, Ph.D.

*Virtual Reality- Bridging the Gap* ................................................................. 30-37  
James Mutter, Faculty Mentor: Jesse Rissman, Ph.D.

*Perceptual Learning in Classical Music* ................................................................. 38-44  
Joselyn Ho, Faculty Mentor: Philip Kellman, Ph.D.

*Analogical Problem Solving In Animation Source Training* ................................................................. 45-54  
Kevin Ruiz, Faculty Mentor: Hongjing Lu, Ph.D.

*Spatial Navigation and Autobiographical Memory* ................................................................. 55-60  
Laryssa Storozuk, Faculty Mentor: Jesse Rissman, Ph.D.

*The Effect of Fractions and Decimals On Reasoning* ................................................................. 61-69  
Marvin Lopez, Faculty Mentor: Keith Holyoak, Ph.D.

*Perceptual Learning and Electrocardiogram Interpretation* ................................................................. 70-77  
Olga Mercer, Faculty Mentor: Philip Kellman, Ph.D.

*Dimensional Change Card Sort, Flanker, and Speeded Math Task* ................................................................. 78-85  
Sarah White, Faculty Mentor: Catherine Sandhofer, Ph.D.

*Reward and Negative Affect Study* ................................................................. 86-89  
Sonia Bhatia, Faculty Mentor: Naomi Eisenberger, Ph.D.
Adaptive Sequencing versus Fixed Spacing Schedule: Known Schedules

Tim Yu, Faculty Mentor: Philip Kellman, Ph.D.

**These research papers are written by undergraduate students as part of the capstone requirement for the Cognitive Science major.**
Photographic Life Logging and the Exploration of Autobiographical Memory

Brittney Medina

Introduction

Autobiographical memory consists of episodes recollected from an individual's life, based on a combination of episodic (personal experiences of objects, people and events) and semantic (general knowledge and facts about the world) memory. Recovering autobiographical memory (AM) requires an effortful reconstructive process in which the individual must search for specific visual images and emotions experienced for a certain event. For this reason, qualities such as emotion and vividness contribute a significant portion to the formation of an autobiographical memory (Cabeza et al., 2007). Many studies have shown that there exist neural differences in brain activation when subjects encounter laboratory memory (LM) (e.g. memories generated by the laboratory or memories from other participants), and AM (e.g. using memories for events from their own life) (St. Jacques et al., 2011). Compared with the laboratory condition (e.g. other’s memories), using real events from a person’s own life elicited greater activity in regions associated with self-referential processing (medial prefrontal cortex), visual/spatial memory (visual and parahippocampal regions), and recollection (hippocampus) (Cabeza et al., 2004). This exemplifies the importance of emotion and vividness in the formation of an AM.

In addition to the vividness or distinctiveness of a memory, the temporal order of a memory is also an integral and defining characteristic of AM (St. Jacques et al., 2008). In other words, it is not only important to know what happened during an event, but also when it happened. A study by Peggy St. Jacques et al. investigated temporal-order memory for autobiographical events vs. laboratory-induced events using an fMRI. The study concluded that temporal-order memory for autobiographical events involves both reconstruction processes
(effortful thoughts that include retrieving contextual details and using them to infer the order of past events) and distance processes (less effortful operations that rely on feelings associated with the strength of the memory), whereas laboratory events involve mainly distance processes. This is an important finding because it shows that memory of self-induced events activates the left dorsolateral PFC while memory for laboratory events activates the right dorsolateral PFC.

The focus of the present study I have been assisting with involves the neural correlates of AM and the differences between self and other photographic sequences as well as the temporal order of the photo sequences (whether scrambled or intact). We hope to show that using one’s own set of photographic sequences versus other’s photographic sequences has a profound impact on the conclusions reached regarding human memory retrieval and its neural substrates. We want to know what happens in the brain when we compare a “first person perspective” memory with a “third-person perspective” one. In essence, through the use of functional neuroimaging techniques and high definition camera devices, the present study is able to delve into the neuroanatomy of our personal past.

**Methods**

The current paradigm captures life events for use as stimuli. So far, we have gathered results for 5 test subjects (3 male, 2 female with a mean age of 20.6). Participants are asked to wear a small, wearable, high definition camera around their neck for a period of 3 weeks. The camera is turned on for a minimum of 8 hours per day. During the course of the experiment, we are able to capture thousands of photos in order to compile a total of 120 event sequences (40 sequences per week) for each participant. Each sequence (comprising of 8 photos) depicts an event that has occurred within a 15-minute window. These event sequences represent unique or
distinctive events that were captured by participants’ cameras. The photographs within an event sequence were chosen to best depict the temporal progression of that event.

I played an integral role in creating these event sequences for two of the participants by creating tags using Picasa, a program for organizing and editing photos. When choosing the images I was going to use, I had to keep in mind a few important points: the aim for a particular sequence was to create a linear progression (tell a “story”) to recreate a unique or personal experience. It was important to avoid “generic” photos that were unmemorable and/or mundane. For example, 8 pictures of different parts of Bruin Walk might not be unique or memorable and is considered a generic event. Sometimes, however, (because participants were UCLA students) it was inevitable to gather generic photos since students spend most of their time in front of a computer screen studying. It was sometimes difficult to gather only distinctive sequences since I only had so many photos to work with. It was also important to avoid using photos with participants’ body parts or articles of clothing in the photos since these features are highly distinctive. This would reveal aspects of the participant’s identity, which, in turn, makes the photo very easily identifiable enough to tell that it is their own photo.

In my opinion, tagging photos was a very interesting experience because I felt like I was trying to reconstruct a person’s life. I was able to see their daily routines and basically walk in their shoes. Doing this was important because, in this way, I am more able to choose the photos that seemed to be memorable or photos that seemed to be noteworthy. I even got to “visit” places on campus I’ve never seen before. The most intriguing part, however, was to actually meet them in person during the testing session of the study (fMRI scanning). It was a bit strange as I felt like I knew their life.
Reverting back to the study, after wearing their camera for the pre-specified time period, participants were given a preview condition (where they were shown a small portion of the event sequences before testing occurred) a week after they last wore their cameras. For this condition, subjects were shown a sequence of 8 photos (each for .8 seconds one at a time with a .2 second interval in between photos) on a computer screen. Participants were asked to indicate how “distinctive” a photographic sequence is. The stimuli are 60 of their own event sequences and 60 of another participant’s sequences.

The testing session of the study (fMRI portion) occurred the day after the preview condition was administered. For this part, participants were shown 240 event sequences during the scan session. Half of all event sequences were from their own lives and the other half were from another participants’ life. This time, participants had 4 seconds to respond with one of the 4 options: Self-Intact (the event sequences were derived from their own photographs with the temporal order unaltered) Self-Scrambled (the event sequences were derived from their own photographs with the temporal order altered) Other-Intact or Other-Scrambled. The scrambled condition refers to the temporal manipulation of the last 4 photos in a sequence (in which photos are not in the same order in which they were originally captured by the camera).

Helping run participants in the fMRI scanner was also a very valuable and worthwhile experience. I was actually able to go into the scanner myself and act as a pilot subject in this study; Tiffany and Andrew scanned me to test for any potential flaws or problems they could encounter while running an actual test subject. These problems ranged from getting the machine up and running to code/script problems. Going into the scanner helped me get a glimpse of what subjects experience while performing the task for the study. I also learned how to prep the subject to enter the scanner. This includes “de-metalling” them, making sure they have ear
protection, informing them about what they will be experiencing and informing them of how to call for help if a situation arises.

**Conclusion**

fMRI data were analyzed with an event-related model. Regions of interest (ROIs) were obtained from a meta-analysis of fMRI studies examining the successful retrieval of autobiographical memories. Condition-specific activity estimates (β values) were extracted from the ROIs and averaged across subjects. Results show activation in several regions previously implicated in autobiographical memory retrieval; this includes the posterior cingulate cortex, retropsinal cortex, hippocampus and parahippocampal gyrus. Specifically, results showed that the posterior cingulate cortex is selectively engaged during the retrieval of personally experience events. These results support the findings found in the Cabeza et al., 2004 article. Additionally, an interaction was found between the source of the memory and the intactness component (whether temporal order was manipulated).

In sum, working in the Rissman Memory lab this quarter with Tiffany and Andrew has been a very beneficial experience for me. I was able to get hands on experience working with the fMRI scanner as well as contributing directly to the acquisition of stimuli for the study. I hope to keep learning more about autobiographical memory to increase my knowledge on the connectivity of the brain during memory retrieval.
References


The Role of the Medial Temporal Lobe in Non-Mnemonic Processes

Edward Owens

I worked in Dr. Rissman’s Memory Lab this quarter analyzing data for a previously conducted study by his graduate student, Andrew Westphal. The data analysis included ROI (region-of-interest) reconstruction, resting-state analysis, and surface ROI visualization. Mr. Westphal’s experiment utilized fMRI to investigate the role of medial temporal lobe structures (mostly the hippocampus, perirhinal cortex and parahippocampal cortex) in episodic memory and analogical reasoning. The study was particularly focused on the role of these structures while working memory was heavily taxed and further investigation of the hippocampus’s role in analogical reasoning. Previous research has shown the hippocampus to play a vital role in episodic memory retrieval primarily from the work done with the amnesiac patient H.M. However, more recent studies have shown the hippocampus and other medial temporal lobe structures to be heavily involved in non-mnemonic functions such as language (Kurczek, Brown-Schmidt, & Duff, 2013). The evidence that the medial temporal lobe is greatly involved in not just memory but multiple cognitive tasks is what inspired the experiment conducted by Dr. Rissman’s lab. I will begin this paper by investigating the role of the medial temporal lobe in non-mnemonic tasks and conclude with proposed theories as to how this might happen.

First, to understand the MTL (medial temporal lobe) we need to discuss the basic forms of memory. We now recognize that memory can be segregated into explicit categories; the declarative and non-declarative. Declarative memory can be thought of as the personal and non-personal factual memory which consists of semantic memory (non-personal), and episodic memory (personal). Non-declarative memory (aka procedural memory) can be thought of as every other type of memory but most commonly refers to habit and skill memory. Patient H.M., an epileptic patient who underwent a bilateral removal of the temporal lobe, including the
hippocampus, provided evidence that the role of the MTL is highly involved in declarative forms of memory, but more recent research, such as that conducted by Shohamy and Browne (2013), has shown the structures to be involved in other forms of cognitive processing as well.

It is well-documented that the striatum is fundamental to learning in operant conditioning paradigms, but multiple studies conducted by Shohamy and colleagues revealed a relationship between reward learning through neuronal dopamine modulation of the striatum and the hippocampus. The striatum is implicated in immediate feedback type learning for example, if a rat were to immediately receive a reward for pressing a lever, the learning that occurs is attributed to the striatum. However, if the rat were to receive a delayed reward or no reward at all, the learning, according to the article, is attributed to the hippocampus as well as the striatum (Shohamy & Browne, 2013). These results suggest that the hippocampus plays a major role in habit learning and most importantly, may not be restricted to declarative forms of memory.

The article by Shohamy and Browne displays the role of the MTL in a type of learning that does not necessarily involve conscious thought but is still intrinsically tied to memory. The most common non-mnemonic cognitive process is language, in which Kurczek, Brown-Schmidt, and Duff (2013) have recently found the hippocampus to play a fundamental role. A key computational function the hippocampus is thought to employ in declarative memory is its capacity to relate novel concepts with memory, this function is called online representational binding. The evidence provided in Kurczek et al.’s paper suggests that the hippocampus uses these capabilities of online representational binding in language as well, where it is able to store the temporary referent of a pronoun to the context of the sentence. For example, “Melissa is playing the violin for Manny as the sun is shining overhead. He is wearing a yellow bracelet, and it looks as if the song is being played well.” (Kurczek, Brown-Schmidt, & Duff, 2013) The
The hippocampus seems to be involved in binding the word “he” and the referent, in this case, Manny.

The study was conducted by recruiting a group of subjects with damage to the MTL, a group of subjects with damage to the frontal lobe and a control group of healthy subjects. The experimenter’s tracked their eye movements as they read through two sentences, such as the one in the previous paragraph, and presented them with a scene that corresponded to the sentence. The data was derived from a ratio of how long the subjects fixated on the correct vs. incorrect referents as the story was being read.

The authors found the MTL damaged patients to do significantly worse than both the patients with frontal lobe damage and the control subjects suggesting confusion during the presentation of the story. Even with referential stories as short as two sentences, the MTL damaged patients had far more trouble discerning whom the pronoun was referring to compared to the other two groups, especially when the characters in the story were the same gender. These findings offer more evidence that the hippocampus is not solely involved in declarative memory processes but also real-time language processing. It also suggests that the frontal lobe is not involved with the maintenance of online referents in language and this process seems to be strictly limited to the MTL.

The MTL is not thought of to be a quick working and real-time processor of information. It has been identified in the field of psychology to be fundamentally tied with the concept of long-term memory and simply aides in contextual binding of situations based on previous experience or learning. Prior research has established the MTL’s role in online representations and relational context over short delays by studying patients with MTL lesions, whom could not perform on tasks with very minimal delays even as low as a second (Warren et al., 2012).
studies reveal the MTL’s role in short-term memory tasks but they still involve memory. In Warren et al.’s study, the experimenters wished to reveal the role of the MTL in relational representation with no delay. Such a study could provide evidence for the MTL’s role in online representations and could also demonstrate its capability to do more than simply relate contexts in novel situations based on memory.

The study was conducted with two groups of lesion patients, one group with lesions to the MTL, another group with lesions to the frontal lobe, and a control group. The subjects were tested with three different experiments. Experiment 1 was an identification task that required the subjects to identify a series of scrambled pictures. A ratio of correct and incorrect answers were taken and analyzed between groups. The MTL lesion patients performed significantly poorer than either the frontal lobe lesion group or the control group.

Experiment 2 employed two different versions of the overlapping figures test which tests the ability of a subject to parse out objects that are superimposed over each other. The standard version yielded no significant results while the altered version, which required the patients to identify the orientation of the objects, revealed the MTL lesion group to perform significantly worse in the task compared to the other two groups. These results suggest that performance in the overlapping figures task is reliant on structures of the MTL.

The final experiment conducted was an object identification task. The subjects were asked to name an object comprised of an outline of fragmented lines. In the first part of the study, differing compositions of fragmented lines were presented of the same object until the subject was able to identify it or until the fifth outline was presented. The second part of the experiment utilized the same type of fragmented outlines of objects but presented four of them simultaneously rather than presenting them sequentially. The results demonstrated that the MTL
lesion patients were unable to capitalize on the extra information provided about the object either with sequentially presented fragmented lines or with the simultaneous presentations of the lines. The frontal lobe lesion group was able to perform at normal levels, revealing the MTL’s possible contribution to object identification in the presence of complementary information.

All of these experiments provide evidence that the MTL is implicated in playing a fundamental part in online reasoning, object discrimination and identification. Most importantly, these results display that the MTL is not purely restricted to tasks that involve memory, but can also aide in complex reasoning. The study conducted by Warren et al. in particular was a key supporting point for the piloting of the study that I am currently working with Dr. Rissman and Mr. Westphal to analyze.

It has been shown that the MTL is involved in multiple facets of cognition and is not exclusive to declarative memory. Anatomically, the MTL, more specifically the hippocampus, has many interconnecting networks throughout the brain (Shohamy & Browne, 2013). This anatomical interconnectedness may be why the MTL has been shown to have a wide array of interactions with differing forms of cognition. Shohamy and Browne propose two theories as to how the MTL interacts with other brain regions to give rise to these wide amounts of interactions, by memory modulation, or by adaptive function. The memory modulation hypothesis posits that the hippocampus and surrounding structures are exclusive to the domain of memory information and modulate actions by utilizing this information. The more general adaptive function hypothesis states that the MTL is not exclusive to memory but is activated through a stream of cognitive processing for its unique computations. The structures are not domain-specific to memory but act as computational functions, being called when needed, taking in inputs and sending outputs.
References


Categorical Generalization in Toddlers versus Adults

Elham Zargarkalimi

Forming categories of distinct objects is thought to be an important aspect in human cognition (Gureckis & Love, 2004). Moreover a crucial step in categorizing different objects for later recognition of novel instances is generalization. This topic has been the subject of many recent studies of cognitive psychologist and has attracted much attention, especially since it is a critical stage in cognitive development in the early years. Therefore, researches have tried to find a mechanism by which generalization is made easier and more efficient for toddlers. Generalization is done by first forming the proper mental abstraction, which takes experiencing many different instances of a category and requires a considerable amount of time (Son, Smith, & Goldstone, 2008). However, researchers have tried to find a more efficient method for generalization by presenting the toddlers with the simplified or already abstracted instances of a category. The researchers refer to this idea as a short-cut in the training process, which directly presents the proper abstraction, leading the toddler to form the proper generalization and therefore have an easier time for later identification of novel instances of the category.

This hypothesis was examined by researchers, Son, Smith and Goldstone in 2008 in four different experiments. These four experiments were targeted to test this idea on young children, generally 15 to 28 months old. In each of these experiments the children were taught categorization of novel objects with either simple or complex training exemplars. In the first experiment there were two different conditions, which consisted of either having the children start with complex training and then go through the complex generalization test case, or having
the children start with simple training and perform a simple object generalization test. Following this study, it was found that the children who were trained and did the generalization test both by simple objects were more successful in the process of generalization to identify a new object. This experiment demonstrated that simplified training instances or simplified generalization tasks lead to more efficient generalization process than complex instances for near transfer. Therefore, it was needed to further examine if the reason behind this advantage was in fact the method of presentation. Moreover, the researchers wanted to know whether this idea was advantageous for further transfers, such as identifying novel objects whose shapes are not identical to the training exemplars. As a result the second experiment was designed with two additional conditions, simple-to-complex, which consisted of training with simple objects and a generalization task with complex objects and complex-to-simple condition, with complex training and simplified generalization task. The result of this experiment showed that participants in the simple-to-complex conditions were in average more successful in generalizing by shape similarity, when presented with complex instances. Accordingly, this experiment supported that simplified learning instances are important in facilitating generalization.

Experiment 3 was quite similar to experiment 1 and 2, only with additional details embedded in the generalization tasks’ stimuli. The result of this experiment demonstrated that the generalization skills of children depend on the amount of their experience with category labels and the children’s vocabulary skills. It was supported that children with greater experience with categories had a better sense of generalization.

The authors had predicted that if simplicity promoted transfer, then presenting an object with fewer features should lead to easier generalization. Therefore, a final experiment was conducted in order to increase the details in the stimuli by manipulating the number of their
features. The results demonstrated that objects with fewer features are similar to simplified exemplars; since it helps children pay attention to only the relevant features. As a result, it was again supported that processing objects with fewer features or basically simpler learning category members advance generalization process.

In summary, it was found that those toddlers who were presented with simpler and more abstract instances while learning categorical labels, were more successful at identifying a novel instance, compared to those trained with complex category members. (Son, Smith, & Goldstone, 2008). In order to advance this finding, this training method was studied on adults’ categorical generalization ability by Khanh-Phuong Thai and Ji Yun Son in 2013. The result of their study also showed the advantage of simpler learning instances for the process of generalization, thus suggesting that the identification of a new category member in adults, is similar to that in children (Thai & Son, 2013). Toddlers and adults who are initially presented with simple instances are more successful in categorical generalization, due to the idea of simple advantage. The simple advantage, according to Thai and Son, is the asymmetry of transfer from simple versus from complex learning exemplars.

Examining categorical generalization by the mentioned researchers was done by recruiting college undergraduates for two separate experiments. In the first experiment, the participants were asked to study a series of Chinese characters with their corresponding English meanings, followed by performing a matching memory task and the generalization task. During the generalization task, the participants were presented with the English meaning of the characters that they had been trained with, and were asked to choose the corresponding Chinese character from four different options. However, the presented choices of Chinese characters were either in the opposite script or the same script as what they had studied. In other words, half of
the participants were randomly assigned to study the Simplified Chinese characters, while performing the generalization task, with the corresponding Traditional Chinese characters, which are more complex in shape and structure. The other half of the participants were randomly assigned to be trained with the Traditional script and do the generalization tasks using the Simplified characters. These two conditions, Simple-first and Traditional-first are similar to the groups of the earlier experiment, simple-to-complex and complex-to-simple, respectively. The results of this experiment demonstrated that participants who were trained with the Simplified Chinese characters in the beginning, were generally more accurate when presented with the Traditional characters, as their training enabled them with a better generalization. However, the amount of time spent for studying each character with its corresponding English meaning was not controlled in this experiment, thus conducting the second one.

In order to assure that the simple advantage was not dependent on the amount of time of exposure to the characters during the learning process, the second experiment was run similarly but with addition of a time factor and removing the training procedure. This experiment was meant to examine the differences in generalization only based on perceptual similarities of the characters when the participant is first exposed to either a Simplified or Traditional Chinese character. In order to control the amount of exposure to each character between all participants, they were presented with the character for 2 seconds and then shown 4 different choices, from which the participants were asked to choose the best match. The tests consisted of two forms, the Generalization tests (simple-to-complex or complex-to-simple), in which the choices were from the opposite script of the initial character, and the Exact Match trials (simple-to-simple, complex-to-complex), in which the participants matched the presented character to the identical character. Analyzing the results of this experiment showed that the Simple-first conditions had a
significantly higher performance in the generalization trials, compared to the Traditional-first condition.

These findings, similar to the first experiment, support that exposure to Simplified characters advances generalization and transfer for Traditional Chinese characters more. This further supports that the simple advantage indeed promotes generalization for Chinese characters even without extended exposure. In other words, the generalization was promoted when the initial training presented the Simplified script, similar to the experiment examining categorical generalization in young children. Therefore, this cognitive process has similar mechanisms in both toddlers and adults, as the simplified instances of a category provides one only with the useful and abstract exemplar of a category. This makes generalization and therefore far transfer more efficient. The authors argue that presentation of simplified instances help making transfer or identifying a novel instance easier, by allowing one to organize his/her perception and by attending only to the important features, which lead to better generalization (Thai & Son, 2013).

In order to examine how the duration of exposure to each character affects generalization, a new study was conducted by Thai during Fall 2013. This experiment recruited 25 undergraduates from UCLA psychology department, who participated for course credit. This study, referred to as experiment 3, was similar to experiment 2, but instead of a two-second representation for each character, the exposure time was varied in each condition. Therefore the four (two-by-two) conditions stayed the same from the previous experiment, which consist of simple-to-simple, simple-to-complex, complex-to-simple, and complex-to-complex. In each trial of this experiment, after presenting the participant with a fixation cross for 0.5 seconds, a character was presented for either 0.5 seconds or 6 seconds in a random order. Each trial was then followed by the generalization task, which required the participant to choose the best match
for the presented character, from four choices. The trials of this experiment similar to the previous experiment were all performed on a computer using E-Prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA, USA). A sample sequence of the procedure for the Exact Match test (complex-to-complex) and a Generalization test (complex-to-simple) is demonstrated in Figure 1.

This experiment is still in progress, and the result of this experiment is pending at this time, as it needs further analysis. However, it is hypothesized that similar to previous studies, presentation of the Simplified script, will enhance the generalization and transfer for identifying the unrelated script. Moreover, it is predicted that greater exposure to the training exemplars would enhance the generalization task and increase the rate of accuracy. The results of these studies can be implied to learn more about such cognitive processes and the mechanism by which one generalizes experienced instances of a category and uses prior knowledge to identify a new instance.
References


Figure 1. (a) Exact Match test (b) Generalization test procedure of the Complex-to-simple condition. Figure extracted from previous studies by Thai and Son (2013).
Identifying Composer Style Using a Perceptual Learning Module

Erika Der Sarkissian

Perceptual learning indicates an improvement in the ability to pick up information through experience or practice (Gibson, 1969). This type of learning includes discovery effects and fluency effects. Discovery effects reveal changes in response-selecting relevant information as well as suppressing irrelevant information. Fluency effects lead to faster problem solving with less effort (Kellman, 2002). The effects of perceptual learning are studied at the UCLA Human Perception Laboratory on various tasks, including solving algebra problems, identifying skin symptoms, and labeling ECG graphs. The designs of most of these studies include a pretest, followed by a Perceptual Learning Module (PLM), which gives the participants feedback on their responses and records their accuracy and response time. One day later, the participants are given a posttest in order to see the effect of the PLM. With the present study, we are interested to see if the PLM would be effective with auditory stimuli. Unlike visual stimuli where the information is presented all at once, auditory stimuli are presented according to time. With the current study, we would like to see if participants could learn to identify the styles of composers through the completion of a Perceptual Learning Module.

In 2005, and Peretz did a study on musical style discrimination based on historical periods. For this study, they tested the Historical Distance Effect, which states that styles should be judged as more similar if they are closer in history, than when they are more distant (Bella & Peretz, 2005). This type of styles discrimination does not actually require the participant to be
able to label a style, but rather assess whether two musical excerpts are similar in style. Therefore, this type of assessment may occur without explicit stylistic knowledge (Bella & Peretz, 2005). For this study, Bella and Peretz (2005) used the four major tonal styles of Western classical music: Baroque, Classicism, Romanticism, and Post-romanticism. Throughout these three centuries of music history, styles properties progressively evolved, as composers deviated from the principles of tonality and created progressively more rhythmically variable music (Bella & Peretz, 2005).

Three groups of participants were used in this study: Western musicians, Western non-musicians, and non-Western non-musicians (Bella & Peretz, 2005). There were twelve participants in each group. The Western musicians group was comprised of students at the Faculty of music of the University of Montreal. The non-musicians were university students without formal musical training, and the non-Western students were raised in China and had grown up listening to Chinese music. They had little exposure to Western music. The participants were given sixteen excerpts imitating the styles from the different time periods. They were instructed to rate the familiarity of each excerpt on a seven-point scale. They were then presented with pairs of these excerpts and given ten seconds to rate the stylistic similarity between them on a seven-point scale.

The results of this study indicated that Western musicians gave a higher familiarity rating to the excerpts than the other groups (Bella & Peretz, 2005). However, all subjects showed low-familiarity with the excerpts. On the similarity test, the Western musicians were also more affected by the Historical Distance Effect than the Western non-musicians, who showed a larger effect than the non-Western non-musicians. This result is expected since the Western musicians have had the most exposure to Western music, and non-Western non-musicians have had the
least. However, the results from all groups presented a significant influence by the Historical Distance Effect. This result indicates that even with little to no exposure to Western classical music, participants were sensitive to the differences between the styles of different time periods.

Unlike Bella and Peretz’s study, our study does not only require participants to be able to differentiate between styles of time periods, but also be able to identify which composer the style belongs to. In his article, “The Development of Sensitivity to Artistic Styles”, Gardner (1971) compares the effectiveness of the completion test, the pairing test, and the instances test in assessing a participant’s sensitivity to style. The completion test comprises of exposing the participant to one section of a work and having him select another section of the same work from an array of options (Gardner, 1971). This method uses the act of comparing between different sections in order to determine sensitivity. Gardner argues that although this test works well with visual art, it is not as effective with music. This difference is due to the fact that visual art is perceived all at once, while musical style requires a longer exposure to give the participant any confidence in matching another section to the first (Gardner, 1971).

The pairing test requires participants to pair different sections of the same work to each other (Gardner, 1971). Gardner also argues that this method is ineffective with music because it causes the participant to rely too much on tiny, trivial details of the piece, such as a particular chord, rather than the overall style. In an instances test, the participant is given instances of a particular style and asked to choose another instance from an array of other works (Gardner, 1971). Unlike the previous tests discussed, the instances test involves a range of an artist’s work and presents more opportunity for the extrapolation of consistent features rather than specific details (Gardner, 1971).
In 2002, Crump ran an experiment to test the ability of participants to discriminate musical style when given five seconds of musical information. He chose to use only five second samples of music because they were long enough for the participants to be able to hold information about style, but small enough to take memory limitations into account. Twenty undergraduate psychology students from the University of Lethbridge participated in this study. The five-second clips were chosen from works by J.S. Bach and W.A. Mozart, who both had highly structured music and were stylistically distinct. The participants received a training phase, where they were trained on either fifty Bach samples or fifty Mozart samples. They were then given a test phase where they were presented with fifty new Bach samples and fifty new Mozart samples and were tested on their ability to recognize whether each new sample was in the same style as the samples in the training or not. Participants also ranked how confident they were on each answer.

The performance of the participants was measured by using hit and false alarm rates (Crump, 2002). Hits were counted as ‘yes’ responses that were correct. False alarms were counted as ‘yes’ responses that were incorrect. Therefore, if a participant received training on Bach samples, and correctly recognized a Bach sample on the testing phase, their response was a hit. If they were given a Mozart sample and incorrectly recognized it as a Bach sample, their response was a false alarm. However, using this type of measurement can lead to issues because a high hit rate may just mean that the participant reported ‘yes’ more often than ‘no’. For this reason, Crump used signal detection statistics to calculate the sensitivity and the bias of each participant. The A’ statistic ranged from 0.5 to 1 and measured discrimination ability. A score of 0.5 indicated no discrimination ability and a score of 1 indicated perfect discrimination ability. The b” statistic ranged from -1 to 1 and measured bias, or how often the participant said ‘yes’. A
score of 0 for the b” statistic meant that there was no bias for reporting ‘yes’. The results showed a significant effect of performance, which means there were more hits recorded than false alarms. The effect of training with one composer over the other did not reveal a significant effect and there was no significant interaction found between trained composer and performance. These results indicate that the participants were able to successfully discriminate between Bach and Mozart with only five seconds of audio.

Our study will test the effectiveness of using a Perceptual Learning Module to train participants on composer style. The participants will be students attending the University of California, Los Angeles and will be given credit for their participation. The PLM will use clips of music by two composers from the Baroque period, two composers from the Classical period, and two composers from the Romantic period. During the PLM, the participants will be given a clip of music and a list of composer names. They will then be asked to indicate the composer they believe composed the music. If they choose correctly, the module will let them know and go on to the next clip. If they choose incorrectly, the module will correct them. The participants will also be given a pretest before the PLM and a posttest after. These tests will include unheard music clips from the same composers used in the PLM, as well as music clips from unheard composers and unheard time periods such as post-Romanticism. They will not give any feedback to the participant. A survey will also be included asking participants various questions about their previous experience with classical music.

The music used in this study was found at the Music Library at UCLA. Songs were cut into fifteen to twenty second clips, depending on the amount of musical information in each clip. Each clip was described based on its pianist, tempo, tonality (major/minor), form (ex. Sonata),
and overall mood or movement. The clips that were used for the study were then chosen out of all the described clips, making sure that all of these factors were varied among the sample.

We hope to find that participants will be able to identify the style of each composer after going through the Perceptual Learning Module. If not, we would be interested in whether participants will be able to differentiate between the styles of different time periods. We have discussed the ways in which written descriptions of the styles of composers or time periods would affect the listener’s training. The study design is still being adjusted to test for these effects with the best method possible.
References


**Virtual Reality- Bridging the Gap**

**James Mutter**

Aims of this paper: This paper focuses on the background definitions as well as design and implementations of the virtual reality literature within several different contexts and practices. Common errors that are observed repetitively are pointed out as well as current pedagogical activity will be explored with the hope of establishing a base to bridge the gap for future research within this field. This revealed gap will hopefully establish its applications to research universities using it for VR’s potential for learning capacities.

Virtual Reality (VR) is a term that applies computer-simulated environments to mimic a lifelike experience, that allows or compels the user to have a sense of being present in an environment other than the one they are actually in, and to interact with that environment (Ming, Ruan, & Gao, 2012). Common practices for researchers are to give a sense of immersion for their participants within the VR environment. Immersion is described as the augmented sense of presence, through virtual embodiment in the form of an avatar with extensive modes of active roles that can be achieved. This amplified presents can impact the affective, empathic and motivational aspects of the experience (Warburton, 2009).

There are several educational practitioners and researchers that believe that virtual reality equipment can support education with strong benefit. There are several modalities to use virtual reality in different contexts such as the ability to facilitate constructivist learning activities for language learning or to focus on its use toward social interaction (Youngblut, 1998). The use of 3D virtual world being integrated into language learning and teaching is still in its infancy. While still in its early years studies have been consistently demonstrating its prospective use for a teaching environment. The 3D interactive technology for prospective language learning through
the use of Moodle, OpenSim and Second life offers an engaging and dynamic framework for language teaching and potential learning for other applications (Chen & Su, 2011). Research institutions and universities are continuously investing heavily in developing and acquiring virtual worlds (University of Delaware, University of Cincinnati, University of Washington, University of Wisconsin, University of the West of England) as examples. However, even with the heavy investment in exploring the potential value of these worlds, the amount of dedication towards cognition is limited. There also remain common practices in the research designs that alienate conclusive results from the current research mythology.

While there is a growing body of research pertaining to the use of VR for educational purposes, there are much of the same common problems that are portrayed in other areas of literature. For example, a study by Bailenson et al. (2008) claimed that there was a significant effect of gender, concluding that men perform better than women. However, they did not look at virtual reality experience beforehand which is a prevailing problem amongst a great majority of the literature. This conclusion may have occurred because, in general, women tend to have a considerable less amount of experience using video game or video game like interfaces (Jenkins, 1998). For this reason they may consequently felt less comfortable overall using the virtual reality environment within the context of the experimental setting. Furthermore, this study and several of the other research pertaining to this field could be susceptible to a number of potential sources of confounding variables. Similar to the finding by Jenkins, researchers Rose and Billinghurst (1995) claim that their own study could be vulnerable to subjects' attitudes and prejudices toward technology that is not looked at before implementation of the experimental design. A study by Dalgarno and Lee (2010) reviewed several of the applications of 3-D virtual realities. The primary goal of their research was trying to identify the sequence of learning
affordances of these types of VR contexts. These affordances include the facilitation of tasks that lead to increased chances for: experiential learning, motivation/engagement, spatial knowledge and contextualization of learning (Dalgarno & Lee, 2010). While these affordances seem noteworthy, other research has pointed out that there is a threshold to reach in order for this potential to truly occur. Researchers Ming, Ruan, and Gao (2012) pointed out that it took around two hours of using their VR technology for the learners to generally reflect that the virtual reality system was proficient and heightened their understanding of mandarin learning within the realm of pronunciations and concepts. This is another frequent problem that abides in the existent research, much of the time the subjects aren’t given an adequate adjustment phase to become immersed in the VR environment. Immersion an important role that is overlooked but also unforeseen are the memory techniques used presently.

A study by Mikropoulos and Natsis (2011) conducted a 10 year overall exposure to the measures of virtual reality and ultimately concluded that constructivism is the modern theoretical model that almost all of their 53 reviewed studies established. This approached is brought to light with much of the same results within the context of learning. Bjork, Dunlosky, and Kornell (2013) found that students have better retention of vocabulary when they correctly recall the word three times or more compared to only one time. The current literature is missing this key aspect in the form of pedagogical learning. For instance Rose and Billinghurst (1995) conducted their Japanese VR experiment with subjects allowed to study what they wanted with only 3 tests total at set intervals to test their understanding. It is noted by Bjork et al. (2013) that people in general have faulty models of how they learn material and therefor their natural heuristics lead to illusions that are counterintuitive. This makes evident that Rose and Billinghurst (1995) results could be subsequently undermining what the potential results could have been. Similarly, Ming,
Ruan, and Gao (2012) experiment regarding Mandarin learning has subjects go through modules throughout the game. Each module they are asked questions regarding the current scene and objects. If subjects get the correct response, they go on to learn new scenarios and cannot practice those old phrases.

Another key point of classical learning techniques that is often overlooked is the testing phases lead to enhanced learning more than the same amount of time studying without being tested (Bjork, Dunlosky, & Kornell, 2013). Using the Rose and Billinghurst (1995) study as an example, they only had one test at the end of session two for their second dependent measure. This further amplifies the research designs flaw for implementation of memory. Additionally, Bjork et al. (2006) states that studying the material after you responded incorrectly for 5 seconds is better than studying only the correct answer for 13 seconds with no testing phase. This simple but pertinent piece is often neglected in a noteworthy amount of the literature of memory retention because retrieval attempts improve future learning (Kornell, Hays, & Bjork, 2009). A growing body of literature has done a vital amount of work with several different contexts that lack one or more of these pieces of information, such as learning languages (Chen & Su, 2011; Ming, Ruan, & Gao, 2012; Rose & Billinghurst, 1995), mimicking real world contexts (Moskaliuk, Bertram, & Cress, 2013; Seymour et al., 2002), episodic memory (Plancher, Barra, Orriols, & Piolino, 2012; Sauzéon, et al., 2012), various theoretical and methodological assessments of VR (Henderson, Henderson, Grant, & Huang, 2010), learning as immersive experiences (Dalgarno & Lee, 2010) and simple experiments having to do with the best approach to learning in VR (Bailenson et al., 2008). Additionally, much work on memory retention within the learning domains of research have persistently gone on to be neglected.
While the growing body of evidence is heightening for the implications and possible uses of VR, it is partly narrow in its attempt to explore different domains especially with that of spaced learning and recall. Bjork, Dunlosky, and Kornell (2013) indicated that over 90% of college students performed better in the context of spacing rather than mass practice. This finding is absent in much of the literature. For instance, a study by Plancher, Barra, Orriols, and Piolino (2012) administered subjects three episodic memory tests after the study but all experimental tests were given within 15 minutes of the mass practice. Another study used to conduct episodic memory by Burgess, Maguire, and O'Keefe (2002) conducted the memory portion after a 10 minute delay. Rose and Billinghurst (1995) tested the effects of retention of learning a language after two sessions spaced one week apart. While the overall gist of the study was incapable of producing conclusive results that are of statistical significance of VR being superior to more traditional methods of textbook learning, it did take steps in order to present long term retention rates of information. Furthermore, they did test the VR context versus a real world context that mimicked the virtual world. However it did have several problems with it, including a lack of background information regarding the subjects use of past video game interfaces which is demonstrated to be a significant factor of using VR, an inconsistency in the amount of people in each condition and several flawed traditional learning techniques that are mentioned above. Without any of the experimental tests being given or subjects being exposed to the stimuli over an extended period of time, results with VR learning so far are quite inconclusive in nature.

The 3D VR technology is established as a concrete context to practice real life scenarios, language development and understanding the fundamentals of complex theories. While each researcher is focusing on developing their thesis within the context on their work, each piece
contributes to the fulfillment of the aforementioned need to progress linearly. This work contributes to a seemingly new research area to understand the parameters that establish the possible internal validity within this field. This validity needs to be understood in order to bridge the gap to the potential learning possibilities that can offer unique features in the VR environment. At this time, the current progress in the design of the environments and extrapolation of possible learning advantages are very much trial and error (Dalgarno & Lee, 2010). This may be because much of this work is driven by intuitions that are seeking very different interpretations of what is attainable. It has been demonstrated that much more thorough investigation is needed to tie the bridge between the education circle and virtual realities closer together in order to truly leverage the potential that VR offers. Only then, when frameworks for VR environments that harvest the use of classical memory techniques, spaced practice and accept the biases that people have regarding technology can facilitation of active learning benefits take place. Without this, the retention of information for student learning is just a guess put on by institutions.
References


Jenkins, H (1998). *Chess For Girls?: Feminism and Computer Games*


Perceptual Learning in Classical Music

Joselyn Ho

As perceptual psychologist Eleanor J. Gibson describes in *Principles of Perceptual Learning and Development* (1969), perceptual learning is the improvement in one’s ability to pick up new information with experience or practice. While the significance of explicit knowledge in learning material has been emphasized, perceptual learning also plays a major role in one’s ability to master information. Perceptual learning involves changes in how an individual responds to or processes the presented stimuli. A person with a lot of experience in a subject, when compared to a person without much experience with the subject, may select the presented information differently by interpreting the information more speedily and efficiently or noticing patterns that the other person might not immediately recognize. An expert of electrocardiogram readings, for instance, would identify patterns in the graphs very quickly and easily because of their training, while a novice would take longer and require more effort to distinguish and interpret the relevant features of the graphs.

Several studies have incorporated perceptual learning modules (PLMs) to investigate the effects and extent of perceptual learning, such as in Kellman et al.’s (2009) study of perceptual learning in mathematics. A PLM consists of computer-based interactive trials that present the participant with classification tasks and give feedback on his or her responses. In one PLM of Kellman et al.’s (2009) study, for each trial, subjects were presented with an equation and four possible algebraic transformations of the equation. When the subjects selected a transformation, they were given feedback on whether their answer choice was correct, along with the time taken
to respond to the trial. The anticipated effect of PLM training is a person’s statistically significant improvement in accuracy and response time to the presented information. Perceptual learning studies usually involve a pretest and posttest to assess the participant’s improvement. The pretest and posttest have the same format as the PLM but do not give the participant feedback. In Kellman et al.’s (2009) study, participants in the experimental group were administered the pretest and PLM on one day, the posttest on the next day, and a delayed posttest a week later. The results for the algebraic transformations revealed that participants improved in accuracy and response time for the posttest and delayed posttest, supporting the positive training effects of the PLM for that study.

Past PLM studies have mainly focused on visual presentations. The proposed study aims to investigate whether the PLM experimental setup can be applied to auditory presentations, specifically classical music. The major difference between visual and auditory stimuli is that sounds are presented temporally and do not remain static. Visual information is presented all at once, while auditory information is dynamic over the time duration. The proposed study focuses on the classification of classical music and was inspired by studies such as Bella and Peretz’s (2005) and Gjerdingen and Perrot’s (2008). The results of Bella and Peretz’s (2005) study suggest that people, regardless of regional background, are able to recognize differences in musical styles. Gjerdingen and Perrot’s (2008) study was inspired by the metaphor of “scanning the dial” which refers to the act of quickly switching among radio stations to determine a preferred style of music. The results of the study suggest that people are able to categorize song clips by genre in very short time durations, regardless of musical training. The outcomes of these studies seem to indicate that music contains features that people can perceive and classify into distinct types. The proposed study aims to investigate this hypothesis and explore whether a
PLM framework can be used to train for the categorization of song clips by composers in the classical genre.

Aims of the Proposed Study

The proposed study will analyze the effects of a PLM on people’s abilities to match classical music song clips to respective composers. The study will also examine the effect of providing descriptions about the composers’ styles.

Method

Participants

Participants will be UCLA students receiving course credit for partaking in UCLA psychology studies.

Design

The control group design will be composed of a pretest, PLM with descriptions of the composers’ styles, and a posttest. The experimental group design will be composed of a pretest, PLM without descriptions, and a posttest. The study will be conducted with a computer program which will be accessible online.

The PLM will be split into multiple sessions and will span over several days. We decided to divide the PLM because pilot testing indicated that the music PLM would take participants longer to complete than with visual PLMs, since the songs are presented temporally. We will request participants to train with the PLM for half an hour each day, for two weeks. Participants will begin training with the PLM right after they complete the pretest. For each trial of the PLM, participants will hear a song clip of about fifteen seconds and match the song to its composer from a list of six composers. Reaction time will not be considered, since the songs are played according to time. Participants will be able to listen to the full duration of the song clip before
each trial times out. Feedback will be given for each response by displaying “Correct” or “Incorrect” on the screen, and the song clip will continue to play until the end of the clip.

![Incorrect Response Example](image)

Fig. 1. Example of an incorrect response in the PLM. This screenshot is taken from a past module. The module in the proposed study will have six composers to choose from, rather than four.

regardless of when the participant selects an answer. When a participant selects an incorrect response, the correct answer will be highlighted. An example of an incorrect response is shown in Fig. 1. After every twenty trials, participants will be presented with a feedback summary of their progress.

The pretest and posttest will have the same format as the PLM but will not include feedback. The pretest and posttest will each have 15 trials.

**Materials**

The music clips in the PLM will be randomly drawn from a pool of 480 clips, with eighty clips per composer. The composers will include Bach, Handel, Mozart, Haydn, Chopin, and Schumann. These composers are selected from the music periods of Baroque (1600-1750),
Classical (1750-1820), and Romantic (1820-1900), with two composers representing each period. To control for instrumentation, all songs will be in the form of piano music. The clips will be cut from songs collected from the UCLA music library. To ensure that the clips for each composer are varied and not concentrated on a specific style, we will first organize the clips based on general features such as tempo, tonality, and mood, and then we will select a balanced number of clips from each category to be included in the PLM. We will make sure to select songs that are not popular, to prevent familiarity from influencing the results.

The descriptions included with the experimental group’s PLM will consist of brief and general explanations of each composer’s style. The information, obtained from music websites, will include the background and common, identifiable trends of the composer. We will use simple terminology that can be understood by people without musical training. We will also include links to videos of works that exemplify the composers’ styles. We will emphasize to the participants that the information in the descriptions is meant to be used as a reference and will not necessarily apply to every song clip that is presented. The descriptions will be presented in the form of a Word document, and the participants of the experimental group will be able to refer to the descriptions at any point during the PLM.

Procedure

Participants will take a pretest and then begin their first thirty minutes of PLM training. They will be asked to continue their daily thirty minutes of PLM training on their own. Participants will be given directions for how to access the PLM on their own computers and will be instructed to train with the PLM in a quiet environment to simulate controlled settings. The participants will return to the laboratory to complete the posttest following their two weeks of PLM training.
Following each PLM session, participants will be requested to complete a brief survey. The survey will ask the participants to rate their alertness during the PLM and the hours of sleep acquired the night before. A separate survey will be administered after the posttest and will ask about the participant’s music listening habits and music training.

Hypotheses

We hypothesize that music contains distinct patterns that people can perceive and use to categorize based on genre. Therefore, we believe the PLM will have a positive effect on people’s abilities to categorize classical music to composers, so both the experimental and control groups will show improvement in their pretest and posttest scores. The purpose of including descriptions is to investigate whether providing the participants with written descriptions will enhance or diminish the perceptual learning effects. We want to see if providing the descriptions will help people who have not had musical training identify the composers’ styles more easily. We hypothesize that the descriptions will enhance the learning process, so that the experimental group will show more improvement over the control group.

Conclusion

Through this experiment, we hope to understand whether perceptual learning can apply to dynamic and temporal auditory information, and whether people can be trained through a PLM to identify classical composers’ styles. We also hope to understand the effects of providing written descriptions. We hope that conducting this study will enrich our knowledge about the extent of perceptual learning and the process of musical training.
References


Analogical Problem Solving In Animation Source Training

Kevin L. Ruiz

Abstract

In a pilot experiment, we examined how well people reason analogically when given an animation as the source analog. Giving participants the tumor problem after being presented with three different animations did this. After each animation participants were given questions that allowed them to form an understanding of the animation they watched. Questions also asked how each animation was different from each other. Results showed a significant rate of transfers. That is, the participants solved the target analog, tumor problem, after being presented with the source analog, animations.
Analogical reasoning is a process by obtaining a solution between two concepts that share similar features. Only when concepts share similar features can they be considered analogous. Analogies allow people to see commonalties between concepts that seem different than each other by allowing them to see the underlying relationships (Day & Goldstone, 2001).

In the current literature, successful use of analogies happens when people solve target analogs after studying the source analog. The source analog is the original concept people learn first in an experiment. Holyoak (2012) noted that the source analog is “better understood” because the participant can have previous knowledge about functional relationships within the source analog. That is, understanding certain aspects or feature within the source analog has “causal, explanatory, or logical connections to other aspects” (Holyoak. 2012).

The target analog is the novel problem participants need to. People have shown to successfully use analogical reasoning when transfer to the target analog is shown after being presented with a source analog (Gick & Holyoak, 1980; Hummel & Holyoak, 2001; Day & Goldstone, 2011).

Analogies can be a powerful tool because it allows people to create a map that will allow people to transfer it to other concepts (Hummel & Holyoak, 2001; Day & Goldstone, 2011; Holyoak, 2012). Mapping is the process of identifying the underlying process and making a connection an element from one analog to another (Emig, 2011). When presented with a target analog, people will create a mapping of the source analog after recognizing the similarities of the two analogs (Emig, 2011; Holyoak, 2012).

Classic analogical reasoning studies use the tumor problem to show the success of analogical transfer (Duncker, 1945; Gick & Holyoak, 1980; Hummel & Holyoak, 2001;
Holyoak, 2012). The tumor problem is a scenario in which the participant is asked to help a fictional doctor cure a tumor with a tumor-killing ray. Participants are told that they can manipulate the amount of rays and the intensity of the rays. However, the ray also affects the healthy tissues around it. The goal of the problem is to kill the tumor without damaging the tissue surrounding it. Duncker (1945) originally called this the “radiation problem” but for the sake of this paper we will call it the tumor problem. The cover story can be found in appendix A. This tumor problem serves as the target analog.

In an experiment done by Gick & Holyoak (1980), analogical reasoning was tested using the success of transferring the schema from the source analog to the tumor problem. In this experiment, the source analog was given through a story Gick & Holyoak (1980) called the military problem. In this military problem a general was to attack and conquer a fortress by sending his army. The general could only capture the fortress if he sent his entire army through two options. He can attempt to conquer the fortress by sending down his army down one path or by dividing his army into small groups and sending them down through multiple paths. However, the general realized that if he went down the former option, his army would thin out and be defeated. The best choice, then, was to separate his army and go down multiple roads to conquer the fortress.

Participants who read the military problem were to use the same strategy to defeat the tumor in the target problem (Gick & Holyoak, 1980). Results showed that participants were not able to solve the tumor after reading the military problem. Only 30 percent of the participants solved it without a hint (Gick & Holyoak, 1980). After a hint – which indicated that the source and tumor analog were similar – analogical transfer improved and 75 percent of the participants were able to solve the tumor problem.
Gick & Holyoak (1980) concluded that people could only transfer a schema to different analogs when participants noticed the similarities. However, noticing the similarities posed an important question for analogical studies: which experimental paradigm better illustrates the function and transfer of analogies?

A notable variation would be an experiment done by Hummel & Holyoak (2001) in which they used static and dynamic diagrams as the source analog. The diagrams had objects converging into a circle either by a single stream or multiple streams surrounding the circle. The objects were either arrows pointing or black blocks moving towards the circle. For static diagrams objects were just in the streams by the circle while in the dynamic diagram the objects were moving towards the circle. After being presented with the source analog the participants were then presented with the tumor problem.

Hummel & Holyoak (2001) found that for the group of participants that received the static diagram with the arrows, only 15 percent solved the tumor problem before the hint. After the hint 55 percent of the participants solved the tumor problem. Much like the Gick & Holyoak (1980) study, a hint was needed before they were able to solve the tumor problem. In the group that received the dynamic diagram with the arrows, 55 percent participants were able to solve the tumor problem before the hint, while an additional 35 percent solved it after a hint. For the static diagrams with the blocks only 5 percent of the participants solved the tumor problem before the hint and 30 percent after the hint. For the dynamic diagrams with the blocks 30 percent of the participants solve the tumor before the hint while an additional 40 solved it after the hint. This study showed that compared to the static diagrams, noticing the similarities was better in the diagrams that were animated (Gick & Holyoak, 1980).
The current study adopts a similar method as the Hummel & Holyoak (2001) study by using animations as a source analog for the tumor problem. However, in this study we situated the animations into a context by creating a cover story. The purpose of the study is to show if people solve the tumor problem when the source analogs are animations. We hypothesize that analogical transfer will be successful. That is, participants will solve the tumor problem after being presented with animations.

**Method**

**Participants**

Eleven undergraduate students (10 women and 1 man) enrolled in classes at the University of California, Los Angeles (UCLA) participated in the experiment as part of a course requirement.

**Design**

The source analogs were three different animations. In each animation the laser(s) had to destroy circle that was situated in a circular barrier by shooting at it. When the laser(s) shot at the circle the barrier also gets damaged. The point of the laser(s) was to destroy the circle without destroying the barrier first. These animations had a cover story, which can be found in appendix B.

In the first animation, one laser was shooting at the circle from the right side. In this animation the laser failed to destroy the circle because the barrier around it was destroyed first. In the second animation, one laser was also shooting from the right side at the circle, however the blast of the laser was a lot smaller than the one shown in the first animation. Much like the
first animation though, the laser failed to destroy the circle because the barrier surrounding it was destroyed first. Lastly, in the third animation there were multiple lasers surrounding the circle shooting at blast at a low intensity as in the second animation. Different than the first two animations, the lasers destroy the circle before destroying the surrounding barrier.

The tumor problem was adopted and served as the target problem in this study. As mentioned above, the tumor problem placed the participant into a situation that asked them to help a doctor destroy a tumor without destroying the healthy tissue surrounding it. This was presented in a cover story after the source analog. The cover study can be found below under procedure sub-section.

After each animation the participants were given a questionnaire that asked a couple of questions regarding the animations they watched. After the third questionnaire participants were given the tumor problem and were asked to solve it. If they solved it, then received the last page. The participants that did not solve the tumor problem were asked reread the tumor problem and try their best to solve it. The last page was a questionnaire that asked if they have seen the tumor problem before and if they noticed the analogy between the animation and the tumor problem. The questionnaire can be found in appendix C.

Procedure

Participants were asked to sit in front of a computer in a private office space. Once the participant was situated, the experimenter asked the participant for some basic demographic information (e.g., age, gender, and major). Once the experimenter wrote the basic information, the experimenter gave the participant a brief description of the experiment. The experimenter said the participant will watch three animations and after each animation the participant will be
answering a set of questions based on the animation the participants just watched. Also, the experimenter asked to read the further instructions on the computer screen and notify the experimenter when asked. The experimenter started the program and left the room.

On the computer screen, participants were to read a set of instructions (Appendix B) that explained the animations even further. After the participant read the instructions the computer asked them to press the space bar to continue on to the first animation. After the first animation finished, the experimenter gave the participant the first hand out, which had questions 1 and 2. Once the participant finished that, the experimenter pressed the computer key for the second animation. Once the second animation finished,

Results and Discussion

Out of the eleven participants that were gathered, two had previous experience with the tumor problem so their data was not included in the analysis. Out of the remaining 9 participants 8 (88%) solved the target problem before the hint. No participants received the target problem with the hint. These results are able to confirm our hypothesis, which means that analogical transfer was successful. That is, the participants solved the tumor problem after being presented with animations.

Compared to the original findings in Gick & Holyoak (1980), there is a significantly higher amount of transfer rates than in the analog presented in their experiments. Unlike the in Gick & Holyoak (1980), we did not explain the animation to the participants. Instead we gave the participants guided questions. We did this so the participants can come up with their own understanding of the animations. We felt that giving them an explanation would be giving the answer away for the tumor problem. By having the participants come up with their own
understanding allows them to create the concept into an understandable framework. Without the explanation the participants would probably not solve the target analog. However, to conclude this we would have to run it as a control group in future experiments.

We expected the success of the transfer to happen the way it did because the target analog was presented right after the source. Also, the mappings between both analogs were explicitly similar. When asked, participants reported a clear surface level similarity between the source and target analog. The source, then, may be too similar to the target problem, which could explain the amount of spontaneous transfer.

Overall, our study provides a good base for future studies. A future study could also look into comparing animations against interactions with the animations. A group of participants would observe the laser animations while another could manipulate the amount of lasers and intensity of the laser(s). Also there would be a control group that would source analog. This would really be beneficial for studies on analogical reasoning because it can show us which of the two create a better understanding of the source analog.
References


Appendix A

Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. It is impossible to operate on the patient, but unless the tumor is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumor. If the rays reach the tumor all at once at a sufficiently high intensity, the tumor will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but they will not affect the tumor either. What type of procedure might be used to destroy the tumor with the rays, and at the same time avoid destroying the healthy tissue?

Appendix B

In the following experiment, you will observe three animations in which one or more friendly lasers are attacking an enemy circle. The enemy circle is surrounded by a friendly barrier. The object of the laser(s) is to defeat the enemy circle without critically damaging the surrounding friendly barrier.

As the enemy circle is damaged by the friendly laser, it gets smaller and gradually changes in color from white to red. As the barrier is damaged it ‘heats’ up and also turns red. Note that the friendly laser(s) cannot damage the enemy circle without also increasing the temperature of the friendly barrier. The objective of the laser(s) is to defeat the circle without allowing the friendly barrier to reach a critical temperature.

If the barrier temperature bar fills up with red, the friendly barrier has exceeded its critical temperature and the friendly laser(s) fail their objective. The objective of the laser(s) is to fill the ‘Circle Damage’ bar before the ‘Barrier Temperature’ bar fills up.

Appendix C

Q1: Why was the laser unable to defeat the enemy circle in Animation #1?
Q2: What could the laser do differently to avoid making the same mistake in the future?
Q3: How was the laser’s behavior different in Animation #2 compared to Animation #1?
Q4: Why did the laser fail in Animation #2?
Q5: How was Animation #3 different from Animation #1 and Animation #2?
Q6: Why did the lasers succeed in Animation #3?

Tumor Problem (without hint)
Tumor Problem (with hint)

Q7: Have you ever encountered the tumor problem prior to this experiment? If so, please describe the context in which you’ve previously encountered it.
Q8: Before reading the question, did you notice any similarities between the animated laser/circle problem and written tumor problem?
Spatial Navigation and Autobiographical Memory

Laryssa Storozuk

Major technological advances over the last several decades have allowed for the development of a new field of memory research utilizing functional magnetic resonance imaging. This research analyses changes in blood flow to detect brain activity. In our current study, we are utilizing fMRI scans to look at autobiographical memories, by showing participants sequences of images taken from events that occurred in their own lives. We are interested in exploring how the source of a sequence of images, (Self vs. Other: i.e. one’s own images vs. someone else’s images), and any temporal alterations made to a sequence of images (Match vs. Mismatch: i.e. intact sequences which are presented in the temporal order in which they occurred vs. “scrambled” sequences, in which the temporal order has been manipulated) affect activity in different areas of the brain, as well as to quantify our conscious abilities to detect such manipulations. Finally, we are exploring any patterns in brain activity – and changes in one’s ability to detect the Self vs. Other, and Match vs. Mismatch conditions - associated with whether a sequence had been previewed by a participant prior to viewing it in the fMRI scanner (Previewed vs. Not previewed).

Innovation plays a key role in this study because of our use of real images capturing events that occurred in participants’ lives. Participants wore cutting-edge cameras, called Autographers, around their necks for a three-week period, during which images were automatically captured up to several times a minute. These cameras, which have been marketed in the United Kingdom as “The World’s First Wearable Camera,” were designed to be worn for all of life’s adventures (Autographer, 2012). Therefore, they contain a multitude of sensors that
will adjust the cameras to the light in a particular location, and will record the ambient
temperature, current GPS location, and direction in which the wearer is moving (Autographer,
2012). Such detectors enrich the images by providing additional contextual cues not encoded in
the average photograph. We hope to use the data that these sensors provide to extend our
knowledge of how people form memories and interpret visual stimuli taken from their own lives.
Using these images as experimental stimuli provides a much more realistic view of participant’s
brain activity as they interact with their world, in contrast to generic stimuli used in more
artificially designed studies (i.e. a stock photo of a grassy hill). We are therefore able to delve
below basic perceptual analyses to uncover how the “semantics” of an image taken from
someone’s life may affect his brain’s response to it - such as the salience of the event being
captured, where the event is neurobiologically stored and retrieved, any emotions associated with
it, and any other personal experiential factors that can be solely attributed to autobiographical
events.

I am especially interested in the link between memory and navigation. Under the
guidance of my graduate advisor, we are currently looking to analyze the GPS data of these
autobiographical events to explore variations in brain activation for variables such as a
participant’s familiarity with a particular landmark. Although much research in spatial navigation
has used the concept of a landmark to mean “any visual stimulus within an environment that
could potentially influence navigation,” very little work has been focused on the specific
characteristics or qualities something must possess in order for it to be classified as a landmark
(Chan et al., 2012). Chan et al. chose to examine a variety of studies focused on navigation and
subsequently defined what features increased the likelihood that an object would be considered a
landmark. They found that objects that are unique within their environment, maintain “a stable
spatial position,” and are fixed at directional “decision-making” locations (i.e. at a major intersection) are more likely to acquire the “status” of being a landmark (Chan et al., 2012). In addition, they categorized landmarks as objects that either act as “beacons for navigation,” provide directional cues, are used as associative cues, or are used as a “reference frame for navigation” (Chan et al., 2012). In our study, we are looking into using this taxonomy as a reference to define landmarks on the University of California, Los Angeles campus, as well as within the greater west Los Angeles area.

Several recent fMRI studies have uncovered specific brain regions involved in spatial navigation using landmarks. Of particular interest is a study in which the features of landmarks - visual salience, size, navigational utility, permanence, portability, and whether they were space defining - were all found to significantly engage the parahippocampal gyrus (Auger et al, 2012). However, objects considered to be the most permanent prompted additional activation in the retrosplenial cortex (RSC), indicating that the RSC may be the storehouse of information encoding object permanence, and that neurons in this region may be influencing place cells in the hippocampus (Auger et al, 2012). Additionally, a study by Mullally and Maguire, which required participants to imagine objects of different types and sizes, found that the largest and least “portable” objects evoked greater activity within the parahippocampal gyrus (Mullally & Maguire, 2011).

Hirshhorn et al. (2012) examined the effect of familiarity of an environment on spatial navigation, by obtaining fMRI scans of participants before and after living in Toronto for a year. They found that, overtime, as participants became increasingly familiar with downtown Toronto, there was a significant decrease in activation of the right hippocampus when executing mental navigation tasks (Hirshhorn et al., 2012). However, the “posterior parahippocampal cortex,
lingual gyrus, and superior temporal gyrus” showed increases in activity (Hirshhorn et al., 2012). These findings are consistent with Multiple Trace Theory, which argues that there are two distinct representations for spatial memories, one that is schematic and is thus useful for recollecting semantic memories use for navigational purposes (i.e. “map-like”), and another that encodes pictorial details and is thus useful for recollecting episodic memories (Hirshhorn et al., 2012). Although both memories seem to be initially primarily stored in the hippocampus, overtime, as exposure to a particular environment increases, its schematic representations seem to become “independent of the hippocampus” (Hirshhorn et al., 2012). This theory predicts changes in both the degree of activity in specific areas of the brain, as well as overall changes as to which areas of the brain are being activated, providing both “quantitative and qualitative” information about changes brain activity (Hirshhorn et al., 2012). It is in opposition to the Consolidation Theory, which only predicts changes in the degree of activity of the hippocampus versus the extra-hippocampal regions, thus neglecting to consider overall changes in which brain regions are being activated (Hirshhorn et al., 2012).

In accordance with all of the aforementioned studies and with Multiple Trace Theory, we are looking to explore the navigational effects of landmarks on the UCLA campus, hypothesizing that increases in participants’ exposure to landmarks on the UCLA campus would induce increased activity in the parahippocampal gyrus, as well as in the retrosplenial cortex. The independent variable would be the number of times a participant encountered a landmark, and the dependent variable would any observed changes in brain activity. In order to maintain consistency amongst participants, as well as to ensure GPS data accuracy, it may be appropriate to restrict our definition of landmarks to include only the most permanent “beacons” (i.e. the object itself is a target location for spatial navigation), which are relatively far apart (at least .1
miles apart), and which are ‘familiar’ to the majority of UCLA students (because our participant population consists of UCLA students). Examples of landmarks may include the Bruin Bear, Royce Hall, and the Inverted Fountain, all of which would be referred to by their GPS location. It may also be interesting to include the John Wooden statue as a landmark, for which object permanence may be ambiguous (i.e. the statue is relatively new to the campus, but is intended to remain indefinitely). We are working to define the term ‘familiar’ as an approximate measure of a participants’ exposure to a landmark – therefore, if the participant has attended UCLA for at least a year, and has consciously encountered a landmark on a regular basis (e.g. an upwards of 20 times within the last 3 months), it will be considered ‘familiar.’

In our study, we are focused in learning more about different aspects of autobiographical memory in the brain, including where it is stored (Self vs. Other), how it is temporally stored (Match vs. Mismatch), and how readily available it is to be retrieved (Previewed vs. Not previewed). We are also interested in using GPS data, encoded in the photographs taken by the Autographer, to explore which areas of the brain are activated when participants encounter familiar landmarks in real-life spatial navigation tasks (i.e. passing the Bruin Bear on one’s way to class).

As a research assistant, I have been given the opportunity to work with participants both behaviorally and in fMRI scanners. I have helped prepare stimuli (i.e. creating photographic sequences of episodic events), ran participants, and researched ways in which GPS data has been used in previous spatial navigation studies. Over the next couple of quarters, I hope to learn more about the innerworkings of fMRI, and will ideally gain an understanding of some of the more basic ways in which fMRI data is analyzed.
References


The Effect of Fractions and Decimals on Reasoning

Marvin Lopez

Abstract

The numerous advantages and disadvantages that fractions and decimals provide when reasoning through different types of problems shall be explored via current research as well as articles that have laid the groundwork for the current research. There are three main differences that are found in the current studies. The first is the bipartite structure of fractions (a/b) lends itself to seeing relationships in discreet and parsed continuous data while providing no advantage in continuous data sets. The second finding is that decimals are more advantageous for making magnitude comparisons due to the fact that human subjects seem to have a number line representation of values in their mind when comparing magnitudes. People appear to have the ability to integrate decimal values more easily into the number lines than they do with fractions. The third observed difference between fractions and decimals is that people can be primed to look for certain value relationships as long as they are using fractions, but the relationships are obscured when decimals are being used.
Introduction

Decimals and fractions represent an idea that people have to use in their every-day lives at least in the form of currency. It is because of this that it would be easy to assume that this is an idea that almost everyone feels comfortable with and have a pretty good understanding of. Yet many students in the community college level still have trouble understanding these concepts and this can start as soon as they are exposed to the idea. The problem may be in the way that they are taught. Usually children are introduced to the idea of counting with natural quantities (i.e. 1 egg, 2 apples). Decimals and fractions are both used to represent values that are smaller than 1. Once young students have mastered counting with natural numbers and simple arithmetic the next step is teach them fractions which is usually done with division (e.g. 9 apples into 2 baskets) (Melissa DeWolfe et al., 2013). This means that children are learning 9/2 which associates fractions more with division rather than a quantity that isn’t a numeral. Once children have mastered this concept they are taught how to convert fractions into decimals. This makes it so that children have a procedural understanding of these ideas without really understanding the idea behind them or their significance.

One possible effect of the way we learn fractions and decimals is that they both have situations where people are more likely to use them. Due to the bipartite nature of fractions (a/b) they seem to be easier to use when the numbers in question can be counted. However, they do not apply well when magnitude is what is important. In the case of magnitude it is more advantageous to use decimals because it is easier to see where they fall with respect to whole values as well as other decimals. For example if a person is thinking of slices of a pie they are more likely to think of 1 out of 8 slices or 1/8 of a pie instead of .125 of a pie. However, when
continuous quantities such as amounts of fluid are compared, people are more likely to use decimals over fractional values.

The studies here aim to verify the differences between fractions and decimals and to gain insight into the mental processes involved. This can lead to breakthroughs in how these concepts are taught which may lead to an optimization of learning and understanding rather than superficial memorizing of procedures.

The first study looks at how decimals and fractions are aligned with certain real world objects (DeWolfe et al., 2013). It is believed that people will be better at recognizing relationships in discreet and parsed continuous objects and will have trouble recognizing relationships in continuous objects. Essentially, people will be better able to see fractional relationships if they can be expressed countable forms. This implies that their accuracy will suffer when it comes to continuous values.

In the second study, the participants’ response times and accuracy are measured for both decimals and fractions. The study intends to show that people have a number line representation of values which allows decimals to fit in easily while giving more trouble with fractions. The advantage of decimals over fractions is that decimals provide a clear measure of magnitude. Even though the relationships are ignored in decimal form values are directly comparable to each other which makes this type of task easier for participants. When comparing fractions, people usually will have to convert it to decimal for to know with certainty which value is larger specifically when the fractions have a different base.

The third study looks at how well relationships are actually perceived with fractions and decimals. This is done by showing the participants consecutive equations that contain reciprocal fractions. It is believed that fractions are better at displaying relationships between numbers. Due
to the way fractions are formatted people should be able to see relationships and use the information to speed up their response time in similar fractions that follow it. When using decimals it is expected that people will not be able to perceive the relationship in the different equations so seeing a related function will not have in an impact on their response time.

Procedure

Study 1

In the first study the subjects were all UCLA students who were participating in the experiments in order to receive course credit. This was a within subjects design so there were two parts in the experiment. In one part of the students were shown a decimal value and were shown an accompanying image. The second part of the test was the same except the student was given a fractional value instead of a decimal value. They were shown images that contained one of three types of quantities, discreet, discretized continuous, or continuous values. The participants were asked if the type of relationship was the shaded part out of the whole, or if it was a ratio relationship against the unshaded area. As soon as the image came up on the screen the timer began and once an answer was input, the timer stopped for that trial. If the participants were incorrect they were notified with a red “x”.

Study 2

For the study on magnitude the participants were all UCLA students who were participating in the study in order to receive course credit. This was a within subjects design which meant that the test was split into two parts. One half of the test tested participants’ response time for fractions and the second half tested students’ response times for decimals. During the training, students were shown a letter and were told to press the “A” key if it was a smaller value and the “L” key if it was a larger value. This was meant to accustom the subjects to
the meaning of the keys. For the test trials the students were shown a value that was either a decimal or a fraction. This was a within subjects design so for one part the students would see decimals and for the other part students would see fractions. For this study no feedback was given for incorrect responses.

Study 3

For this study the students were asked to decide whether the displayed equation was true or not. For true responses they would press the “A” key for false responses they would choose the “L” key. This was a between subjects design so student in the decimal condition were only tested on decimal equations and students in the fraction condition were only shown fractional equations. During the test some of the equations came as pairs that were meant to prime students to the relationships between the numbers. If the participants noticed the relationships between consecutive equations then there response time on the second trial should be faster. For this experiment the students were not given any feedback for correct or incorrect trials.

Results

Study 1

In this study incorrect answers were excluded from the results. There was no significant difference between types of relationship (part of a whole vs. part to part ratios). However, there was significant difference in accuracy between fractions and decimals. For trials that contained discreet quantities or parsed continuous quantities, the accuracy for fractions was significantly high than the accuracy for decimals. For continuous quantities the accuracy for fractions and decimals showed no significant difference.
Study 2

Response times for the experiment showed that on average people can make much quicker judgments when the magnitude comparisons are presented in decimal form. There is also a distance effect observed, which means as the test value gets further from the given comparison value, the response time gets faster.

Study 3

In this study it was found that people in the decimal condition had no change in response time for the paired trials. However the people in the fraction condition had a significantly faster response time on the second of the two paired trials.

Discussion

The finding of all three of the experiments support the initial claims. Differences in accuracy show that people identify relationships in discreet and discretized continuous values better with fractions than with decimals. The fact that they were significantly more accurate during the fraction trials is evidence of this.

Fractions are a better display of relationships, evidence of this is the increase in response speed in the paired trials. It can be assumed that people were primed to look for the relationship on the second trial which made it possible to recognize the relationship faster. During the debriefing of the experiment participants in the fraction condition did, for the most part, say that they had noticed that the trials came in pairs. This means that the relationships were explicitly recognized. However for the decimal condition, even though the same values were used except in decimal form, people did not notice this pairing. This provides further evidence that people are better at detecting relationships within values when displayed in fractional form.
When people need to compare two values, in terms of reaction time, it is apparent that the type of number used makes a significant difference. It is evidence that people can make better judgments about magnitude when presented with decimals. The big disadvantage that fractions have in this type of procedure is that fractions can have different bases which makes them impossible to compare directly. Some amount of computation is usually required in order to compare two fractions which may explain the slower reaction times.

All of these findings show that the way people perceive values relies at least to some degree on their understanding of fractions and decimals. The problem with this is that fractions and decimals are two different ways of representing the same type of value. This should imply that there would be no difference in response times and accuracy but in fact there are differences present in the results. There are several reasons why there might be the first of which is the value representations themselves. It might simply be because the forms really represent different types of information better but it might also have to do with how well people understand the two forms. Due to the way the curriculum is set, people really might be getting through school by never understanding concepts and simply memorizing the procedures to get a passing grade. In our subject pool, which is all UCLA students, people may have a better understanding than the average person. If our subject pool came from a larger demographic people the differences may be even larger. Math is a topic that always builds on itself and gets progressively more difficult conceptually and if people do not have a solid foundation as early as fractions are taught then the effect it will have on students by high school and college age can be profound. Due to the fact that for many people their understanding of math, or lack thereof can be the deciding factor of whether or not they pursue a higher education. Modifying the way these concepts are taught can
possibly lead to better education and more possibilities for students who think they are no good at math but in fact are just victims to the way the curriculum has been established.
References


DeWolfe, M., Bassok, M., & Holyoak, K.J. (2013) Analogical Reasoning with Rational Numbers: Semantic Alignment Based on Discrete Versus Continuous Quantities

DeWolfe, M., Bassok, M., & Holyoak, K.J. (2013) Rational Numbers as Relational Models: Analogical Reasoning with Fractions and Decimals
Perceptual Learning and Electrocardiogram Interpretation

Olga Mercer

Introduction

Perceptual learning (PL) can be defined as the process of improving perception, that is getting better at simple sensory discriminations as well as complex categorizations of spatial and temporal patterns that apply to real-world expertise. Eleanor Gibson sets the stage for research on perceptual learning and defined it as “an increase in the ability of an organism to get information from its environment as a result of practice” and she attributes this to the detection of properties, patterns and distinctive features (Gibson, 1969). According to Gibson, perceptual learning makes a significant contribution to cognitive development; In her experiment of the visual cliff, she observed changes in depth perception in infants which is an indication of PL (Gibson & Walk, 1960). But perceptual learning seems to be an important component to any learning, not only in infancy but throughout life. In another experiment, a group of infants and a group of adults were exposed to and learned how to differentiate complex graphic ‘scribbles’ (Gibson & Gibson, 1955) only by being exposed to many different scribbles. This is an example of PL in a ‘low-level’ perceptual task in which participants pick up raw sensory information but don’t need to make sense of it. However, further research shows that perceptual learning also applies to high-level perception, that is discovering and encoding abstract relations. Kellman and colleagues have examined the effects of PL in a variety of domains: for instance, seeing complex patterns in Chinese characters (Thai, Mettler & Kellman, 2011) and in another instance, seeing equation structures across transformations (Kellman, Massey & Son, 2010).
Furthermore, a computer-based learning technology known as a perceptual learning module (PLM) has been developed to study perceptual learning and optimized by Kellman and colleagues. They describe an adaptive learning system that “uses both learning accuracy and response time (RT) as direct inputs to sequencing” the material being learned (Mettler, Massey & Kellman, 2011). Basically, PLMs remarkably improve the efficiency of learning and it adapts to the learning faculty of individual students. In addition, interactive computer simulations, in which graphical concreteness, narrative contextualization and diagrams are being manipulated, have been devised by the Percepts and Concepts Laboratory at Indiana University also to learn more about how perceptual learning works.

Our research is interested in the application of perceptual learning to the domain of medicine, more specifically electrocardiography. Reading and interpreting electrocardiograms (ECGs) is an important skill to have for anyone working in the medical field. However, it is often seen as a real challenge for medical students and medical staff to learn this skill. We have found that even junior doctors have difficulty interpreting ECGs and show a relatively low percentage of accuracy when asked to do so. In summary, it takes a lot of practice with and seeing lots of different cases before a person can interpret ECG tracings correctly. In the following experiment, we used perceptual learning technology to teach people how to read ECGs.

Pilot study

Method

Participants

Undergraduate students of the University of California, Los Angeles took part in this study and received course credit for their participation. They signed up for the study through the Sona systems, a human management pool software. The participants did not have any prior
experience reading electrocardiograms. However, most of them had heard of electrocardiograms prior to this study.

Materials and Apparatus

To run the experiment, we used a computer with internet access. We created a perceptual learning module that pulled images from a database of real 12 lead-ECG recordings. In total, we train participants to diagnose 9 different heart patterns: normal sinus rhythm, Acute Anterior ST Segment Elevation Myocardial Infarction, Acute Inferior ST Segment Elevation Myocardial Infarction, Left Bundle Branch Block, Right Bundle Branch Block, Left Axis Deviation, Right Axis Deviation, Left Atrial Enlargement and Right Atrial Enlargement. For each part of the study we use a different set of images (i.e., the ECGs in the posttest are different from the ones in the training) to prevent rehearsal effects. The PLM uses a sequencing algorithm well-known in the literature and optimal for learning to determine the sequence of presentation of ECG traces (Mettler, Massey & Kellman, 2011). The participants are asked to match each image with the proper heart diagnosis name by picking one of 4 options at the bottom of the screen. Figure 1 shows a screenshot of the PLM for one of the trials. Moreover, we duplicated each ECG tracing in the database so that the second version of a tracing has the diagnostic features marked with arrows and bars as well as a small description of the diagnosis. This serves as a feedback in the module. Figure 2. shows what a typical feedback screen looks like. The feedback screens are not timed in contrast to trial screens where time can run out. Finally, we put together a brief primer that gives basic knowledge about ECGs. It consists of a series of PowerPoint slides with a brief explanation of the 12 leads on an ECG tracing, how to measure the widths and heights of the waves for an ECG on the grid and examples of a typical ECG trace as well as a description of the
distinctive features for each of the 9 heart patterns. We have a small quiz where the participants have to match the name of each heart pattern with their corresponding description to test whether the participant retained the necessary information.

**Design**

In the pilot for this study, we only had one condition, that is one type of training: the active classification condition, in which, for each trial, participants had to choose the correct heart diagnosis depicted on the ECG recording out of 4 options. In the actual experiment, participants will be assigned to one of three conditions randomly: in addition to the active condition, there will be a passive classification condition, in which, for each trial, participants will be shown ECGs as static images with all the relevant features marked and a brief description. They will be asked simply to pay attention and pick out the relevant features for each heart diagnosis. Finally, there will be a passive-active training condition in which participants will go through a set of passive trials and then move on to a set of active classification trials and so on. Before the PLM starts, participants will be randomly assigned to one of these conditions. No matter the condition, accuracy and response time (RT) are recorded for each trial. Also, at the end of the module, we calculate an overall accuracy and response time across all trials, for each part of the study (pretest, posttest, delayed posttest) and for each participant.

**Procedure**

Participants, generally 2 per session, were asked to sit in front of a computer. After recording their name, gender and assigning them a username, the experimenter briefly explain what an ECG is and the value of electrocardiography in the medical field. Before starting the procedure, participants were asked to answer the following questions: “how much sleep did you
get last night?” and “how alert are you on a scale from 1 to 6?” with ‘1’ being not at all alert and ‘6’ being very alert. Then, the participants were instructed to study the primer for a total of 20 minutes and take the quiz at the end to see whether they were ready to start the module. If not, they could look over the primer once more. The module started with a pretest, which consisted of 18 questions, to assess how much prior knowledge the participants had. Again, for each question the participants needed to match ECG tracing with one of 4 options below it, but there was no feedback in this part of the study. Then, the participants continued with the training, which consisted of many trials, with a feedback screen after each trial and a summary of their overall accuracy and response time every 20 trials. In order to complete the module, the participants have to go through 9 mastery levels, meaning they had to steadily improve on their percentage in accuracy but also in their response time for diagnosing each of the heart patterns. It usually took an hour and a half for participants to complete the module. After the training, the subjects were given a posttest, which was pretty much identical to the pretest, that is, there were a total of 18 questions and no feedback. This enables us to see how much they learned from the module. At the end of the session, the participants answered a few demographical questions such as years in college, major, language fluency as well as some questions about their experience participating in the study. Lastly, after a week, the students came back to the lab for a delayed posttest in order to assess how much they retained a week later and how much decay occurred. The delayed posttest was similar to the pretest and posttest. Before the participants left the lab, the purpose of the study was explained to them.

Expected results and Implications

Through this ECG perceptual learning module, we hope to teach heart patterns effectively and efficiently such that the overall response time and accuracy will have significantly improved from pretest to posttest and that there will be virtually no decay over time,
meaning that the percentage accuracy and response time will not have significantly decreased from posttest to delayed posttest. Also, we hypothesize that the participants in the *active* classification condition will perform best since they are more actively engaged in the module and they get more hands on practice with the material. In summary, we expect participants to pick up the general patterns of ECGs with this perceptual learning module.

We have given a similar training to medical students at the David Geffen School of Medicine and they dramatically improved on reading ECGs with virtually no decay observed after 20 weeks suggesting long-lasting gains. We are working on making the module even more adaptive and interactive so that it can track the student’s progress even better and difficulties in learning the material. In the future version of this module, we will try out various learning methods such as “compare and contrast” trials; this will tell us what is most helpful for students performing a cognitive task. Finally, this research will give us information on cognitive processes in general and how we learn a skill. It may be the use of PLMs is an unprecedented strategy for people to become experts in a field.
Figure 1. Example of a trial in the PLM for interpreting electrocardiograms.

Figure 2. Example of a feedback screen in the PLM for interpreting electrocardiograms.
References


doi:http://dx.doi.org/10.1037/h0048826


doi:http://dx.doi.org/10.1038/scientificamerican0460-64


Dimensional Change Card Sort, Flanker, and Speeded Math Task

Sarah White

Methods

Participants

We plan to have approximately 600 participants with a mean age of 7.5. The participants will be from eight different schools in the Los Angeles area, one first-grade, one third-grade, and one fifth-grade classroom per school from each of these eight schools, approximately 20 participants from each classroom. These eight schools were selected because the socioeconomic status (SES) of the students is split with half coming from high-SES backgrounds and half from low-SES backgrounds, as determined by the percentage of the student body that qualified for the free or reduced-lunch program. Half of the schools have a high percentage of dual-language learners; half have a low percentage of dual-language learners; this division was made base on the percentage of students enrolled in English-Language-Learning programs. These eight schools are spread out throughout Los Angeles, representing a broad distribution of racial and academic backgrounds. Other demographics will be determined from the information collected via a language-background questionnaire that we will send home to the guardians of the participants.

Materials

The participants will be completing a battery of executive functioning tasks. The executive functioning tasks are as follows.
Dimensional Change Card Sort

The dimensional change card sort (DCCS) is based upon the task used by Zelazo, et al. (1996) and modifications made by Carlson & Meltzoff (2008). Our version consists of twenty-four bivalent test cards and six practice cards. We had two sets of cards that had inverted values on the two attributes of shape and color. Each 3 in. by 5 in. laminated-paper card has a computer-generated picture of either a car or a flower centered in left-to-right and top-to-bottom on the card. The cars and flowers were either red or blue with a black outline on a white background. Each set of cards had twelve cars and twelve flowers. In one set, all the cars were blue and all the flowers were red; in the other set, all the cars were red and all the flowers were blue. Six cards – three flowers and three cars in each set - had gold stars in the upper-right-hand corner of the card to indicate a rule change; this number was similar to the proportion of cards with stars as used by Carlson and Meltzoff. The sets were pseudo-randomly ordered with each quartile of the cards containing three red cards and three blue cards; three of the cards in each quartile had car pictures and three had flower pictures. There was at least one card with a star on it in each quartile; the other two cards with stars were randomly distributed in the set. The practice cards had the same images; two each of cars without stars, two flowers without stars, and one of each picture with a star. Half of the practice cards had blue pictures; half had red pictures. The boxes that these cards were placed in resembled those used by Zelazo. They were approximately 5.5 inches wide, 4 inches deep, and 8 inches tall with three vertical walls and a floor, open to the participant. There were four boxes, each with a bivalent test card on the center-back wall. Two boxes went with each set of cards and the bivalent test card had opposite values for picture and color than the cards in the set with which the boxes were used; the set of cards with blue cars and red flowers was used with the boxes that had a red car target card and a blue flower target card.
**Flanker Task**

The flanker task was based on the task used by Eriksen & Eriksen (1974) with modifications based on those used by Huizinga (2006). There were two sets of cards with sixteen cards each. Each card was made of laminated-paper and was 3 inches by 5 inches in size. In the center of each card there were five greater-than or less-than signs in black, 100-point, Comic Sans font. 12 of the cards had all of the signs pointing the same direction. 6 cards had all greater-than signs; 6 cards had all less-than signs. On four of the cards, the center sign was the reverse of the four other signs. Two of the cards had a greater-than sign flanked by two less-than signs on each side; two cards had a center less-than sign flanked by two greater-than signs on each side. Four of the cards had green star-shaped stickers centered above the central sign to indicate a rule change; each type of card (all greater-than, all less-than, one greater-than, one less-than) had one card with a star on it in the set. The order of each of two sets of cards was pseudo-randomly generated with one starred card randomly placed in each quartile and one reverse-center card in each quartile.

**Speeded Math Task**

The speeded math task was based on the popular “Mad Minute” worksheets used in American classrooms. Each worksheet consisted of four vertical-half sheets of paper, 3.75 by 11 inches in size. Each page had 25 math problems listed vertically, without numbering. The math problems were written in size-12 Times New Roman font with a space between each number, the operation, and the equals sign. Every math problem only used single digits on the left side of the equation. All of the math problems that were used had solutions that were whole numbers, no matter which operation was placed between the two numbers (addition, subtraction,
multiplication, and division). Only the digits 1 through 9 were used to avoid the confusion of calculating with zeroes that was observed in pr-testing. There were two conditions – blocked and random. In the blocked condition, the operation switched every five problems. In the random condition, the operation switched at random, two problems of the same operation were allowed to follow each other.

Procedure

The guardians of the participants filled out consent forms and a language background for each participant. The speeded math task was conducted with each class of participants all at once. They were instructed to complete all of the problems as quickly and accurately as possible; if they didn’t know the answer to a question, they could guess and move on, but were not allowed to skip any questions. They were allowed to correct their answers by crossing out their first answer and writing their new answer. They were given sixty seconds, using a conventional stopwatch, to answer as many questions as they could and were promised a reward at the end for whoever could finish the most problems. The number of problems completed, the percentage correct, and the number of perseverative errors (errors consistent with the operation prior to the error) were calculated.

For the DCCS, participants were assessed individually. They were told to sort the cards one at a time as quickly and accurately as possible. The participant was told to either sort by color or by picture; then the star indicated to sort by the other attribute. The ordering of which attribute the participant had to sort by first was counterbalanced across participants. Before starting the trials, the participant’s understanding of the rules was assessed using the practice cards. The experimenter corrected errors during the practice trials but gave no feedback during
the trials. During the experiment, the experimenter recorded which box each card was sorted into. The time taken to sort all of the cards, the percentage of correct sorts and number of perseverative errors were recorded. On the flanker task, the experimenter instructed the participant to respond to each card as quickly and accurately as possible by pushing a button with either their right hand or their left hand depending on which direction the center arrow pointed. When there was no star, the participant was instructed to push the button on the same side as the center sign. When there was a star above the center sign, it indicated to push the button on the opposite side of the direction that of the center sign. The participant’s understanding was assessed with six practice cards before beginning the trials and having them re-explain the rules in their own words. For the experimental trials, the experimenter held the cards and changed the card after every response. Another experimenter timed how long it took for the participant to finish the set as well as the participant’s first response for each card (left or right button). Feedback was not given during the trials.

The accuracy and performance on switch trials on the speeded math task will be compared with accuracy and response time on the switch trials of the other two executive function tasks, accounting for ordering of which EF task was conducted first after the speeded math task as well as which version of the speeded math task the participant had. Any differences across language background or socioeconomic status will hopefully not be significant.
References


Reward and Negative Affect Study

Sonia Bhatia

My position as a research assistant at the Social and Affective Neuroscience Lab began in the summer of 2013. Initially, I did not know what to expect as a research assistant, but was extremely excited to begin this journey. During the summer, I began working on a study investigating the relationship between reward and negative affect, which I continued working on during the fall quarter.

Working on the Reward and Negative Affect study was extremely eye opening for me. As a Cognitive Science student, many of my past professors offered extra credit to students who signed up to participate in psychology studies at UCLA. While many of my peers seemingly found this to be an inconvenience, I thoroughly enjoyed participating in these studies. Every aspect of participating in these research studies was enjoyable for me, from interacting with research assistants to learning more about the research that was being conducted. What I liked most about participating in these studies was that through each study, my knowledge in the field of psychology was broadening. Consequently, as a research assistant, it was very interesting to be on the “other side” of the research study, and interact with my peers in a different way than I was used to.

The tasks in the Reward and Negative Affect study were very computer based, as participants would spend at least half of the duration of the study on the computer. During this study, the participants’ initial mood and frustration level was recorded. Then, participants partook in three rounds of a computer task that was meant to be frustrating. During the computer task, participants were asked to press a button when they did not hear a “beep” sound. However,
participants did not know that the computer task was programmed so that in some rounds, it would beep after the participant had pressed the button, which made the task impossible to be good at. Participants were also given a “reward” (M&M candies) after a certain round of the computer task, depending on the subject condition. After each round, the participants’ mood and frustration level was recorded.

My tasks during the Reward and Negative Affect study included running the study on participants, and conducting data entry and analysis. Students would sign up to participate in this study online for course credit, and would come into the lab on their scheduled date and time. I would explain the nature of the study, and walk them through everything that they would do during the study. After they had completed the study, I would debrief them, and explain the true nature of the study, as we had used deception.

Conducting this study on participants was one of the most interesting and exciting things I did as a research assistant. I thoroughly enjoyed interacting with all the students who came into the lab- each student was so different to work with, an aspect that made research so exciting. As a student who participated in many research studies myself, it was easy for me to detect when certain participants were bored, frustrated, or excited with the study. Most research studies are quick for students, but this study took almost an hour- which is a fairly long time for most students. To my satisfaction, most students were very interested in the study, and were very focused and attentive throughout the duration of the study. Inevitably, there were some students who were inattentive throughout the entirety of the study, which was disappointing. Through this study, one of the most useful skills I acquired was thinking on my feet. Many times, I would be faced with problems during the study- MATLAB would act up, the participant would come extremely late, or the computer task would have a glitch. It was my responsibility to think on my
feet for a solution to these problems, because I knew that I only had one hour with each participant, and the data collected would be very valuable.

Working on this study was particularly interesting to me because I had just completed the Psychology 100B: Research Methods course. It was so fascinating to see the correlation between this course and everything I had done during the Reward and Negative Affect study. Additionally, although I used MATLAB during Psychology 100B, I was glad that I had more exposure to the software program during my time in the lab.

After I was finished running the study on at least fifty participants, I compiled and entered all of the data onto an Excel spreadsheet. This was an extensive task simply because some of the data was on the computer, and some data was on loose sheets of paper. Once this process was complete, I moved on to the next project- which was a fMRI study in which I simply assisted with Data Entry.

The last project that I worked on investigated the relationship between certain neuropeptides and social behavior. In this study, participants were given one of two neuropeptides- oxytocin or vasopressin- or were given a placebo. Each participant was given this neuropeptide through a nasal spray, and was then asked to write two passages. The first passage asked the participant to discuss a recent time that they provided support to someone, and the second passage asked the participant to recall a recent time that they got into an argument with someone. The topics of these passages were centered around the hypothesis of the study. We do not have an extensive knowledge on oxytocin and vasopressin, but hypothesized that oxytocin is linked in some way with social bonding and rearing, while vasopressin is linked more to aggressive behavior.
My task during the Oxytocin/Vasopressin study was data analysis. After participants had written both of their passages, I was given about 200 passages to read and rate. Each type of passage had its own rating sheet; I would need to read each passage thoroughly and then rate it on numerous factors. It was interesting for me to read about the different life experiences that people had gone through, and inspiring to learn how much these people had supported their family and friends. It was also unfortunate to read about the arguments that subjects had been involved with, as many times, they mentioned that the argument was not resolved.

Due to the opportunity I was given to be a research assistant, my knowledge about all aspects of research have grown. It was so fascinating for me to witness concepts that I had learned about in my courses come to life in the research lab. In the future, I would definitely love to continue during research, and perhaps begin or pilot a study of my own. I am thankful that I was given the opportunity to conduct research, because I know that it is difficult to find a research lab to work in.
Adaptive Sequencing versus Fixed Spacing Schedule: Known Schedules

Tim Yu

Introduction

With the dominance of the Internet and mobile devices in the modern world, the dissemination of information is more widespread and ubiquitous than ever before. Because of the ease of access to all this information, people’s consumption of information have increased to levels never seen before. However, the retention of knowledge, or the information learned, is a whole different story. This is especially relevant in educational institutions, which are responsible for imparting large amounts of information to students, but useless if the knowledge base acquired by learners is not durable and ultimately not retained. It is not just about the retention though. Efficiency of the learning process is critical as well, due to the typically limited amount of time available for learning and studying. However, to maximize the effectiveness and efficiency of learning requires the realization and understanding that every individual learner has different mental capabilities and needs. As more and more research in cognitive psychology improves the understanding of the various principles of learning and memory, it only makes sense to integrate these principles into instructional methods, and to use the power of technology to customize such methods to the needs of individual learners and tasks, to see if learning could be enhanced and optimized. While there are many suggested methods for enhancing learning, such as combining graphics and verbal descriptions, asking deep questions, and integrating exercises into examples (Pashler, Bain, Bottge, Graesser, Koedinger, McDaniel & Metcalfe,
2007), the current study focuses on spacing effects in learning over time, tests as retrieval method, and effective allocation of the study interval.

While it is general knowledge that repeated exposure to the same bit of information improves the chances of long-term memory retention for it, it is the spacing between the exposures that really helps with more effective learning and better long-term retention. This phenomenon is known as the spacing effects—a temporal delay or space (time or intervening trials) between presentations in order to improve memory. Studies have shown the distribution of spacing to be more advantageous than massing, or cramming (Baddley & Longman, 1978). Because it takes time for working memory to transition into long-term memory, the spacing allows more time for consolidation of these incoming working memories. The various contexts present during encoding also helps improve long-term retention, but does not account for why the spacing effect backfires with extremely long delays, nor for the benefits from short delays. Last but not least, there is the retrieval difficulty factor, in which long term memory strength improves when the items are more difficult to retrieve during practice (Pyc & Rawson, 2009). In addition, when the learning strength is low, taking a retrieval test is more effective than re-studying (Roediger & Karpicke, 2006).

Though there has been much research focused on the size of the optimal spacing, there has been less research focused on finding the optimal spacing schedules, which can be described as a list of spacing gaps between a number of presentations. One way to design a spacing schedule is to increase the size of delay between presentations when the forgetting rate slows, thus sustaining the retrieval difficulty and improving memory. Such a schedule is called a fixed expanding schedule, which was first explored by Landauer and Bjork (1978). They found that expanding intervals during retrieval practice led to better retention rates in comparison to the
fixed uniform and fixed contracting spacing schedules (Landauer & Bjork, 1978). In fact, expanding spacing schedules have been found to be even more effective when learners are forced to process intervening material between the testing items, and when the material is more prone to be forgotten (Storm, Bjork & Storm, 2010). In general, expanding retrieval practices are more effective than equal interval retrievals (what most school textbooks follow) and contracted retrievals.

However, most of the research on spacing schedules have focused on fixed schedules, which do not account for the fact that different learners have different learning strengths at any given time for different things. That is why the Adaptive Reaction Time-based Sequencing (ARTS) learning system was developed—to provide a way in which optimal spacing could be predicted by the different reaction times of different individuals to answer questions (Mettler, Massey & Kellman, 2011). The ARTS method relies on a priority score system to dynamically determine the sequence of the presentation items. High learning strength is assumed for items that are answered quickly and correctly, while low learning strength is assumed for items that are answered slowly, even if correct. The learning strength scores and reaction times are then used to determine and assign priority scores to all of the items, representing the relative importance of their future presentation. To determine the actual presentation sequence, the system automatically selects the highest priority item, which is usually the missed items, for the next trial, meaning missed items would make quick reappearances. On the contrary, quick and accurate responses would result in longer delays for the retention interval.

Because there does not exist much literature on adaptive scheduling, our present study aims to further investigate adaptive scheduling’s ability to optimize spacing schedules and thus, long-term retention of learned items. It is currently not clear how effective trial spacing works in
such adaptive models. More specifically, our study aims to compare fixed scheduling to adaptive scheduling, with the number of delays held constant, in order to see whether adaptive schedules have similar pattern of delays to fixed schedules or not. That is also a comparison that has not been investigated before. Another comparison our study examines is between equal and expanding fixed schedules, also with the number of delays held constant. It is especially critical to control as many variables as possible in order to isolate the effectiveness of the space scheduling methods. Since the ARTS method has already proven to be more effective than a previous model developed by Atkinson ((Mettler, Massey & Kellman, 2011), it is the adaptive sequencing method used in this study. We hypothesize that in terms of performance measured by accuracy and change-scores at the post-tests, the adaptive schedule would produce the highest performance, followed by the expanding schedule, then the fixed schedule.

Methods

Participants

The participants of this study consisted of 72 University of California Los Angeles college students recruited from the University of California Los Angeles Psychology Department Subject Pool. In addition, the participants were granted required study credits for various Psychology classes for participation in this study.

Materials and Design

The stimuli were the names of African countries and their locations on a map of Africa shown on a desktop monitor. There were 24 target items and 14 filler items.
There were 3 conditions set up in a mixed design. The adaptive and fixed (both equal and expanding) conditions were used in a between-subject design, with both conditions containing 36 participants each. Meanwhile, the fixed equal and fixed expanding conditions were used in a within-subject design, for a total of 36 participants.

The experiment featured a pre-test/post-test design, broken up into two separate experimental sessions. The first session consisted of a pre-test, followed by a learning phase, and finished off with an immediate post-test. The second session occurred 1 week later, consisting of a 1-week delayed post-test. All the tests were identical, with each containing 38 target and filler items, and no feedback.

The Adaptive Reaction Time-based Sequencing (ARTS) algorithm was used for the adaptive scheduling. For the fixed scheduling, 24 items were randomly divide between the equal and expanding schedules. In order to help the fixed schedules fit together, minor changes to the size of spacing delays, otherwise referred to as ‘jitter’, were allowed. Otherwise, the fixed equal interval schedule was 5-5-5, while the fixed expanding interval was 1-5-9. The filler items were extra items used to allow spacing delays at the end of training when training items were depleted. They were also used during training to help the fixed schedules fit together. Each target item was presented exactly 4 times by all conditions.

**Procedure**

First, the participant was asked to sit down in front of a designated computer, and to turn off his/her cellphone. Then, the experimenter read instructions verbatim to the participant, telling the participant that he/she was not expected to know any African geography, and was going to be taught African geography. Should the participant have indicated that he/she did have prior
knowledge of African geography, it was recorded in the subject log. The participant was told that for the first phase of the experiment, he/she would be assessed on his/her knowledge of African geography. Then, the experimenter provided instructions for the pre-test. The participant was told to put on noise-cancelling headphones to eliminate any distractions for the duration of each phase. Upon starting the pre-test, the participant was presented with a map of Africa on which a country was highlighted. The participant had to select the name of the highlighted country from a list of African countries using the mouse pointer, and was encouraged to respond within 30 seconds regardless of whether or not he/she knew the answer. After finishing the 24 trials, the participant notified the experimenter, who then provided instructions for the next phase—the learning phase. For the learning phase, the participant answered each question as he/she did in the pretest, except this time, the participant received feedback after each trial, informing him/her of whether the answer was right or wrong. The correct answer was highlighted and the name of the country displayed as well. This feedback was used by the participant to learn all the countries presented. The participant received a feedback screen with bar graphs indicating his/her performance after every 10 trials, and was allowed to use this time to take a break. After successfully “learning” every country, the participant notified the experimenter, who then provided instructions for the last phase of this first session—the immediate post-test phase. The post-test phase was the exact same thing as the pre-test phase. Upon finishing the post-test, the participant was told to return in a week and to not practice any geography until then.

For the second session of the experiment one week later, the participant was asked to take another post-test, which was the same exact test as the ones administered the previous week. When the participant finished, the experimenter debriefed, thanked and dismissed him/her.
Results

As evident from Figure 1, a bar graph of the average accuracies in percentage for each phase of the experiment, adaptive sequencing produced significantly higher accuracy in the delayed post-test than the fixed expanding and fixed equal conditions. This is also evident in Figure 2, a bar graph showing the average change score (the per-test accuracy subtracted from post-test accuracy) from pre-to-post and pre-to-delayed. Adaptive sequencing produced a significantly higher accuracy change score there as well in comparison to the fixed expanding and equal conditions. The same applies to the average gain scores (accuracy improvements from pre-to-post-tests), as shown by Figure 3. Last but not least, Figure 4—which shows the size of adaptive delays by correctness—depicts adaptive sequencing to have the largest average delay between re-presentations of an item, in comparison to the fixed expanding and equal conditions. However, it is also important to note that no significant difference was found between the fixed expanding and fixed equal conditions, as evident by the figures as well.

Discussion

As expected, the adaptive sequence produced the best learning performance in comparison to the fixed conditions. This is because the adaptive nature of the method allowed the delays to more closely match the needs of the different individual learners and items. In addition, that the adaptive sequencing method produced the largest average delay between the re-presentations of an item is consistent with the idea that longer delays between presentations lead to better learning, supporting the spacing effect theory. Despite the positive results, there is still room for improvement for the learning performance. The parameters could perhaps be tweaked a
bit to help with that—for example, some of the spacing delays appeared to contract, which is not what we wanted to happen. The second delay may have been overshot a bit as well, potentially making it a reason why no significant difference was found between the fixed expanding and equal conditions.

Accuracy results by experiment phase

![Bar chart showing average accuracy by experiment phase]

Figure 1. Average Accuracy, in percentage, of each of the 3 scheduling conditions for each of the 4 phases of the experiment.
Figure 2. Accuracy change scores (post-test accuracy minus pre-test accuracy), in percentage, of each of the 3 scheduling conditions from pre-to-immediate post phase and pre-to-delayed-post phase.
Gain scores by experiment phase

Figure 3. Average gain score (accuracy improvements from pre to post-tests), in percentage, of each of the 3 scheduling conditions from pre-to-immediate post phase and pre-to-delayed-post phase.
Size of adaptive delays by correctness

Figure 4. Average delay in trials between re-presentations of an item for overall incorrect and correct answers for each scheduling conditions.


In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory (pp. 625-632). London: Academic Press.


