# Winter 2014

## Table of Contents

- **Relational Reasoning with Fractions and Decimal** .................................................. 3-9  
  Andrew Molica, Faculty Mentor: Patricia Cheng, Ph.D.

- **Photographic Life Logging and the use of GPS data for Spatial Navigation** ................. 10-15  
  Brittney Medina, Faculty Mentor: Jesse Rissman, Ph.D.

- **The Metacognition of Perceptual Fluency: Effects and Potential Influences of Achievement Goals** ................................................................................................................................. 16-21  
  Charlie Hall, Faculty Mentor: Alan Castel, Ph.D.

- **The Effects of Sleep Quality on Virtual Reality Learning and Overnight Forgetting** .......... 22-29  
  Daniel Lin, Faculty Mentor: Jesse Rissman, Ph.D.

- **The Human Connectome Project** .................................................................................. 30-35  
  Edward Owens, Faculty Mentor: Jesse Rissman, Ph.D.

- **The Role of Abstract and Concrete Representations in Multimedia Learning** ................. 36-42  
  Elham Zargarkalimi, Graduate Mentor: KP Thai

- **Perceptual Learning in Music** .................................................................................... 43-48  
  Erik Shiboski, Faculty Mentor: Philip Kellman, Ph.D.

- **Rational Numbers as Relational Models** ....................................................................... 49-55  
  Eugene Goh, Faculty Mentor: Patricia Cheng, Ph.D.

- **Perceptual-Motor Sequence Learning of General Regularities and Specific Sequences** .... 56-63  
  Giancarlo Sanguinetti, Faculty Mentor: Zili Liu, Ph.D.

- **Presence and Immersion: Bridging the Gap between VR and the Physical World** .......... 64-72  
  James Mutter, Faculty Mentor: Jesse Rissman, Ph.D.

- **Optimal Sequencing in Adaptive Learning** ................................................................... 73-78  
  Jason Ku, Faculty Mentor: Philip Kellman, Ph.D.

- **Measuring Infants’ Abilities to Predict an Effect through Causal Inference** ................. 79-85  
  Jingxi Zhai, Faculty Mentor: Scott Johnson, Ph.D.
Dual Process Theories of Memory.................................................................86-91
Kaitlin Swinnerton, Faculty Mentor: Jesse Rissman, Ph.D.

Animation Understanding and Problem Solving........................................92-101
Kevin Ruiz, Faculty Mentor: Hongjing Lu, Ph.D.

GPS Data and Autobiographical Memory...................................................102-108
Laryssa Storozuk, Faculty Mentor: Jesse Rissman, Ph.D.

SGLT2 activity in the Rat Brain.................................................................109-116
Liseth Magana, Faculty Mentor: Jorge Barrio, Ph.D.

Multiplicative reasoning and Rational Numbers........................................117-123
Marvin Lopez, Faculty Mentor: Patricia Cheng, Ph.D.

Metacognitive Judgments and Memory .....................................................124-130
Monica Vu, Faculty Mentor: Elizabeth Bjork, Ph.D.

Perceptual Learning in the Interpretation of Electrocardiograms..................131-141
Olga Mercer, Faculty Mentor: Philip Kellman, Ph.D.

Causal Parsimony and the Appreciation of Fractal Artwork .........................142-150
Priyanka Mehta, Faculty Mentor: Patricia Cheng, Ph.D.

Animation Understanding and Problem Solving........................................151-161
Sana Tauqir, Faculty Mentor: Hongjing Lu, Ph.D.

Reward and Negative Affect Study............................................................162-164
Sonia Bhatia, Faculty Mentor: Naomi Eisenberger, Ph.D.

Perceptual Learning ....................................................................................165-170
Victoria Groysberg, Faculty Mentor: Philip Kellman, Ph.D.

Desirable Difficulty as A Product of Audio Disfluency.................................171-177
Zach McFadden, Faculty Mentor: Robert Bjork, Ph.D.

**These research papers are written by undergraduate students as part of the capstone requirement for the Cognitive Science major.**
Relational Reasoning with Fractions and Decimals

Andrew Molica

Abstract

Research has looked at the challenges students face in regards to fractions and has attributed this misunderstanding to be related to the magnitude of fractions due to its visual representation. New studies that focus on educated adults have shown that the fraction notation can actually be advantageous when solving relational tasks when compared to decimal form. This study explored fractions in a traditional mathematic form that fractions are commonly used: multiplication. College students were asked to judge fraction and decimal multiplication equations as correct or incorrect, and results concluded that fraction multiplication was faster and more accurate than decimal multiplication. When reciprocal primes were added to the problems, participants were able to solve the second fraction problem faster than the first. This study demonstrates that adults understand and can identify relations in multiplication in fraction notation better than decimal form.

Introduction

The American mathematics education system has been shown to focus less on the understanding of concepts in mathematics and more on the memorization of procedure to compute an answer (Behr & Post, 19992). This is problematic when a conceptual understanding of mathematics is required to solve a problem that does not fit the mold of the standard procedure that has been memorized. A sample of eighth graders were asked to add $12/13$ and $7/8$ and select the closest answer from the choices of 1, 2, 19, and 21. The correct answer, 2, was chosen less than both 19 and 21 (Carpenter, Corbitt, Kepner, Lindquist, & Reys, 1981). Both $12/13$ and $7/8$ are close to 1 which should allow for an easy estimation of the answer 2, but
students were attempting to add numerators and denominators because that is the process they had been taught.

An important concept in mathematics is the understanding of mathematics as a system of relations between quantities (Richland, Stigler, & Holyoak, 2012). Before higher order mathematic concepts can be learned, a detailed, conceptual understanding must be established in order to create a foundation for more complex concepts.

When students begin multiplication with decimals and fractions, decimal multiplication is usually taught in a similar way to whole numbers. Students are initially challenged by locating the decimal place, but the procedure to obtain a product is almost exactly the same as multiplication with whole numbers (Hiebert & Wearne, 1985). The similarity in multiplication does not carry over to fractions.

Students have trouble when rational numbers are expressed in the $A/B$ format. This form is much different than a decimal which contains only a value. Fractions are a relation between the numerator and denominator. $1/2$ and $3/6$ represent the same magnitude, but $1/2$ represents the magnitude as a relationship between 1 and 2 and $3/6$ represents the magnitude as a relationship between 3 and 6. These infinite possibilities of expressing the same magnitude require students to understand how one number relates to the other, unlike decimals where only $.5$ could represent this magnitude. When students begin multiplying fractions, they must be able to conceptualize multiplication and the relational expression of the fraction. It has been proposed that a true comprehension of multiplication requires an understanding of reciprocal relationships the multiplied factors and their product (Thompson & Saldanha, 2003). Once a fraction is added to the process another element of relational reasoning is present to complicate the problem further.
Since multiplication with decimals is only a slightly more complicated version of multiplication with whole numbers it should be easier than multiplication with fractions that have the added level of relational reasoning.

To see if adults have a better understanding of the relationships in fraction multiplication, a relational priming paradigm could show that adults use the relationship from the first problem to help them solve the second problem faster.

Methods

Participants

The participants were undergraduates from the University of California, Los Angeles. There were a total of 60 participants (mean age 20, 47 females) who volunteered for the study to receive course credit. Thirty participants were randomly assigned to two between-subjects conditions.

Design and Materials

The study was a 2 by 2 by 2 design (number type: fractions vs. decimals, trial type: first trial vs. primed trial, problem type: true vs. false). Number type is a between-subjects factor and trial type and problem type are within-subjects factors.

There were a 240 problems total, and half of the problems were correct. The stimuli were multiplication problems in the form of \( A \times \frac{B}{D} = C \). \( \frac{B}{D} \) was represented either as a fraction or by the equivalent decimal rounded to two decimal places (15 \( \times \frac{1}{3} = 5 \) or 15 \( \times .33 = 5 \)).

Half of the trials were prime pairs where a prime was followed by a primed reciprocal. The correct primes were represented as \( A \times \frac{B}{A} = B \) and \( B \times \frac{A}{B} = A \). \( \frac{B}{A} \) and \( \frac{A}{B} \) were reciprocals of each other. The other half of the problems were primed pairs that were incorrect,
represented by \( A \times B/A = C \) and \( C \times A/B = A \). \( A/B \) and \( B/A \) still share a reciprocal relationship.

Each participant was randomly assigned to receive \( A/B \) or \( B/A \) first in the primed pairs.

The unprimed trials had many foils and fillers. The form of the fillers was the same as the primed trials, but some of the fractions were equivalent reduced forms \((15 \times 1/5 = 3 \text{ instead of } 15 \times 3/15 = 3)\). Half of these unprimed trials were correct. The problem order was randomized so the some problems appeared true followed by true, true followed by false, false followed by true, and false followed by false.

**Procedure**

Each experiment was run on a Macintosh computer using Superlab 4.5. The response times and accuracy of each participant were recorded. Participants were given the following instructions.

*Decide whether or not the following problems are correct or incorrect. If the problem is correct hit the a key, and if the problem is incorrect hit the l key. All of the answers have been round to the nearest whole number. Answer the problems as quickly as possible while still being as accurate as possible.*

**Results**

**Accuracy**

Participants scored an average accuracy of 90% for fraction problems and 78% for decimal problems \((t(58) = 5.30, p < .001)\). A 2 by 2 by 2 (number type, trial type, problem type) mixed factors ANOVA showed a strong effect of number type favoring fractions at 94% over decimals at 78% \((F(1, 58) = 41.19, p < .001)\). Number type and trial type showed no significant interaction, which means there was no differential priming effect in accuracy for fractions over decimals. There was no effect of trial type which indicates there was no general priming effect on accuracy. There was no reliable three way interaction.
Response Times

The average response time for fraction problems was 2.76 seconds, which faster than the average response time for decimal problems at 4.03 seconds (t(58) = 3.93, p = .001). Each participant’s response times were averaged over each trial for each of the true and false primed pairs. The response times on incorrect answers were excluded from the analyses. A 2 by 2 by 2 (number type, trial type, problem type) mixed factors ANOVA showed a significant three way interaction (F(1,58) = 8.27, p = .006). There was a significant two way interaction for the true primed pairs (F(1,58) = 10.72, p = .002). This indicates a differential priming effect for fractions compared to decimals. A planned comparison showed that the second (primed) trial averaged 2.21 seconds and was significantly faster than the first (prime) trial (1,29) = 3.08, p = .004). The decimal problems did not reveal a significant difference for the primed trial which averaged 3.84 seconds compared to the priming trial which averaged 3.58 seconds (t(29) = 1.72, p = .10). The response times support that there is a priming effect with fractions and no priming effect with decimals.

False prime trials showed no effect on trial type (F(1, 58) = 3.158, p = .081) or number type (F(1, 58) = .08, p = .778). There was also no significant two way interaction (F(1, 58) = .537, p = .467). A priming effect for response times was only evident when the problems were correct.

Discussion

Overall, fraction multiplication tended to have higher rates of accuracy and speed over decimal multiplication. When college students saw the fraction multiplication, they were able to simplify the problems to allow for faster solving, which is not the case with decimals. A problem such as 14 X 2/7 = 4 is easy to reduce to 14 X 2 = 28 and then 28 / 7 = 4 or 14 / 7 = 2 and then 2 X 2 = 4. The decimal example of this does not have a simplification that is as efficient or
effective. 14 X .29 = 4 could be changed to 10 X .29 and 4 X .29, but 4 X .29 will most likely either take several seconds to compute or require a guess of the value.

The decrease in reaction time for solving primed fraction multiplication problems, but no significant change between primed decimal multiplication problems supports the idea that adults do understand the reciprocal relationship between fractions. The relationship is visually obvious in fraction problems (2/9 and 9/2) where in decimal problems it is harder to see (.22 and 4.5). Adults saw the relationship in fractions which allowed them to answer the question without having to solve it again, because they had just solved the same problem in reciprocal right before. They did not need to redo the calculations for fractions as seems to be the case with decimals.

The hypothesis that fractions have a rational component that college students have learned to use when solving fraction multiplication problems is supported by the findings of this study. The deep understanding that college students have allows them to solve the fraction multiplication problems with greater speed and accuracy than the decimal problems. The vital relationships of reciprocity and multiplication can be taught using fractions due to their particular form and this understanding cannot be achieved as effectively with other rational numbers.
References


DeWolf, M., Bassok, M. & Holyoak, K. J. (under review). Rational numbers as relational models: Analogical reasoning with fractions and decimals.


Photographic Life Logging and the use of GPS data for Spatial Navigation

Britney Medina

Successful navigation throughout one’s environment is a crucial feat humans undertake in every day life. Whether it is recognizing one’s way to the usual grocery store or driving to an unfamiliar location; humans use internal cognitive maps to navigate the world. It has been demonstrated that the hippocampus is required to learn the layout of a new environment, however, it is debated whether it plays a role in navigation and spatial memory retrieval for already learned layouts or layouts that one has become familiar with (Hirshhorn et al., 2012). This is an interesting finding as it posits the idea that other regions (independent of the hippocampus) might be responsible for the navigation of familiar places as well as the retrieval of these kinds of spatial memories. A longitudinal study tracked changes in brain activation that occurred as participants became familiar, over time, with an environment. This study found that the hippocampus is not necessarily required to support mental navigation in a familiar environment. Instead, at least under some circumstances, mental navigation can be supported by a network of extra-hippocampal regions (Hirshhorn et al., 2012). It would be very interesting, in fact, to track brain activation (hippocampus or elsewhere) for UCLA students as they become more and more familiar with the campus.

In addition, an area that has not been thoroughly studied, however, is how humans integrate these routes into a cognitive map and whether these cognitive maps actually preserve real-world distance relationships. More broadly, what role do medial temporal lobe regions play in the retrieval of spatial navigation? These questions were recently raised to me and another undergraduate student as newly extracted GPS information has become available through the use of participant’s wearable, high definition camera devices. The current study I am involved with
investigates the neural correlates of Autobiographical memory and the differences in brain activation when subjects are asked to make judgments of the distinctiveness of a memory among other tasks (whether it is from their own life or someone else’s life). This study requires a participant to wear a portable, lightweight camera device that takes automatic pictures of their life throughout a period of three weeks. Within this period, the autographer camera takes thousands of pictures that are later used to create sequences or events to recreate a memory of an event. These sequences are made to depict a linear progression (tell a “story”) and recreate a unique or personal experience that occurred within a 15-minute window. These sequences are presented to them, while in the fMRI scanner, to test different aspects of memory retrieval. It was recently discovered that these cameras also contain latitude/longitudinal GPS information for a given picture giving us information about the whereabouts of participant’s locations for a given image captured.

Extracting, organizing and figuring out what to do with this newly found data has become a side project for Laryssa and I. We have created a MATLAB program that will extract GPS data in degrees, minutes, and seconds for each participant and organize it into a matrix to decipher exactly how much data we have to work with. So far we have only analyzed the data for one participant, however, the program is able to organize information for an arbitrary number of participants. We found that this particular participant had GPS data for 256 images out of 960 (about ¼ of all images contained data). Also, for any sequence that had at least one image with data, the mean number of images with data for a sequence was about 5/8. Sometimes the autographer camera does not record any GPS data for a given sequence (might be due to satellite obstruction) or it might not record data for every single image in a sequence. For the time being, however, it is still unclear if the data is sufficient enough to deduce distance measurements or
spatial cognitive relationships between given locations. In any case, this GPS information does relate to other articles that have examined spatial navigation and memory using other methods.

Reverting back to the question asked earlier, a study by Morgan et al. used fMRI data to test for a key aspect of a cognitive map: preservation of real-world distance relationships and their role in the human hippocampus. For their study, 15 University of Pennsylvania students, who had at least 1 year of experience with the campus, were shown color photographs of 10 prominent landmarks from the UPenn campus (22 images for each landmark → 220 images total). In the scanner, subjects were asked to covertly identify each campus landmark and make a button press once they had done so. Subjects were not told to physically or mentally navigate between landmarks. Instead, subjective “distances” between landmarks were determined by asking subjects to estimate number of minutes required to walk between each pair of locations. Take note, however, that this distance calculation is a spatial-temporal distance as opposed to an actual physical distance. The study suggested that the hippocampus calculates the extent to which the current stimulus is consistent or inconsistent with contextual spatial relationships. In other words, the hippocampus considers locations that are physically closer in space to be more representationally similar and locations that are further apart in space to be more representationally distinct. To put it more clearly, by viewing a familiar landmark, participants must have established a “context” for each stimulus for which to deduce representationally similar distances as opposed to representationally different distances. Decreased activity was found in left anterior hippocampus for proximal landmarks due to lesser contextual changes (representationally similar) and increased activity in left anterior hippocampus for distal landmarks due to greater contextual changes (representationally different). This hippocampal response indicates sensitivity to the spatial relationships between
landmarks in regards to their contextual proximity as opposed to their physical proximity. In regards to a cognitive representation of physical proximity, a different study shows that the entorhinal cortex, in fact, encodes metric information about the spatial relationships between landmarks (Doeller, Barry, & Burgess, 2010). With this information in mind, another article also suggests that other parts of the medial temporal regions are associated with spatial navigation.

For this study, 13 participants (who had moved very recently to Toronto) were scanned in an fMRI scanner while performing various mental navigation tasks involving newly encountered Toronto landmarks. Eight of these participants returned for a second fMRI session after living and navigating in the city for about 1 year. This study concluded that the hippocampus is involved in mental navigation for a recently learned location; however, it might not be as pertinent or involved in mental navigations for a highly familiar environment (Hirshhorn et al., 2012). In fact, the study showed increased activation in the parahippocampal cortex, lingual gyrus, caudate, posterior cingulate/retrosplenial cortex, and prefrontal cortex regions during the second session after participants had become familiar with the city. This study is the first to track brain activations in spatial memory for short term vs. long-term retention intervals (familiar vs. unfamiliar environments) (Hirshhorn et al., 2012).

In sum, it has been discussed above that landmarks play an essential role in human’s ability to spatially navigate their environment. Human’s cognitive maps have different schematic representations depending on if the person is familiar or unfamiliar with their environment. The studies suggest that one’s mental representation of the environment changes and becomes independent of the hippocampus with increasing familiarity (Hirshhorn et al., 2012). According to another study, the hippocampus, however, does play a crucial role in encoding contextual distances to form spatial relationships in one’s environment. In fact, another
study has shown the entorhinal cortex may play a role in the representation of actual psychical, metric distances of space (Morgan et al., 2011). This knowledge is relevant to the spatial navigation project we have been working on because the students in our study are a mixture of all grade levels and have different levels of familiarity with the campus. Perhaps, retrieval of a memory for a given sequence will also depend on how long they have been a student at UCLA navigating the campus. Additionally, the distance traveled within a sequence could be used for correlation studies that deal with brain activations in extra-hippocampal structures such as the entorhinal, parahippocampal and other medial temporal regions. Structuring the GPS data for participant’s images, in order to extract useful statistics, has been a very interesting endeavor this quarter and we hope that this information can be useful for future studies investigating how spatial maps and navigational tactics are represented in the brain.
References


Motivation affects cognitive processes in profound ways and often results in differences in attention, memory, decision-making and problem-solving. In cases when motivation rests on achievement, researchers have identified the differences between mastery and performance goal states and the effects each goal state has on metacognition. Mastery goals are defined as attempts to develop competency, whereas performance goals focus more on the demonstrative of competence relative to others. In 2011, Murayama and Elliot attempted a study to observe the effects of these two achievement goals on short-term and long-term memory. Separately, a 2008 study conducted by Rhodes and Castel exposed metacognitive illusions of varying perceptual fluency. These two studies analyze vastly different principles of memory; however, by understanding the proposed mechanism for the results of each study separately, inferences can be made about the possible effects of achievement goals on the metacognitive illusions of perceptual fluency.

Achievement Goals

Driven by the absence of achievement goals effects on memory, Murayama and Elliot attempted to examine the effects of each goal state on short-term and long-term memory. In two experiments, participants from either Japan (Experiment 1) or The United States (Experiment 2) were separated into two groups; each group was given instructions eliciting either a mastery goal or performance goal state. After these performance goal states had been established, participants were presented with word lists, including word generation tasks, lures, and, in Experiment 1,
stroke counting tasks. At the conclusion of the encoding phase, participants were tested for recognition of the words on a remember-know test that required participants to make a judgment about whether a certain word appeared on the lists. This test was repeated after one week to test long-term retention. The results of the study indicated that performance-based goal states during encoding resulted in better immediate recognition than mastery-based goal states; however, the effects were reversed for the long-term recognition test.

The results of this study indicate that mastery and performance goals may influence participants to use different memory strategies to accomplish the same task. Murayama and Elliot propose that mastery goals lead to “rich and variable contextual associations that likely consolidate over time” while performance goals lead to “narrow but concentrated memory traced that help memory in the short-run but are likely to rapidly decay”. This distinction highlights that mastery and performance may result in different learning due to difference in encoding. Encoding-based explanations of differences are further supported in this study because participants were asked to discard their achievement states before the recognition task. Therefore, the present results seem to support a theory in which different methods of encoding are employed by participants depending on their purpose when learning. Furthermore, the recognition tasks given to participants were supposedly completed in neutral achievement states which would indicate that the effects must have occurred during encoding.

*The Metacognitive Illusions of Fluency*

Perceptual fluency refers to the relative ease by which a particular item can be perceived. In memory research, perceptual fluency has not been associated with increased memory performance. However, when asked to make judgments involving the likelihood of future recall, participants in a study conducted by Rhodes and Castel in 2008, showed signs of a
“metacognitive illusion” in which items more easily perceived were judged more likely to be remembered. In each of six experiments, participants were asked to memorize words presented in either large font size or normal font size. In addition, participants indicated their predicated likelihood of recalling a particular word directly after its presentation by indicating a judgment of learning (JOL). As previously found, font size did not result in a difference in the ability to recall an item in the free recall test. However, even after experience, explicit and knowledge of the effect, a “metacognitive illusion” that more easily perceived items are more easily remembered persisted in the relative JOLs for normal or large font sized words. As the authors indicate, “these data indicate that participants made predictions of future memory performance based on salient perceptual information available at encoding”. The tendency to rely on perceptual cues for predictions of memory is part of the cue-utilization framework theory that postulates that learners rely on all contextual information of cues in order to make predictions of memory performance. Because other contextual information was controlled, font size provided the only possible information for participants asked to predict their memory. While font size is used in this study as an operational manipulation of perceptual fluency, the evidence it provides for a “metacognitive illusion” exposes a disjunction at the time of encoding between actual memory performance and predicted performance.

Proposed Effects of Achievement Goals on the “Metacognitive Illusion”

Thus far, two seemingly unrelated examinations of memory performance have been conducted resulting in the identification of two effects on memory that occur during the encoding of new information. First, alternative achievement goals during encoding have been shown to produce different results in short-term and long-term memory. Next, perceptual fluency has been shown to create a “metacognitive illusion” characterized by an increased confidence in memory
for more fluent items. Furthermore, each effect approaches metacognition very differently. On one hand, the effects of achievement goals appear to be caused by the conscious metacognitive employment of different memory systems while the overconfidence induced by fluent items appears to point to a metacognitive breakdown. In order to better understand the scope of each effect, the “metacognitive illusion” of perceptual fluency may be analyzed in relation to achievement states. If a study were to induce achievement goals before presenting items with differing perceptual fluency, the results might shed additional light on the broad-based nature of mastery goal encoding and the narrowly focused nature of performance based goals. The fundamental purpose of combing these effects would be to determine if either broad-based or narrowly focused encoding resulted in the reduction or growth of the “metacognitive illusion” of perceptual fluency. If performance goals induce narrowly focused encoding, it is possible that performance goals could make people more susceptible to the belief that a more fluent word would be more easily remembered because they are attending to that particular word’s features more than mastery goal-based elaborative encoding that may disregard perceptual features for more elaborative encoding. Alternatively, the competitive nature of performance goal learning may cause individual JOLs to exhibit the desire to remember that word better rather than an accurate prediction, resulting in the masking of the “metacognitive illusion” rather than its elimination. If this were the case, mastery goal states might elicit greater accuracy than performance goals in JOLs. Clearly many possibilities exist when predicting the result of combining achievement goal states with perceptual fluency and a properly designed study may be able to provide more evidence for each effect.
References


My Experience

I am so grateful for the opportunity to capture a glimpse into the world of graduate research provided to me by this class and lab. I enjoyed Psych 100A and 100B, but being able to see the concepts I learned in those classes come to life in the study design, data collection, presentation, and in the discussions in the lab meeting has been incredibly exciting, if not overwhelming. The most rewarding experience for me was attending the lab meetings and listening to the logical way the graduate students and Professor Castel approached each presentation and discussion. I look forward to graduate school when I can be surrounded by similarly talented and passionate people. While I was incredibly busy this quarter with other obligations and felt like I couldn’t take advantage of every opportunity available to me by participation in this lab (e.g. submitting to PURC), it was a great way to get my feet wet in the research community at UCLA. I would love to be able to increase my exposure to research related skills such as data analysis, study design, lab presentation, etc. I can’t wait to build on the knowledge and exposure to research this quarter has given me in future quarters. I also cannot be more grateful to have had a wonderful graduate student advisor who took extra time to talk with me about graduate school admissions and went out of her way to inform me of additional opportunities for growth in the psychology department 😊
The Effects of Sleep Quality on Virtual Reality Learning and Overnight Forgetting

Daniel Lin

Abstract

As virtual reality (VR) becomes increasingly used as a research and education tool, understanding what factors modulate its effectiveness has become a pressingly necessary, yet often neglected, area of research. Research shows that sleep is important for memory consolidation and thus retention. The current study examines the effects of sleep quality on VR learning—particularly, overnight forgetting. While immersed in VR, participants learned the pronunciations and meanings of 42 Swahili nouns in three learning sessions, each followed by cued recall as retrieval practice. They were then tested again the next day. The difference in memory recall performance between the final test of Day 1 and the initial test of Day 2 was defined as an overnight forgetting score. Sleep quality was measured on Day 2 using the Pittsburgh Sleep Quality Index (Buyess, 1988). We hypothesize that sleep quality will negatively correlate with overnight forgetting, such that the better a one sleeps, the less one will forget the learned material (i.e., better retention). If confirmed, our results might inform user-end adjustments to make the most of VR-based education.

Introduction

In recent years, many institutions have heavily invested in VE-based learning to enhance presentation of information and to make learning more accessible. Whole campuses are being reconstructed in virtual worlds, for example, where students can roam in and around digital classrooms replicas and learn. Virtual education and computer-based learning have the exciting potential to present and teach information in ways that aren’t feasible in the real world, with the hopes of being a freely accessible or relatively low-cost educating tool. Innovative techniques are
being developed to utilize this potential: students can already instantly be immersed in digital reconstructions of thriving ancient cities, and giganticallly-scaled 3D models of macromolecules can be built in seconds, for example. But overall, there is a lack of research that investigates the effectiveness of learning information specifically learned in virtual education, and what factors affect it.

There is a paucity of research that investigates what factors modulate the effectiveness of information specifically learned in virtual environments. With this project I hope to begin bridging the gap between classic psychology learning and memory studies with virtual reality education to help inform the end-users of this new technology on how to maximize their learning experience while using it. Specifically, I wish to examine the effects of Sleep Quality on memory of information learned in virtual environments.

I believe this is important because a sizeable portion of VR-learning users use VR remotely in their own homes, late at night after they come home from work, risking the chance that they cut into their sleep in order to finish a VR class. Understanding the relationship between virtual environment learning and sleep can help these users understand whether it’s worth sacrificing sleep to finish a virtual reality lesson, for example. 

Sleep is well-known to play an important role in memory consolidation. Better sleep holds a strong relationship with better long-term retention of information and better academic performance (Howell, 2004). We hypothesize that sleep quality will negatively correlate with overnight forgetting, such that the better a one sleeps, the less one will forget the learned material (i.e., better retention).
Methods

24 unpaid participants were recruited from the University of California Los Angeles Psychology Department Subject Pool. The participants were naïve to the purpose of the experiment, and had given electronic acknowledgment that they have read the human subjects and privacy policies, and agree to the terms therein.

Participants are tasked to learn and recall a list of 42 Swahili words. These stimuli were presented in OpenSim, a software used to create virtual environments, and presented on an LCD screen in a dark room. The procedure proceeds as follows: On the first day of the experiment, participants begin the session by customized their own virtual avatars—which they will be using for the rest of the experiment—for fifteen minutes. Participants move around the virtual world using a joystick, and communicate with their experimenter via virtual avatar communication and in-world voice chat.

For learning sessions, participants watch a video within the virtual environment at a pre-determined location within the virtual context or ‘island’, which then displays a digital representation of an object, it’s English name, and read aloud the Swahili translation three times before moving on to the next stimulus word, as seen in Figure 1. Participants are instructed to repeat the Swahili pronunciation each time they hear it.
For testing sessions, participants are given a cued-recall test to aid their learning as a form of retrieval practice, but we also simultaneously score their performance to track their progress and information retention throughout the experiment. Testing stimuli include a video that is also played within the virtual environment, showing a series of unlabelled real-world representations of a previously-trained object from the learning session. Participants were given ten seconds for each picture to recall the correct Swahili translation, while their voices and answers were recorded. Performance was determined by the proportion of syllables in the word that participants were able to recall and pronounce correctly.

On day one of the experiment, participants went through three learning sessions of the stimuli list, with each learning session proceeded by a testing session. They would go home and return the next day during the same time of day, and start the session with a testing session.
before being allowed to review any of the learned words. Stimuli presentation order was randomized for each testing and learning session.

The difference in performance between this first testing session of day two and the last testing session of day one were used to construct an ‘Overnight Forgetting score’ as follows: for each tested word, a score of 1 was given if the participant performed worse on the second day session than for the first day testing session; a score of 0 for no change in performance; and a score of -1 if they performed better than the first day testing session. A value was assigned for each tested word on the stimuli list, and an Overnight Forgetting score was formed by summing all of these values together. Therefore, participants who performed relatively poorly on the second day testing session should have a higher Overnight Forgetting score as a result from forgetting more words overnight, and vice versa.

Sleep Quality was determined by the Pittsburg Sleep Quality Index (PSQI), given through a post-experimental survey at the end of the experimental session on day two (Buyess, 1988). A scoring algorithm determined a single-value Global PSQI score, with a higher score indicating poorer sleep quality. We divided determined “Good” sleepers as those with <6 score, “Poor” sleepers with scores between 6-9, and “Very Poor” sleepers with scores >9.

In our analysis, we ran a one-way ANOVA to determine the differences in overnight forgetting between the three different sleeper quality groups.

Results

Figure 2 presents the average Overnight Forgetting score across each sleep quality condition. Looking at the pattern of results displayed in Figure 2, it appears that overnight forgetting score decreased across each poorer level of sleep quality. In other words, found that on
average, better sleep quality causes participants to perform worse on more words than those with better sleep quality, and this effect was significant.

To test these apparent effects, the data were analyzed using a one-way between-subjects ANOVA, which revealed a significant main effect on sleep quality, such that the average Overnight Forgetting score was significantly higher in participants with ‘Good’ sleep quality than those with ‘Poor’ or ‘Very Poor’ sleep quality, $F(1, 24) = 5.206, MSE = 79.68, p = .015$.

![Graph showing the relationship between sleep quality and overnight forgetting score.](image)

**Figure 2.** Those with lower sleep quality were forgetting less compared to those with better sleep quality. *Y label: Overnight Forgetting score. X label: Sleeper Group category.*

**Discussion**

These preliminary results show that overnight forgetting decreases as a participants’ quality of sleep becomes worse. This was inconsistent with previous research claiming that better sleep leads to better retention and performance on memory-based tasks. Perhaps further directions for this experiment could include exploring the fact that our participants may on average be chronically sleep deprived either from being in college, or specifically in UCLA, and generally maintain a poorer sleep quality against students from other schools. One possible
explanation for these results could be from the initial day one performance of those with good sleep quality: perhaps they were forgetting more because there was more to forget—as in they, on average, were getting more correct on the first testing sessions, leading them to have a higher chance of obtaining a high overnight forgetting school than, say, someone who could only remember two words at day one and forgot one of them on day two. So to investigate this, I could analyze and compare the initial performance of participants at the end of day 1. For future directions of this experiment, we plan to examine how sleep quality affects students in the long term by looking at their testing session performance on day four and day eighteen, both of which we are in the process of collecting data for. Perhaps it may be the case that poor sleep quality leads to better short-term performance but ultimately worse long-term performance. Overall, we were left intrigued by the interesting preliminary results, and we look forward to further investigating this effect and finding an explanation.

Acknowledgements

I would like to thank Joey Essoe, Jesse Rissman, my fellow undergraduate research assistants of ALIVE, and the rest of the Rissman Memory Lab team for their valuable input and discussions that supported and guided me to finishing this project.
References


The Human Connectome Project

Edward Owens

This quarter I continued to work in Dr. Rissman’s Memory Lab under his graduate student Andrew Westphal. My data analysis has shifted from processing of ROI’s (region-of-interest) utilizing SPM and other various pipelines in the medial-temporal lobe (MTL) to data analysis and processing of resting-state data. After ending my work on the MTL project, I began a resting-state analysis script on the data provided by the third quarter of the Human Connectome Project (HCP). fMRI analysis utilizes resting-state scans for various reasons and the HCP has streamlined the process of acquiring the data. During the winter break, I coded a resting-state analysis script that takes an ROI and finds correlated brain regions which are active during a subject’s rest. I am currently working on a script that will allow X, Y, Z coordinate input, and find correlating resting-state brain regions with the inputted coordinate. The script utilizes the Human Connectome Project Workbench and the data from the HCP.

The HCP is a five year long project that will recruit 1,200 subjects for various tasks, diffusion imaging and resting-state analysis. The project utilizes the latest analysis and scanning techniques to acquire the data and a custom 3T scanner to perform the scans. The project is particularly focused on myelin mapping and also uses MEG (magnetoencephalography) and EEG (electroencephalography) to acquire supplemental data. Subjects are scanned in a 3T Skyra at Washington University for diffusion imaging (dMRI), resting-state fMRI (rfMRI) and task fMRI (tfMRI). All subjects perform behavioral tasks and their results are measured along with a blood draw for genotyping. Two-hundred subjects are selected from the original 1200 for additional scanning in a 7T Siemens scanner. A subset of 100 subjects is selected from the original sample to be scanned using MEG and EEG during rest and while performing a task (Van
Essen et al., 2012). The HCP wishes to correlate brain structures and activity with certain behavioral measures such as working-memory, or visual acuity. By utilizing the best-to-date analysis techniques and scanning hardware, the HCP will be able to change the way in which neuroimaging data collection and analysis is performed in fact, the change has already begun.

In a study conducted by Markov et al. (2014), tracers were injected into the cerebral cortex of macaques to investigate the neural connections of the monkey’s cortex. The experiment revealed 1,615 pathways, a third of which had not been discovered. The study also found an unexpectedly high amount of unidirectional links in the connectivity matrix (Markov et al., 2014). The information from the study is used to generalize into the connectivity of humans which would not be nearly as effective if the noninvasive techniques of the HCP had not been developed. The HCP implements a standardized technique for acquiring fMRI data, resulting in far more effective generalizations.

Investigating the relationships of differing brain regions has powerful medical implications as well. Although the physical damage of a brain region is contained within the affected region, the extensive connectivity of the brain leads to far more complicated effects of lesions. For example, Carter, Shulmann, and Corbetta (2012) utilized the resting-state data from the HCP to investigate problems in neural connections following a lesion caused by a stroke. Resting-state data was used due to its effectiveness in imaging the reorganization of interconnecting brain regions after a stroke. Seeing how the brain reorganized itself after a lesion provides invaluable information for the treatment and rehabilitation of those affected. Studying how connections are affected by lesions has commonly been used with task-evoked fMRI data which has been incredibly fruitful. A problem with task-evoked data is the patient’s ability to physically perform the task. If the patient can perform the task, controls must be introduced in
order to compare the deficit in the affected subject’s brain images with a control group. Tasks also only provide insight into the impairment of a particular task. If more than one region is affected (the most common case) a single task will not reveal the breadth of the affected brain regions. Resting-state data paired with a connectivity-based approach allows the clinician to investigate a wide range of debilitated regions without the use of invasive techniques or endless hours of uncomfortable scanning of the patient (Carter, Shulman & Corbetta, 2012).

The HCP is useful due to the breadth of information it covers. Blood drawing from every subject allows for extensive interpretations that are still correlational, but far more powerful. Due to the fMRI’s ability to look at a brain in vivo in a non-invasive manner, it has many uses beyond the intended uses of scanning healthy adults. For instance, the machine can be used to scan term-born infants. Because motion and noise is a problem for the scanner, special equipment must be used but it is very possible to get an in vivo image of a term-born baby without negatively effecting its development. In a paper by Ball et al. (2013) a team of scientists were able to utilize tractography and the data from the HCP to gain information on the effects of premature birth in the thalamocortical connective tracts. The study was able to utilize the immense precision of DTI (diffusion tensor imaging) to find diminished connections between the thalamus and the cortex (Ball et al. 2013). This finding can be used to explain many of the cognitive ailments of those who were born prematurely.

The HCP can also be applied to the field of pediatrics. The use of invasive techniques to investigate traumatic brain injury in adolescence introduces a risk of harming neural developmental growth. With the use of fMRI, pediatricians can gain insight to the damage of a patient’s brain without harming development. In a study conducted by Wilde and Bigler (2012), the effects of traumatic brain injury to the cortex of an adolescent were investigated. They found
severe impairment not only in the areas affected by the lesion but also to neighboring structures. As I stated previously, neural networks are incredibly complex, and damage to a single area of the brain is not isolated to the lesion alone. Neuroplasticity offers yet another complication in deciphering the breadth of the damage caused by a lesion early in development. The brain will adapt to the injury and form new paths that may perform the function of the lesioned path in an almost identical manner. However, the newly formed experience-dependent path may be maladaptive and slow. To investigate these effects in a non-invasive method, the rfMRI and dfMRI data acquired by the HCP lends itself to be an incredibly helpful tool. With the HCP, rehabilitation clinicians are able to get an image of the damage caused by a lesion early in a person’s development and investigate the possible damage to their connectonomics (Wilde & Bigler, 2012). Clinicians will be able to investigate the effects of a lesion patient much more effectively due to the data acquired from the HCP.

Because the HCP employs a vast amount of techniques such as EEG and genotyping to gather as much information on the human connectome as possible, it has far reaching applications. However, the HCP’s contribution to contemporary science is not just in its data, it also offers a new method of data-acquisition and distribution. The convenience of mass amounts of accurate and highly rich data allows the modern scientist far more time to think of solutions rather than collecting and analyzing data. This trend of mass data-acquisition is not limited to the HCP. Henry and Hohmann (2012) are using the idea of mass data-acquisition employed by the HCP to map genes heavily implicated in the development and maintenance of the spinal cord in mice. This mass distribution of information will allow easy access of the entire genome of the central nervous system in mice. Scientists will now be able to spend less time acquiring data and
more time conducting research and finding solutions to critical problems. The mass distribution of standardized data is an imperative contribution to modern science (Henry & Hohmann, 2012).

rfMRI has broad and wide-reaching applications in the medical field. However, it is very new and in the early stages of its development. Practical applications in the medical field have been found such as presurgical localization of brain regions (Lee, Smyser & Shimony, 2013). rfMRI’s biggest positive compared to task-based fMRI is its application for patients who are unable to perform the task. With rfMRI, a clinician will be able to localize affected brain regions and perform the surgery accordingly. Also, because motion is one of the largest problems for fMRI, this method of scanning proves to be very effective for patients with altered mental states as well. Through an analysis of the connectivity of the default mode network, clinicians are able to discover problems of consciousness impairment in patients and apply treatment accordingly (Lee, Smyser & Shimony, 2013). Resting-state analysis may prove to be a very powerful tool in medical as well as scientific research.

The HCP has affected many areas of science from pediatrics, to inspiring the mapping of the genome of the mouse’s central nervous system. Its implications are incredibly powerful for those in the scientific community through its mass distribution and easy access. The HCP has streamlined the process of fMRI data acquisition with the latest developed scanners and cutting edge data analysis. The project has not concluded, and will not conclude until 2016 yet has already seen promising effects in the scientific community. Hopefully the ideology of the HCP will spread further to more areas of science than just fMRI. With its use of mass distribution, and standardized methodology, the HCP may be the stepping stone to a new era of scientific discovery in the field of neuroscience.
References


Henry, A. M., & Hohmann, J. G. (2012). High-resolution gene expression atlases for adult and developing mouse brain and spinal cord. Mammalian genome, 23(9-10), 539-549.


The Role of Abstract and Concrete Representations in Multimedia Learning

Elham Zargarkalimi

Numerous researchers have spent a great amount of time examining the different methods by which learning can be enhanced, while presenting the learner with more information. In order to overcome the limited capacity of working memory and avoid overloading cognition, multimedia learning has been introduced. According to Mayer (1997), multimedia learning occurs “when one receives information presented in more than one mode”. Different presentation modes refer to the format in which the information is presented to the learner such as in written words or pictures. Both written words and pictures are processed visually. Another type of multimedia learning refers to having multiple sensory modalities presented at the same time. For example, this could be done when the information is processed not only visually but also acoustically. When using multiple sensory modalities, the split attention effect we may get when different types of information is presented in the same modality is reduced, and information processing can be more efficient (Mayer & Moreno, 1998). Techniques of multimedia learning have been increasingly studied, and researchers are constantly looking for more ways in which multimedia learning can become more effective by integrating other factors in it.

Recent research by Son, Doumas, and Goldstone (2010) examined the effect of adding verbal labels on learning, specifically on transfer abilities and relational reasoning. Relational reasoning is highly important in human cognition and is cognitively demanding, since it is based on the representation of the relations of the objects involved and bindings of these objects to
particular relational roles (Son et al. 2010). Based on previous studies, understanding relational similarities and acquiring novel abstract relations is a difficult task for students. To examine how relational reasoning can be facilitated by presentation of verbal labels, Son et al., compared two different training conditions in their experiments: a Word condition, which presented the relational labels in the tutorial, and a Control condition that did not include those labels. They trained participants about Signal Detection Theory (SDT), a method of understanding decision making under uncertainties, and tested their ability to reason relationally on a new transfer problem. In a SDT situation, categorical decisions are made based on the decision (e.g., signal vs. noise) and whether or not the signal or noise actually exists. The goal is to maximize the likelihood of deciding the signal exists, when it is actually present and to minimize the false alarms, which is deciding that the signal exists, when it actually does not.

Son et al. (2010) manipulated the alignment of stories presented (whether the tutorial story semantically aligned or comparable with the transfer story), and the alignment of relational words with the tutorial stories, creating four different experiments. Experiments 1 and 2 used aligned tutorial and transfer problems, in contrast to experiment 3 and 4. For example, in experiments 1 and 2, participants viewed tutorial and transfer problems that both involve semantically positive targets (such as healthy athlete or sweet melon) or negative targets (sick patient or infected melon), but in experiments 3 and 4, participants viewed non-aligned targets (e.g., positive target in the tutorial and negative target in the transfer problem). In addition, experiment 1 and 3 used relational words that semantically overlapped with the tutorial stories (e.g., “healthy athletes” as the positive target), as opposed to experiment 2 and 4 (e.g., “sick patient” as the positive target). It was predicted that if the semantic overlap of the words with the tutorials is crucial, experiments 1 and 3 would benefit from the relational words. It was also
predicted that the mismatch between the words and the referents would be harmful to the 
participant’s learning in experiment 2 and 4. In order to measure the participants’ performance, 3 
different tasks were done, a learning quiz to test the effect of words on initial learning, a transfer 
quiz, and an analogy quiz, in order to examine how participants can make connections between 
the tutorial and the transfer stories. The results of these experiments showed that presence of 
verbal relational labels produced better transfer, especially when the two situations were 
alignable. Also, mismatches between words and their labeled referents did indeed decrease 
learning.

How multimedia learning interacts with the visual presentation of learning materials is 
also under investigation. Moreno, Ozogul, & Reisslein (2011) examined how the students’ 
learning can be influenced by the different levels of concreteness of the visual presentations and 
written information, and how that may influence their problem solving abilities. To do this, 
Moreno et al. trained participants with electrical circuit analysis using concrete or abstract visual 
and written information. Concrete visual representation and problems’ cover stories refer to 
those that replaced the standard symbols used in physics by real-life objects, such as light bulbs 
and batteries. They predicted that concrete representations would facilitate learning because real-
life objects should be easier to connect to and be more accessible in long term memory. 
Moreover, concrete or realistic diagrams could promote learning by promoting interest and 
motivation in the student.

Moreno et al. defined the abstract visual representations and cover stories as those that 
use the conventional symbols and labels used in electrical circuit analysis, such as “resistors” and 
“voltage sources”. This type of abstract representation is predicted to support the process of 
knowledge transfer. Abstract representations can be better than concrete representations,
considering that concrete representations have more irrelevant information that could even
distract the learner by the superficial information. Therefore abstract visual representations can
potentially guide the learner to only focus his/her attention to the crucial information and spend
his/her cognitive resources on what is necessary to know and encode.

Moreno et al. were also interested in how effective abstract and concrete visual
representation when used together, combining both real-life objects corresponding to a problem’s
cover story and the conventional labels for representing the relevant information of the cover
story. They hypothesized that concrete representations promote problem solving abilities more
than abstract visual representations, and therefore students learning with both concrete cover
stories and concrete visual representations should perform better in problem solving than those
learning with abstract materials. Moreno et al.’s alternate hypothesis was that abstract visual
representations may promote better problem solving skills for students, since abstract
information can help them focus on the necessary characteristics of the problem and enable them
to generalize the problem’s structure for better transfer skills. Their third hypothesis stated that
the combined abstract and concrete visual representations would promote students’ problem
solving more than the concrete or abstract representations alone, since they offer the cognitive
advantages of both abstract and contextualized information.

In their experiments, participants were presented with abstract cover stories and abstract
diagrams (group A), concrete cover stories and concrete visuals (group C) or a combined group
(group AC), which was presented with concrete cover stories and a combination of both kinds of
the diagrams. In a follow-up experiment, another condition was added in which the participants
were presented with concrete cover stories and abstract visual representations. Learning was
measured by conducting near-transfer tests, which consisted of asking the participants to solve a set of novel electrical circuit problems by applying what they had learned in the training.

The results of this experiment demonstrated that the participants who received the instruction with a combination of abstract and concrete representations performed better on the transfer test than the students who were taught with only abstract or only concrete representations. Therefore, Moreno et al. concluded that it is most effective for novice students to be presented with both kinds of instruction, in order to guide their attention to what is necessary using the abstract representation, while the concrete representations help them in relating to the superficial features to their prior knowledge to promote their understanding. Moreover, with adding the second type of combination in the follow-up study, Moreno et al. found that group of participants who were instructed with simultaneous abstract and concrete representations outperformed the groups who learned with the concrete diagrams alone, in the transfer measure.

*Proposed Study*

This study aims to further examine the integration of abstract and concrete representations in multimedia learning by adding verbal explanations and animations. The study materials of this experiment, as well as the post-tests are adapted from the previous study mentioned earlier, by Son et al. (2010). During this experiment, the participants are presented with one of the four videos about Signal Detection Theory. These videos consist of abstract or concrete visual presentations (Av, Cv) that are paired with abstract or auditory narrations (Aa, Ca). More specifically the four videos consist of combinations of these manipulations, concrete visuals paired with concrete auditory (CvCa), concrete visuals with abstract auditory (CvAa), abstract visuals with concrete auditory (AvCa) and abstract visuals paired with abstract auditory
(AvAa). Each participant’s performance is tested by measuring their ability to transfer and apply the knowledge learned from the videos to novel problems.

Consistent with Son et al. (2010), I predict that instruction with abstract representations (AvAa, AvCa, CvAa) would result in better transfer compared to instruction with concrete representations (CvCa). Additionally, consistent Moreno et al. (2011), I predict that mixed representations (CvAa, AvCa) would promote better transfer than the consistent representations (AvAa, CvCa). Thus, I expect that combining different presentation formats and using different modalities can have unexpected outcomes, which could have potential implications for designing effective multimedia learning materials.
References


Perceptual Learning in Music

Erik Shiboski

As a research assistant in Human Perception Laboratory, I have spent the last ten weeks helping my graduate research advisor design a study that investigates the effect of perceptual learning on dynamic auditory stimuli, namely composer styles in classical piano music. Perceptual learning is characterized by changes in the ability to easily extract relevant information as a result of experience (Gibson, 1969). These changes are what differentiate experts and novices in their ability to process pertinent information, features, and patterns. In this paper, I will examine several studies that have implemented technology to facilitate perceptual learning of several abstract, complex concepts, patterns, and representations. I will then discuss the ways in which our purposed study will apply similar perceptual learning technology to train subjects to recognize abstract patterns in classical piano music.

In their study “Basic Information Processing Effects from Perceptual Learning in Complex, Real-World Domains”, Kellman and colleagues (2011) investigated process and encoding changes due to perceptual learning. They used the learning of Chinese characters to demonstrate changes in processing and ultimately provide evidence of perceptual learning. To facilitate perceptual learning, the experiment implemented technology to train subjects. Using a system called a Perceptual Learning Module (PLM), participants were trained to categorize Chinese characters. The system itself involves a series of trials that provide feedback and dynamically assess a participants level of mastery based on response time and accuracy. In the
training phase of the experiment, learners were trained specifically to distinguish characters at 3 different levels: stroke, component, and structure (Kellman et al., 2011). These tests involved visual search tasks and matching with novel stimuli that were never presented in the PLM phase. The study also included a control group in which participants judged the stroke count of characters.

To test if the system really was promoting perceptual learning, and not some other type of learning (through memorization of stimulus for example), participants were administered tests for basic information processing differences. The results indicated significant effects on visual matching and searching task efficiency from the PLM training, when compared to control condition. The results also showed certain effects that were related to specific levels of training within the PLM. For instance, structure training showed a significant improvement in test task performance when characters shared common structure across a given matching trial.

In another study investigating perceptual learning, Mettler & Kellman (2014) implemented “adaptive response-time based category sequencing” to test whether this would enhance learning of abstract categories. To explore the ways in which perceptual learning may be applied to learning and classification of real-world tasks, two experiments were conducted, in which participants were trained to classify 12 different butterfly genera. The stimuli were learned in a random presentation condition, an adaptive category sequencing condition, and adaptive category sequencing of within-category variability (meaning much less variability among exemplars for a given category) (Mettler & Kellman, 2014). In the first experiment, subjects were administered a pretest which involved random presentation of categories, and matching one of four images to the category label. The participants were then put through a learning phase, where they completed trials that involved matching images to genus labels. Two learning
One of the adaptive sequencing conditions implemented sequencing with ‘mini-blocks’, which involved presenting mini-blocks of exemplar butterfly images before the adaptive sequencing (which was identical to other adaptive condition). Subjects then completed an immediate post-test which involved matching tasks with novel stimuli. Participants were also given a delayed post-test one week after their first session. The second experiment repeated the procedure of the first experiment, but tested for learning using within-category variation (overall less variability).

Results from both experiments showed evidence of increased learning efficiency and accelerated learning for the adaptively sequenced conditions versus random presentation. Experiment 1 demonstrated the most significant effect size for efficiency advantage in posttest. Items in the delayed post-test showed a 29% efficiency advantage over those in random presentation (Mettler & Kellman, 2014). The adaptive condition also showed significantly higher accuracy on test trials relative to the random presentation. Experiment 2 showed the same learning effects, but to a greater degree. Efficiency was around 38% better in the adaptive sequencing condition. The results indicate that adaptive perceptual learning methods can help accelerate learning, and implicate real-world education and training applications for this type of learning.

The previous research shows the ways in which perceptual learning can accelerate and enhance learning efficiency. But these results were only produced in visual tasks, which involved static stimuli. The goal of the present research is to apply the perceptual learning module (PLM) and adaptive spacing, to learning composer style in classical piano music. Unlike static visual
stimuli, music presents patterns and relevant features over time. We want to explore whether perceptual learning can be applied to this non-visual domain.

In his paper “The Development of Sensitivity to Artistic Styles”, Garder (1971) discusses the properties of artistic style, and how it can be classified. “In making an art work or any other behavioral product, the individual is ‘creating’ an object in a style reflective of his intentions, values, personality, experience; in contemplating an artistic or other object, the viewer or listener ‘perceives’ numerous details and a global impression which may enable him to identify or classify the ‘style’ of that object” (Gardner, 1971, p 516). Music style can be classified in this way. Individual composers not only have their own personal style, but compose within the style of the musical period in which they worked. Gardner also points out the ways in which cultural style can regulate the differences between individuals. There are stylistic elements that constitute the period in which a piece was composed. Identifying these elements is an important part of classifying a composer’s style. But how can we train participants to classify the musical style of a composer?

In our study, we will be implementing PLM technology that trains for the recognition of 6 classical composers based on their composition style. We selected composers from the Baroque, Romantic, and Classical periods, and used only piano music selections. The series of learning trials will involve comparing musical excerpts, matching music to a composers name (or period), and receiving feedback for responses. We will implement PLM technology similar to that of Kellman et al. (2011) as well as the adaptive spacing techniques implemented in the Mettler (2014) experiments. Subjects will progress through the PLM based on accuracy. Based on an individual participant’s responses, the module will dynamically calculate the participant’s level of mastery. The study will also test subjects before and after the series of PLM learning
sessions to test ability to recognize novel clips from learned composers. Tests will involve matching a music clip to the composer’s name, with a total of 6 choices for each trial. Novel clips and composers will be introduced during the tests to control for memorization of the excerpts. We hypothesize that the training will show similar patterns of enhanced learning to that of visual perceptual learning found in previous research. Furthermore, the PLM will improve the subjects’ ability to extract relevant features and patterns to classify the style of the selected composers. This will demonstrate that perceptual learning can be applied to dynamic auditory stimuli, and could offer implications for further educational applications of perceptual learning. The training of verbal language is a real-world example of dynamic pattern recognition. Our research could offer applications for the training and learning of verbal language, and other dynamic stimuli (visual or auditory). By extending the implementation of perceptual learning to new kinds of stimuli, we hope to broaden the educational implications of previous research conducted by Kellman et al. (2011) and Mettler (2014).
References


Rational Numbers as Relational Models

Eugene Goh

Abstract

We propose that rational numbers can serve as relational models. Previous studies showed that mathematical and real-world relations guided people’s ability when solving arithmetic word problems and when constructing algebraic expressions. A fraction’s structure \((a/b)\), also known as its bipartite structure, encourages estimation or counting of the size of two sets. In particular, fractions are a more accurate counting strategy when the sets comprise of discrete elements. Discrete elements refer to elements that are not continuous. On the other hand, decimal notations encourages a one-dimensional estimate of a ratio, which typically leads to a procedure less accurate than counting, which is the procedure that fractions encourage. We constructed analogical reasoning problems that required subjects to map rational numbers (fractions or decimals) onto displays, which either depicted part-to-whole or part-to-part ratios between two quantities. Furthermore, we even varied the nature of the displayed quantities, which could be discrete, continuous, continuous but parsed into discrete components. When reasoning about discrete elements, fractions were more accurate than decimals. However, when quantities were continuous, accuracy was lower for both number types, with decimals being more advantageous. These findings support the general idea of mathematical thinking as a form of modeling, where mathematical formats serve as a bridge between the real world and quantitative concepts.
Experiment

Method

Participants

We had 58 undergraduate participants from the University of California, Los Angeles. The mean age for the participants was 20.4. Of the 58 participants, 49 of them were females. They were randomly assigned to the two between-subjects conditions in equal numbers. We provided participants with course credit for participating in the study.

Design and Procedure

The study was designed to be a 2 (number type: decimals vs. fractions) X 2 (relation type: part-to-part vs. part-to-whole ratios) X 3 (display type: continuous, discretized, discrete) study. The relation and display types were within-subjects factors, whereas the number type was a between-subjects factor.

Discrete items included circles, squares, stars, crosses, trapezoids, and cloud-like shapes. Continuous items included displays of rectangles that differed in width, height, and orientation (both vertically and horizontally). Discretized items were identical to the continuous items, except for the fact that in the discretized case, the rectangles were divided by dark-lines into equal-sized objects. Figure 1 below shows all three of the display types. The red and green colored items depict the stimuli used in actual test trials, whereas yellow and brown colored items were used in practice trials. We varied the color such that it represented both the larger subset and smaller subset. Half of the trials had the red subset as the numerator, and the other half had the green subset as the numerator. We used Superlab 4.5 on Macintosh computers to display these stimuli. Participants’ response times and accuracy were recorded.
We matched the magnitudes of fractions and decimals. The values of the decimals and fractions were always less than one, with decimals being rounded to two decimal places. The values of these rational numbers represented one of two ratio relationships: part-to-whole ratio (PWR) or part-to-part ratio (PPR). Figure 2 shows an example of one of these relationships.

Figure 1. Examples of continuous, discretized and discrete displays used in test trials.

Figure 2. The display above shows a part-to-whole ratio of 9/10 or 0.9. The corresponding part-to-part ratio of this display would be 1/9 or 0.11.

Results

A 3(display: continuous, discretized, discrete) X 2 (relation type: PPR vs. PWR) X 2 (number type: fraction vs. decimal) mixed factors ANOVA was used to assess differences in response time (RT) and accuracy. Only accuracy and mean RT on correct trials were computed for each participant in each condition. Figure 3 shows the results of accuracy identification for
the mixed ANOVA. The patterns exceeded 50%, which means it exceeded the chance level for all conditions. We found an interaction between display type and number type ($F(2, 55) = 24.57$, $MSE = 1.7, p < 0.001$). Planned comparison tests revealed that participants were more accurate when they used fractions rather than decimals for both discrete (81% vs. 63%; $F(1, 56) = 23.64$, $MSE = 8.2, p < 0.001$) and discretized displays (78% vs. 63%; $F(1, 56) = 13.92$, $MSE = 9.7, p < 0.001$). However, accuracy was no a significant factor of number types for continuous displays (62% vs. 67%; $F(1, 56) = 1.52$, $MSE = 10.4, p = 0.22$).

![Graph showing accuracy of relation identification across all three displays.](image)

**Figure 3.** Accuracy of relation identification across all three displays.

The mean RT is displayed in Figure 4. A significant interaction was obtained between number and display type ($F(2, 55) = 3.52$, $MSE = 1369, p = 0.037$). For the discrete displays, the difference for RT for fractions and decimals was not reliable (4.20 s vs. 5.97 s; $F(1, 56) = 2.40$, $MSE = 7607, p = .13$), even though RTs were generally faster for fractions than decimals. Discretized displays was also shown to be not reliable (3.93 s vs. 4.59 s; $F(1, 56) = 0.585$, $MSE = 4357, p = .45$). This means that we cannot attribute to accuracy advantage of fractions over
decimals to speed-accuracy trade-offs. Participants might have been taking a longer time in order to arrive at the correct solution to a particular problem. In contrast, participants were slower with fractions than with decimals for continuous displays (3.77 s vs. 2.88 s; $F(1, 56) = 4.77$, $MSE=9202$, $p = .03$).

**Figure 4.** Mean Response times for fractions and decimals across all three displays.

**Discussion**

The findings of this study showed that fractions are much better suited for discretized and discrete displays, both in terms of accuracy and RT. Fractions were shown to be advantageous for identifying ratio relationships when these ratios were represented by both fractions and decimals. DeWolf et al (2013) assert that fractions are better for representing the relations between two distinct sets because of their bipartite structure. In this experiment, the visual displays might have been a limiting factor as these results were shown only with these displays. Ratios were evaluated more accurately when the number was represented as a fraction rather than a decimal when displays conveyed countable entities, which consist of either sets of discrete objects or discretized objects. However, continuous quantities showed no such advantage for
fraction or decimals. Accuracy was similar in evaluating ratio relations, with decimals being slightly more accurate, but not significant enough to be considered advantageous. RT was also slightly faster for decimals than fractions for the continuous displays, but there was no significant advantage there as well.

These findings show the importance of relational structure that is internal to fractions. Participants did better in identifying ratio relationships using fractions because the values within the fraction were countable, which means that they could be mapped to particular entities. In contrast, it is much harder to map integrated decimals values to two separate subsets within a display, which highlights the slow RT and lower accuracy rates for discrete and discretized displays for decimals. However, people are forced to use some sort of estimation strategy when presented with a continuous display, as decimals had about the same accuracy as fractions, with lower RT. A possible explanation as to why people performed worse on continuous displays might be that adults exhibit misconceptions of the complex conceptual structure of fractions (Siegler et al., 2011; 2013; Ni & Zhou, 2005; Stigler et al., 2010). Another possible explanation is that students often find it difficult to understand how whole numbers that are within the fraction contribute to its overall magnitude (Ni & Zhou, 2005; Vamvakoussi & Vosniadou, 2010). Since continuous displays do not have the dark lines to show clear equal-sized separation of objects, students find it hard to determine an overall magnitude for the fraction. In comparison, this is exactly what decimals do, they show a one-dimensional magnitude, which makes it easier to determine the ratios to the continuous displays.
References


Perceptual-Motor Sequence Learning of General Regularities and Specific Sequences

Giancarlo Sanguinetti

Background

The authors of this study sought to investigate the underlying mechanisms between perceptual learning through self-generated motor sequences. They explain that this kind of mapping happens in a variety of real world skills, such as becoming proficient at typing. Marsolek & Field formalize this kind of relationship as perceptual-motor sequence learning, and define it as “learning to associate unfamiliar perceptual sequences with unfamiliar motoric sequences via familiar perceptual-motor elements that are juxtaposed to form the novel perceptual-motor mappings.” Pertinent research in this domain of psychology has involved the use of artificial grammar studies and procedural skill-acquisition studies. Learned information from these kinds of studies prompted the authors to differentiate between two types of perceptual-motor sequence learning: general regularity learning and specific sequence learning.

General regularity learning is the unconscious detection of a pattern from a set of stimuli that helps increase processing efficiency when prompted with new stimuli that adhere closely to the same pattern. Marsolek & Field used artificial grammar learning studies performed by Reber (1967, 1969) to demonstrate this type of learning, where participants were presented with nonsensical strings of consonants whose construction followed a specific order, despite never being explicitly stated. When later tested, participants processed new strings that adhered to the same rule more quickly than strings that adhered to a different rule or random rule.

Specific sequence learning is the effect of participants processing items in the test session that were present during the training more quickly than any other kind of item. Such evidence can be seen in digit-typing studies performed by Cohen et al. (1997) and Poldrack et al. (1999),
where the digit strings adhered to sequencing rules, much like the aforementioned artificial grammar study, and participants typed strings more efficiently if they were present during training than new strings that adhered to the same sequencing rules.

In light of all this research, Marsolek & Field were most curious as to explaining the underlying cognitive architecture that could produce these effects. Their primary question of interest was whether through the correct experimental design you could have participants place cognitive demands that could help delineate the general regularity learning effects from specific sequence learning effects. Specifically, they sought to create an experimental design that developed from previous digit-typing research (Cohen et al., 1997; Poldrack et al., 1999). To accomplish this, they designed a digit-typing experiment that enforced sequencing rules, which would help elicit general regularity learning. To determine if specific sequence learning could be separated from this general regularity learning, Marsolek & Field implemented an experimental condition that would force participants to remap their learned motor mechanisms from training to a new mechanism at test. This was put into effect by having the digit sequence typing occur on two different areas of the keyboard: the numerical pad, labeled the numerical key configuration, and the number keys above the alphabetical portion of the keyboard, labeled the horizontal key configuration. This manipulation asserts that disparities in test performance will be displayed between different types of stimuli because the key configurations will create a conflict between observable perceptual-motor sequences learning across different variations of the same stimuli. The resulting predictions from this hypothesis are: groups whose key configuration remained constant between training and testing should demonstrate both general regularity and specific sequence learning effects, whereas the group whose key configurations differed across training and testing should demonstrate a general regularity effect but no specific sequence effect.
Experiment 1

The experimental procedure consisted of three sessions: training, testing, and recognition task. All sessions took place in front of a computer that presented the various types of stimuli on the screen; participants were required to utilize a chinrest to maintain a controlled viewing distance from the screen and interacted via a standard keyboard. Experimental materials consisted of two sets of 576 four digit strings, none of which used the number zero or used the same digit twice in a singular string. Strings were created using a strict rule standard, wherein the first set (hereafter base rule) only four numbers could appear after the first digit (i.e. a 5 can only be followed by a 1, 2, 6, or 8) whereas the second set (hereafter opposite rule) used the other four numbers (i.e. a 5 can only be followed by a 3, 4, 7, or 9).

Training sessions involved the presentation of three sets of 180 trials for a total number of 540 trials. Two sets, labeled unique sequences, consisted of digit strings that adhered to a base rule, with every string being unique, and one set, labeled repeated sequences, consisted of 15 digit strings that adhered to that same base rule but would be presented 12 times during the training session. A trial consisted of a brief presentation of a digit string in a vertical orientation that would prompt the participant to type out each of the digits in order and then press the nearest enter key. Participants were instructed to use only their right hands, be consistent with the fingers they used to press each key, and not correct any errors they realize they made during a trial.

The testing sessions also consisted of three sets of 180 trials: the first set, labeled old rule unique, consisting of unique sequences that adhered to the base rule used in the training session, the second set, labeled new-rule unique sequences, comprised of digit strings that adhered to the rule opposite to the one participants were accustomed to in training, and the same set of 15 digit strings adhering to the base rule appearing 12 times each, labeled old-rule repeated sequences.
The experimental condition was whether or not a participant’s key configuration assignment during the testing session differed from the key configuration assigned during the training portion. Participants whose assignments did not differ between sessions were labeled the same-key configuration group, and those whose assignments did differ were labeled the different key configuration group. In addition, the testing session took place 48 hours after the training session, to minimize any recency effects from the training session.

The recognition task was a presentation of 30 trials which only required the participant to give simple yes or no judgments of whether the sequence had appeared in the previous two portions of the experiment. 15 of the trials, labeled targets, were the old-rule repeated sequences from the training and testing session, and the other 15, labeled distractors, were unique sequences that had not been presented previously, but adhered to the base rule from training.

A visual example of the different aspects of the experimental design is displayed below:
Results for the test session were measured through interkeypress intervals, error rates, and initiation latencies while results for the recognition task were measured through absolute score and response times. The most significant results, acquired through repeated measure ANOVAs, were that the same-key configuration group displayed both general regularity learning and specific sequence learning by typing old-rule unique sequences more efficiently than new rule unique sequences and typing old old-rule repeated sequences more efficiently than any other set. In contrast, the different key configuration group only exhibited general regularity learning because they typed old-rule unique sequences faster than old-rule repeated sequences.

However, the most surprising result was that both groups demonstrated strong recognition score because they responded to targets significantly faster than distractors, and were able to identify nearly 85% of all the targets accurately. The authors interpreted this finding as evidence that declarative memory had a strong influence in the specific sequence learning results. This led to the design of a second experiment that sought to investigate the relationship between perceptual-motor mechanisms from declarative mechanisms by eliminating the perceptual-motor component from the training session and observing what effects it would have on the testing session.

Experiment 2

For the purposes of this paper, experiment 2 will only be described as a near replication of experiment 1, with the exception that the training sessions required participants to speak the sequence aloud into a microphone rather than type them. All other aspects of the experimental procedure were identical to experiment 1.

Results for the test session were measured through interkeypress intervals, error rates, and initiation latencies while results for the recognition task were measured through absolute
score and response time. The most significant result, acquired through repeated measure ANOVAs, was that no general regularity learning was observed since old-rule unique sequences were typed no more efficiently than new-rule unique sequence, but specific sequence learning was demonstrated by old-rule repeated sequences being typed substantially more efficiently than the other two sets.

Conclusion

Marsolek & Field’s results serve as an exemplary model of good experimental design; their study demonstrated that even the most interdependent of cognitive skills can be disassociated given the right task. They observed general regularity learning effects when their participants typed sequences regardless of experimental condition, whereas the different key configuration group failed to demonstrate specific sequence learning effects when their same-key configuration cohorts could. In addition, when participants used verbal processes rather than motor processes, as in experiment 2, no general regularity learning was observed, but specific sequence learning was seen. These effects in total represent the strength of perceptual-motor sequence learning. By engaging dual modalities, motoric and visual, in training, participants could demonstrate unconscious processing of an abstract pattern when tested by the same modality. Even the immediate research could be developed by recreating experiment two with slightly manipulated conditions: having one group study with speaking sequences and then testing them through that same verbal modality contrasted with a group whose study session consists of typing sequences and then testing them through the speaking sequences. This experiment could help see if general regularity learning can occur across different modalities, so long as the experimental measure is held constant. Findings like these could also have a wealth
of application in different domains of perceptual-motor knowledge acquisition, such as music learning via instruments or even martial arts expertise.
References


Presence and Immersion: Bridging the Gap between VR and the Physical World

James Mutter

Abstract

This virtual reality experiment investigates the effects of immersion and presence on language learning. Immersion is defined as the virtual reality mimicking the physical world whereas presence is the degree to which a VR user identifies with his/her avatar. Research showed that VR might be a beneficial tool for learning, but it is unclear how presence and immersion facilitate learning. Fox et al. (2009) demonstrated that the VR can be used to impact real world behavior with presence as the mediating variable but only looked at eating behavior with little regard for user-avatar interaction (i.e., participants were not controlling the avatar). Although presence and immersion may be of importance, it is generally neglected in much of the current literature (Burgess, Maguire, and O'Keefe, 2002; Plancher, Barra, Orriols, and Piolino, 2012; Rose and Billinghurst, 1995). As both presence and immersion are dependent on the user, our goal was to measure these using a 10 question post-experimental survey to evaluate learning outcomes after users were allowed to navigate the virtual space with an avatar for a significant amount of time of ~3.5 hours.
Virtual Reality (VR) is a generated computer environment that can simulate space in a fictional or physical world. In order to differentiate other modes of displaying information, VR is unique in two main domains. VR allows the user to experience immersion which is the degree to which one experiences a virtual environment as if it were real. This allows the user to interact with the environment in as many ways possible to the physical world. Presence is the degree to which a VR user identifies with his/her avatar on a psychological level. Although presence and immersion may be of importance, it is generally neglected in much of the current literature (Burgess, Maguire, & O'Keefe, 2002; Plancher, Barra, Orriols, & Piolino, 2012; Rose & Billinghurst, 1995). Furthermore, of the research that does measure these characteristics (Fox, Bailenson, & Binney, 2009), little attention is paid to the amount of time the user is in VR to feel engaged. Researchers Ming, Ruan, and Gao (2012) validated that it took around two hours of using their VR technology for the learners to generally reflect that the virtual reality system gave them a sense of immersion.

Characteristics of VR have been demonstrated to affect behaviors outside of the virtual environment (Ratan & Hasler, 2010), including exercising (Fox & Bailenson, 2009) and eating behavior (Fox, Bailenson, and Binney, 2009). Furthermore, research has established how VR affects behavior inside the virtual world, including, affording users to give self-judgments regarding body size (Normand, Giannopoulos, Spanlang, & Slater, 2011), full body ownership (Kilteni, Bergstrom, & Slater, 2013) and feeling a social presence with other avatars through interactive communication (Jin, 2010).

Although a great body of literature has developed a means to express behavior inside and outside the virtual environment, this literature implicitly entails subjects need to feel present and immersed to some degree to reach full benefit. However, previous literature addressed how these
associations are overlooked and how no two researchers operationally define these concepts in the same light (Ratan & Hasler, 2010). Furthermore, there is no standard for post experimental surveys dealing with presence and immersion. Very few collaborative tests have been done with these measures to compare validity amongst questions and several researchers make their own questionnaire or draw from several different sources.

We aim to examine the relationship between VR learning and subjects subjective report of presence and immersion using Fox, Bailenson, and Binney’s (2009) post survey questionnaire. They conducted a ten item measure to assess participants’ experience of presence and immersion in the virtual world. The researchers demonstrated that the VR can be used to impact real world behavior of eating behavior with presence as the mediating variable. While they looked at how presence inside VR can affect behavior outside in the physical realm we are looking at memory retention outside of the VR. However, they only looked at eating behavior with little regard for user-avatar interaction (i.e., participants were not controlling the avatar). Our study differs in that the environment is interactive with users having the ability to move through space and drawing on past literature (Ming, Ruan & Gao, 2012), subjects will have ~3.5 hours of game play which should give them a higher degree of presence. This research aims to bridge the gap between the physical world and VR with which subjects connect to their virtual self by defining and measuring presence with regard to memory retention.

Method

Participants

Twenty three adults (6 males, 17 females, age range: 19-27 years) participated in this experiment. Participants were freshman, sophomore, junior or seniors from the University
California, Los Angeles as well as individuals living in Los Angeles. All participants in the experiment were compensated by being given cash or SONA credit.

**Design**

This research was conducted to order investigate the effects of domain type (i.e., presence or immersion) and gender have on memory retention. The experiment was conducted using several statistical methods including a mixed-subjects 2 (domain) x 2 (gender) factorial design, linear regression with presence and immersion as predictors, linear regression with presence as predictor and T-tests using high and low mean splits. All possible combinations of the independent variables were presented in this way. Immersion is the degree to which one experiences a virtual environment as if it were real. Presence is the degree to which a VR user identifies with his/her avatar on a psychological level. Each level of the domain independent variable contained 5 items for each condition. The dependent variable was the amount of correct recall of Swahili words on T5.

**Materials and Procedure**

The post survey questionnaire was adopted from Fox, Bailenson, and Binney (2009). They conducted a ten item measure to assess participants’ experience of presence and immersion in the virtual world. The first five items were used to measure presence and the second five items were used to measure immersion.

Participants were run by research assistants through a total of 5 testing and 4 learning sessions spread over a two day period. The first day lasts ~2.5 hours while the second day lasts ~1.25 hour. During the first day research assistants instruct participants how to maneuver the world using the joystick then allow subjects to explore the terrain and customize their avatar. Following this the subjects were told to watch an instructional video that states the nature of
procedures that will happen. This include exploration phases, learning phases and testing phases. Exploration phases are times in which subjects are allowed to explore freely for 5 min. Learning phases in which a VR representation of an object will appear with the Swahili name portrayed at the top and the English name on the bottom. A voice will repeat the Swahili name 3 times while the object rotates in which subjects repeat after for a total of 42 words. Participants are further instructed that they need only to recall the Swahili names. The testing session includes real world representations of the objects with no Swahili or English text. Participants are instructed to pronounce the Swahili word out loud and that only the last pronunciation is used for scoring. Each image appears for 10 sec and turns red during the last 3 sec.

After the instructional video is completed participants proceed to do 3 learning/testing sessions. The experimenter scored based on correct amount of syllable pronunciations. Each syllable in the word is worth 1 point. If the consonant or vowel is wrong, .5 points get taken off the final score for the word. If the whole syllable is swapped with another syllable then .5 points are taken off but if only half the syllable is swapped then only .25 points are taken off the final score. After the last testing session they are instructed to not look up words until the experiment is completed. The second day is composed of one learning and two testing sessions followed by the post experimental survey. Participants marked all 10 items on a 5-point scale ranging from “not at all” to “extremely.” Their responses are coded such that “not at all” was converted to a 0 while “extremely” was converted to a 4. The middle points were coded as well in order to conduct statistical tests.

Results

Across all statistical tests there were no significant results to be reported except for the 2x2 mixed subject anova which showed a significant effect for presence type ($p<.05$) but post
hoc tests did not show a significant difference for either effect when compared to T5 scores ($p>.05$). These mean tests were conducted in line with research by Fox, Bailenson, and Binney (2009). As an example, in order to examine the role of self-reported presence we performed a mean split in order to separate participants into high and low conditions. The scores that were below the mean ($M=2.6$) were categorized as low presence ($N=12$) while those above were categorized as high presence ($N=10$). One data point was thrown out as it fell on the mean.

**Discussion**

While the results are not conclusive by any means there are a couple important implications that can be interpreted from the data. First, there is great deal of controversy over gender effects on VR retention. 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 in relation to task performance. The study by Fox, Bailenson, and Binney (2009) claimed that their main effect for sex bordered on significance with women reporting less presence than men. 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 (Cassell & Jenkins, 1998). For this reason they may consequently felt less comfortable overall using the virtual reality environment within the context of the experimental setting. Our experiment controlled for this in that subjects that played second life or other games similar to ours were not able to participate. This was one of the first studies that controlled for this type of variable. We conclude that gender has no relationship with degree of presence or immersion.
Another interesting result is that are dependent measure (T5) was not influenced by either independent variable. Fox, Bailenson, and Binney (2009) ended up collapsing across scenarios to just measure presence in order to find significant results. While the results in our experiment were not significant that may be due to the problem of the questions asked. It may very well be possible that results would be found if using a different measure. The literature is slim regarding post experimental surveys pertaining to not only presence and immersion but also consistency between measures. However, there are few research papers addressing this issue, including Witmer and Singer (1998) who found consistent measures with high internal validity between both domains of presence and immersion. They interestingly measured immersion not in terms of VR implementation but rather in the form of the physical world. That is, they measured this to the degree of which an individual can experience immersion in day to day life. This includes questions asking how involved they get in terms of sports, reading, work et cetera. Using this method they found immersive tendencies predict the amount of presence that users experience in the VR. Furthermore, they conclude that there is a positive consistent relationship between task performance and presence (Witmer and Singer, 1998). The implementation of this more consistent measure may lead to significant results across both measures. Furthermore, this questionnaire uses many more questions to address both issues which will lead to more statistical analysis that can be done to find the mechanism behind presence and immersion on memory. At this stage we are developing a method to contribute to a seemingly new research area. As this progresses and we understand the parameters to meet the idea of interest we can possibly revolutionize the gap between the VR and the physical world.
References


Optimal Sequencing in Adaptive Learning

Jason Ku

The effects of spaced retrieval practice on learning have been extensively studied and the general consensus is that spaced practice produces better long-term results in learning than massed practice. However, the optimal spacing between repeated tests that produce the best results is still under investigation. Some candidates include: equally spaced retrieval sequences, fixed expanding retrieval sequences, random sequences and adaptive sequences. In equally spaced retrieval sequences, the spacing between all repeated tests for every item in the sequence is the fixed and the same for each item. For fixed expanding retrieval sequences (T.K. Landauer & R.A. Bjork, 1978), the spacing is initially very small and gradually increases with subsequent tests. A random sequence is exactly as its name implies; the spacing between tests for all items are random. Finally, adaptive sequences (Atkinson, 1974; Mettler & Kellman, 2010) are dynamically generated using an algorithm which factor in trial-by-trial accuracies and reaction times. Several studies have compared the effectiveness of the four different types of spaced sequences.

Karpicke and Roediger (2007) compared short-term and long-term retention of expanding retrieval sequences to that of equally spaced retrieval sequences. Participants studied vocabulary word pairs and were tested on them according to the two different schedules. In a series of experiments, they initially found that an equally spaced sequence produced better long term (tested 2 days later) results while an expanding sequence produced better short-term (tested 10
minutes later) results. This finding was inconsistent with the widespread belief that expanding retrieval usually produces better long-term learning results (Landauer & Bjork, 1978). In a subsequent experiment, they found that the first spacing for a learned item is critical. Specifically, delaying the initial retrieval attempt of a learned item enhances long-term retention, regardless of how the remaining tests were spaced. Since the original expanding retrieval sequence did not have a delayed initial retrieval, the spaced sequence which did have that initial delay, produced better long-term learning. They proposed that delaying the first retrieval increases its difficulty which is important for long-term retention (Karpicke & Roediger, 2007). These results indicate that increasing the difficulty of testing, by manipulating spacing, increases long-term retention of tested items.

Atkinson (1974) developed an adaptive sequencing algorithm which uses a student’s test accuracies and their overall performance history to dynamically generate a spacing sequence to optimize learning. Adaptive sequencing has been shown to produce better long-term retention than equally spaced sequencing as well as random sequencing (Mettler & Kellman, 2011). The Atkinson model of adaptive sequencing used a learner’s accuracies to compute the sequencing. Mettler, Massey and Kellman (2011) improved upon Atkinson’s model by developing a new adaptive algorithm called Adaptive Response Time Based Sequencing—ARTS, which factors in both a learner’s accuracy as well as their response times. The goal of ARTS is to optimize spacing between presentations by increasing or decreasing delay based on how quickly the answer was computed. A fast reaction time indicates a high learning strength and spacing is increased for that item to make it more difficult. Slower reaction times reduce spacing between presentations. Mettler, Massey and Kellman (2011) found that long-term retention, measured by learning efficiency (accuracy divided by total number of trials) was greater for ARTS system
than the Atkinson model. Participants in the ARTS condition showed 54% increased efficiency over participants in the Atkinson condition in the immediate post-test and 76% more efficiency in the delayed post-test. The results indicate that response times contain important information regarding learning strength and can be used to enhance learning.

Mettler and Kellman (2014 in preparation) compared the learning outcomes of the ARTS system to the outcomes of random sequencing. In this study, they utilized learning to criterion and dropout as new features to the adaptive learning sequence. Learning to criterion means that after a set number of successive correct recalls, an item would be removed from the sequence completely. This has been shown to increasing learning efficiency (Mettler & Kellman, 2014 in preparation). They found that long-term retention, measured by learning efficiency (accuracy divided by total number of trials) was greater for adaptive sequencing than random sequencing. This was true of the immediate post-test as well as the 1-week delayed post-test. Between random sequencing and adaptive sequencing, it is clear that adaptive schedules produce better long-term learning.

Landauer and Bjork (1978) have shown that expanding sequences lead to greater long-term retention than equally spaced sequencing. Mettler and Kellman (2014 in preparation) are currently comparing learning outcomes of the ARTS system to that of expanding sequencing to determine which produces the best learning results. In the current study, participants learn a number of science facts related to chemistry and chemistry nomenclature and are subsequently tested on their retention. Participants first take a pretest to access their chemistry knowledge. No feedback on performance is given in the pretest. They then begin the learning phase in which the chemistry facts are presented in either a continuously expanding sequence or in an adaptive sequence. During the learning phase, participants are given the correct answer after each
response. A post-test is administered immediately following the learning phase and again 1-week later.

In the expanding sequence, the order of presentations are fixed and predetermined. Due to the nature of expanding sequences, it is difficult to create a perfectly linear expanding schedule such as (1-5-9-14) without the use of filler items. For the purposes of the current research, a sequence was developed such that the early spacing between each item were held relatively constant while the later spacing were allowed more variation; the first few spacing of items were almost identical (1-5-9-...) while the subsequent spacing may deviate more from the ideal sequence (17-24-25-27). This was based on the idea that the initial delay between item presentations is critical to learning while later delays are less important (Karpicke & Roediger, 2007).

One effect currently under investigation is the systematic difference in reaction times between different types of questions. In the current study, participants enter their responses to either multiple choice or free response questions. Multiple choice questions are inherently faster to answer as they do not require typing. Since reaction times are critical in determining the spacing between item presentations, it is important to find any systematic differences in reaction times due to different question types. If this difference does exist and is substantial, the algorithm should be revised to reflect the differences to calculate more appropriate item priorities. Specifically, a linear or scalar adjustment could be made to equalize the means and distributions of reaction times between the problem types.

A preliminary analysis of 14 participants’ reaction times showed that RT’s were in fact significantly different between the two types of questions. Participants were on average about 850 millisecond slower in answering the free-response questions (M = 4998.7, SD = 706.4) than
the multiple choice questions (M = 4156.8, SD = 817.3), \( t(26) = 2.916, p = 0.007 \). This difference should be taken into account when calculating priority scores for the two types of questions.
References


Measuring Infants’ Abilities to Predict an Effect through Causal Inference

Jingxi Zhai

Abstract

The purpose of this experiment was to demonstrate how infants could detect patterns and anticipate events from a predictive visual sequence of three looming shapes. Infants between 9-to-18-months were tested to see if they are able to use causal information to anticipate outcomes, using eye-tracking technology to measure data. Infants participated in one of three conditions: ABC, BC, and C. Each corresponded with a certain number and position of shapes that would be indicative of a reward stimulus. If the sequence of shapes followed the rule, the trial was true and an attention getter (AG) appeared on the screen. If the shapes did not follow the rule, no attention getter is displayed. Latencies of infants’ eye movements toward the reward locations were analyzed. Over time, these latencies shortened, with infants looking quicker to the attention getter in true and false trials demonstrating infant ability to associate end of the sequence with the area. Effects of both infant age and complexity were observed and older infants displayed more anticipatory eye movements in true trials than in false trials.
Measuring Infants’ Abilities to Predict an Effect through Causal Inference

Even within the first months of birth, infants show abilities of statistical learning in more than one sensory modality (Kirkham, Slemmer, & Johnson, 2002). Infants are also able to detect co-occurrence frequencies between causes and effects in patterns of temporal coincidence (Griffiths & Tenenbaum, 2009) and by combining multiple sources of information about what is necessary for a result across repeated demonstrations (Buchsbaum, Gopnik, Griffiths, & Shafto, 2011). Through children’s ability to detect causal relationships and recognize patterns they may be able to learn to anticipate outcomes after viewing various conditions in event sequences that are consistently predictive and others that are not. Such learning may also be reinforced by “rewarding” correct predictions. Here, we examine these hypotheses by testing infants’ learning of “causal action” sequences.

To test this we showed infants various sequences of three looming shapes in one of three conditions. Each condition a specific rule determined whether the trial was true, resulted in a reward stimulus presented. In trials where the specific rule was violated, no reward stimulus was presented. For example, in the ABC condition, all three shapes must be in the correct sequence to result in a reward while the BC condition only required the last two shapes to be in the correct sequence to predict a reward. The C condition required only the last shape to be in the correct position to predict a reward.

The dependent measure used was the number of predictive gaze shifts (PGS), defined as directing the point of gaze away from the frame displaying the shape sequences toward the frame displaying the reward 200 milliseconds or less prior to the onset of the animation, suggesting anticipation. Using this definition, we inferred how well infants learned what sequences predicted the effect.
We hypothesized that older infants would perform better (e.g. produce more gaze shifts prior to the onset of the animation in anticipation in predictive trials relative to non-predictive trials) than younger infants, regardless of which condition. We also predicted that infants would perform best in the C condition, followed by the BC condition, followed by the ABC condition. This is due to the higher number of required tracking of conditional probabilities in ABC and BC.

Methods

Participants

The subjects for this study consisted of sixty-four 9- to 18month-olds who were recruited from a data of new parents. There were 33 males and 31 females who had an average age of 13.6 months. An additional 28 infants were tested but eliminated from data analysis due to excessive fussiness (25) or experimenter error (3). Testing was performed at the University of California, Los Angeles Baby Lab. Participants were given a piece of clothing as a thank-you gift in exchange for participation in this experiment.

Procedure

Participants were presented with 48 displays on a monitor while their eye-movements were recorded with an SR EyeLink 1000 (SR Research, Ltd) at 500 hz. The stimuli for the visual sequences were looming colored shapes that were created and animated beforehand and saved as video files. The reward effect for predictive trials was one attention getter animation from a set shown to attract infants’ attention. These stimuli and rewards have been shown to be distinguishable and attractive to infants (respectively) in previous visual statistical learning experiments (Kirkham et al, 2012).

Each display contained two grey frames. In one frame, a sequence of three looming
shapes appeared (750 millisecond per shape) follow by a 1500 millisecond inter-stimulus interval (ISI) after the third shape. The other frame was designated for the reward effect if applicable in that trial. Infants were randomly assigned to one of three conditions (ABC, BC, or C) with each corresponding to a certain number and position of shapes there were predictive of the reward.

For example, in the ABC condition, all three shapes must be in the correct sequence to result in a reward while the BC condition only required the last two shapes to be in the correct sequence to predict a reward. Five shapes were used which were a blue square, a red circle, a green cross, yellow triangle, and orange diamond. In each trial the shapes in sequence, which shapes corresponded to which position) was randomized.

Half of the total sequences presented were true trials (AG), which meant they predicted the appearance of an attention-getter, the reward in this experiment. The other 24 sequences were all false trials in which shapes were not presented in the correct sequence of that condition, resulting in no attention-getter. The order of true trials and false trials were randomized.

The dependent measure used was the number of predictive gaze shifts (PGS). This is defined as directing the point of gaze away from the frame displaying the shape sequences toward the frame displaying the reward 200 milliseconds or less prior to the onset of the animation, suggesting anticipation. Using this definition, as well as mean predictive latencies in true versus false trials, we inferred how well infants learned what sequences predicted the effect.

**Results**

Figure 1 presents the differences in predictive gaze shifts in true and false trials as a function of condition and age. Looking at the pattern of results displayed in Figure 1, there is a significant overall correlation between infants’ age and the difference between the number of PGS on true trials and PGS on false trials ($r = .28, p = .025$). Within each condition (Figure 1),
the correlation between age and the true vs. false PGS difference scores was statistically significant for the C condition \( (r = .48, p = .036) \), marginally significant for the BC condition \( (r = .39, p = .10) \), and nonsignificant for the ABC condition \( (r = .16, p = .45) \). The overall positive correlation between age and PGS difference scores supports our first hypothesis that older infants would perform better than younger infants. This finding suggests a developmental trajectory during the period of 9 to 18 months. The within-condition correlations suggest that development in this period is strongest for the C condition, followed by BC. The latter result provides tentative support for our second hypothesis, suggesting that infants are better able to make causal predictions when the number of conditional probabilities they need to track is relatively low.

Discussion

This experiment and model provides valuable data in evaluating developmental chances in infants between nine and 18 months of age. Based on our hypotheses and data, two developmental trajectories seem to appear around this age: shape identification increases with age and a bias exists that favors simple causes and shorter sequences. While infants are capable of complex statistical learning, these results suggest that tracking multiple conditional probabilities in predicting an effect involves additional abilities that require further development.

These results show promise in studying infant development beyond both statistical learning and causal perception research.

Acknowledgements

Primary investigators in this project are Scott P. Johnson and Bryan D. Nguyen. I made contributions to this study under their supervision. Figure 1 courtesy of Bryan D. Nguyen.
References


Figure 1. Differences in predictive gaze shifts in true vs. false trials as a function of condition and age. Points correspond to individual infants, plotted by their age and the number of their PGS during true trials minus the number of their PGS during false trials. Chance level performance rests at the 0-line.
Recognition memory involves two distinct processes: recollection and familiarity. Recollection refers to episodic memory retrieval about the instance of learning a piece of information. Familiarity, on the other hand, refers to a general strength of memory. Furthermore, recollection is a slow, controlled, threshold-based process whereas familiarity is a fast, automatic, signal-detection based process whose strength varies continuously. Receiver operating characteristic curves can be used to identify the degree to which each process is being used in a given task. As I will discuss, this information can then be used to test hypotheses about the nature of recollection and familiarity, such as those posed by the dual process signal detection theory and the dual process hypothesis of memory and aging.

In “Recollection and Familiarity: Examining Controversial Assumptions and New Directions,” Yonelinas et al. (2010) discuss the dual process signal detection model and how it can be used to distinguish between recognition and familiarity memory. The DPSD model further distinguishes recollection and familiarity by posing that recollection uses a threshold model and involves the hippocampus, whereas familiarity uses a signal detection model and results from repeated neural processing in regions outside the hippocampus (Yonelinas et al., 2010). In this paper, the authors identify and evaluate the three most controversial aspects of the model, which are, “(i) the two retrieval processes differ in the sense that recollection is a threshold process, whereas familiarity is a signal detection process; (ii) familiarity can support accurate associative and source recognition under certain conditions; and (iii) the hippocampus is critical for recollection, but not familiarity.”
The first way, and the way that is most critical to this discussion, the authors evaluate the assumptions of the DPSD model is by looking at item recognition tests in which subjects must discriminate between items they previously studied and new items and rate their confidence in their assessment on a six point scale with one being the item is definitely new and six being the item is definitely old. If the subject were using recollection memory, we would expect a high confidence rating, whereas if the subject were relying on familiarity, we would expect a wider range of confidence ratings. Data of this sort can be used to generate ROC curves. A curved, symmetric ROC curve would indicate the use of familiarity, a hockey-stick shaped ROC would indicate the use of recollection, and a combination of the two would represent a mixture of recollection and familiarity. Given a particular ROC, one can use a quantitative measurement model to identify the amount of familiarity versus recollection present. In reviewing the literature, the authors found that tasks designed to use more familiarity yielded more curved, symmetrical ROCs. Additionally, tasks designed to use more recollection yielded more asymmetrical ROCs. This supports the DPSD model’s distinction between recollection being threshold based and familiarity being signal-detection based.

In “Recollection and Familiarity in Recognition Memory: Adult Age Differences and Neuropsychological Test Correlates,” Prull et al. (2006) discuss the differences seen between older adults and younger adults in the strength of recollection and familiarity memory measures. The dual-process hypothesis of memory and aging predicts that older adults exhibit deficits in recollection, but not in familiarity when compared to younger adults. Prull et al. used three different methods to test this hypothesis and found mixed results. They consistently saw a decline in recollection in older adults, which is consistent with they hypothesis, but the familiarity results were not consistent across methods.
In the inclusion/exclusion task, participants studied lists and then were presented with words that were either in the studied lists or not. Participants were asked if the word presented was from the list they studied. In the remember/know task, after recognizing an item, participants then had to state whether they remembered specific information about learning the item (which would indicate the use of recollection), or if the item simply felt familiar (which would indicate familiarity). In the ROC conditions, like in Yonelinas et al. (2010), participants had to rate on a six point scale their recognition confidence with one being very sure new and six being very sure old. ROC curves were created and then used to measure recollection and familiarity where recollection was represented by a probability and familiarity was represented by the signal detection measure $d'$. Results from both the inclusion/exclusion and ROC tasks showed impairment in recollection but not familiarity in older adults, which supports the hypotheses of the DPHMA. However, the remember/know condition found that older adults were impaired in both recollection and familiarity. These results show that aging might have a broader impact on memory function than originally thought.

A study by Castel et al. (2011) also looks at how memory changes with age. This study looked specifically at the value-directed remembering paradigm. In this paradigm, subjects are presented with lists of items, each of which has a point value, and are instructed to remember as many items as possible in order to maximize their score. The best strategy is to selectively remember the high value words, which requires “attentional control, goal maintenance, and inhibition of less-relevant information” as well as metacognitive awareness of one’s memory capacity. Castel et al. found that while older adults recalled fewer items than younger adults, they maintained high selectivity indexes. This suggests that while their memory capacity might be impaired, their cognitive control is not.
These three studies show how recognition memory can be broken into components and that these components may change differentially over time. Yonelinas et al. (2010) used ROC curves, among other methods, to find support for the dual process signal detection model that identifies recollection and familiarity as threshold based and signal-detection based phenomena respectively. Prull et al. (2006) also used ROC analysis in analyzing the dual process hypothesis on memory and aging. They supported the prediction that recollection exhibits increased impairments with increased age, but they found mixed results regarding the effect of age on familiarity. Castel et al. (2011) used the value-directed remembering paradigm to test for changes in memory throughout the lifespan. They found that older adults exhibit decline in memory capacity as exhibited by the ability to recall fewer items. However, older adults did not see declines in selectivity, suggesting their cognitive control and metacognitive awareness was still high.

My Experience in the Lab

I have very much enjoyed my quarter spent working in the lab. While running subjects is not the most thrilling task, I have found the experience valuable. Having spent the past four years learning about various psychological experiments, it has been very interesting to be directly involved in the research process. I have a better understanding and appreciation for the work that goes into conducting research as well as a better understanding of some specific experimental methods.

The most valuable experience I had in the lab was working with the older adult data and running the ROC analysis. At first, I was simply going through the motions, not completely understanding what I was doing. However, after reading through the articles by Yonelinas et al. and Prull et al., I gained valuable insight into the meaning of all of the numbers and curves I was
generating. Furthermore, I gained a better overall understanding of what the goals of the study were. Additionally, the process of drafting a PURC abstract about the results was similarly insightful. While it did not lead to an actual submission, I learned so much by going through the drafting process.

One thing about this quarter that was slightly disappointing was the amount of cancelled lab meetings. Being new to the lab, I do not know if this is normal or not, but I would have loved to been able to attend more meetings. Even though I didn’t always understand everything being discussed, the topics were fascinating and I enjoyed simply being able to sit and listen. Though if there were papers I could be pointed to that would provide background information about the topics discussed, I would be interested in reading them so I could get even more out of the meetings. Overall, this quarter has been a great experience and I look forward to next quarter.
References


Animation Understanding and Problem Solving

Kevin L. Ruiz

Abstract

The current ongoing study explores the role of animations in analogical problem solving. The purpose of the study is to show if people can solve analogies with animations that are situated within a context. The current study had participants solve a target problem after being shown three different scenarios. Additionally, the way the scenarios were presented varied between the participants – the participants either received a full animation, a low frame rate animation, or a static diagram. As previous studies suggest, it is expected that participants who receive the full animation to have a higher transfer rate to the target problem. The results show that there seems to be a trend between the conditions without the hint, but these are non-significant.
Animation Understanding and Problem Solving

Analogies allow people to see commonalities between bodies of information that seem different than each other by allowing people to see underlying relationships. By noticing the underlying connections, a person can apply the same solution from one body of information to a novel one through a process called mapping (Gick & Holyoak, 1983). The body of information one refers to is called the source analog, which provides information that can lead to understanding to a new body of information, or target analogs (Holyoak, 2012). Although feature similarities aid people to understand analogies, research in the field focus whether people understand the schema, or “core idea”, driving the analogs.

A reoccurring analogical problem that illustrates the power of analogies is the tumor problem (Duncker, 1945; Gick & Holyoak, 1983; Pedone, Hummel, & Holyoak, 2001; Holyoak, 2012). The tumor problem is a scenario in which the participants are asked to pretend to be a doctor trying to 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 rays also affect the healthy tissue surrounding the tumor. The tumor problem can be found in appendix A. With these instructions the tumor problem has a goal, resource, constraint, solution, and outcome. The goal is to kill with the tumor with the resource of the rays, but the constraint is that the rays cannot be administered from one direction because it will damage the healthy tissue surrounding the tumor. When this problem is given alone the solution and outcome cannot be figured out; in the original study only 10 percent of the participants were able to solve it (Duncker, 1945).

Adapting the tumor problem, Gick & Holyoak (1983) replicated the results from Duncker (1945), but also added another condition in which the participants received a story preceding the tumor problem. This story is called the military story (Gick & Holyoak, 1980). Since it comes
before, the military story is the source analog and tumor problem serves is the target analog. In
the military story, a general wanted to defeat a dictator that is in a small fortress. He had the
options to conquer the fortress by sending down his army down one path or dividing his army
into small groups and send them down multiple paths. The general realized that if he we went
with the former option, his army would thin out and be defeated. Thus, he best choice was to
choose the latter option.

When given the military story, participants are able to use the same strategy to destroy
the tumor problem (Gick & Holyoak, 1983). Like the tumor problem, the military story has a
goal, resource, constraint, solution, and outcome. The goal is to use army to capture the fortress,
with a large army, but the constraint was that the general could not send the army down one road
because it will result in defeat. As given by the story, the solution the general used is to send his
army down multiple roads simultaneously to get the desired outcome of capturing the fortress.
After reading the military story, 30 percent of the participants were able to solve the tumor
problem without a hint; when given a hint to think about the military story, an additional 50
percent solved it (Gick & Holyoak, 1983). As the results show, their participants were able to see
that solution to the tumor problem was to administer low-intensity rays from multiple directions,
to get the desired outcome of destroying the tumor by rays.

As illustrated by Gick & Holyoak (1983), both problems have a common goal, resource,
constraint, solution, and outcome. Known as the convergence schema, participants had to
understand that the analogs have the goal to use force to overcome a central target, with
sufficiently great forces, but with constraint of not being able to apply force along one path. The
only solution, then, is to apply weak forces along multiple paths simultaneously to reach the
outcome of overcoming the target by force. If, then, the participants solved the tumor problem
after being presented the military story, then they probably have some understanding of the convergence schema.

In the same study, Gick & Holyoak (1983) studied analogies by providing diagrams as a source analog besides a passage. The participants had success – combining the total of participants that received hint and without hint, 67 percent solved the tumor problem. Adopting this method, Pedone et al. (2001), went further by using dynamic diagrams in addition to static diagrams. The diagrams had objects converging into a circle either by a single stream or multiple streams surrounding the circle. The objects were arrows converging on the circle, arrows diverging from the circle, or black boxes going towards the circle. For static diagrams objects was a picture, while in the dynamic diagram the objects were moving towards the circle like an animation. Participants were separated to receive one of the three objects and one of the two movements. After being presented with the source analog the participants were then presented with the tumor problem.

Pedone et al. (2001) found that the participants that were able to solve the tumor problem more were those with arrows converging on the circle. More specifically, the participants that received dynamic diagrams with converging arrows solved the tumor problem with a total of 55 percent before the hint and an additional 35 percent after the hint. Fifteen percent of the participants who received the static diagrams with converging arrows solved the tumor problem before the hint, and an additional, 55 percent after the hint. Although both of these groups did not differ, Pedone et al. (2001) shows that using dynamic diagrams may be useful for analogical problem solving.

The current study adopts a similar method as the Pedone et al. (2001) study by using diagrams as a source analog for the tumor problem. However, in this study we extended the use
of diagrams by situating it into a context by creating a cover story. The purpose of the study is to show if people are able to solve analogies with the dynamic diagrams situated in a context. To compare this the experimented is designed to show participants either a diagrams in full animation, animation with low-frame rate, or a static diagram with a verbal explanation. Given what’s been found, it is expected that animations that are more perceptually fluent will lead to better mapping and transfer, or solution, rates. That is, it is expected that participants that receive a full animation will solve the tumor problem more than the other groups.

Method

Participants

Thirty-three undergraduate students enrolled in classes at the University of California, Los Angeles (UCLA) participated in the experiment as part of a course requirement.

Materials and Design

The experiment was designed to have participants solve the target after being shown the source analogs, which were three different scenarios.

As seen in cover story, which can be found in appendix B, each scenario had cannon(s) attempting to destroy a circle that was situated within a circular barrier by shooting at it. When the cannon(s) shot at the circle the barrier also gets damaged. The point of the cannon(s) was to destroy the circle without destroying the barrier first. In the first scenario, one cannon was shooting at the circle from the right side. In this scenario the cannon failed to destroy the circle because the barrier around it was destroyed first. In the second scenario, one cannon was also shooting from the right side at the circle, however the blast of the cannon was a lot smaller than the one shown in the first animation. Much like the first scenario though, the cannon failed to destroy the circle because the barrier surrounding it was destroyed first. Lastly, in the third
scenario there were multiple cannons surrounding the circle shooting at blast at a low intensity as in the second animation. Different than the first two scenarios, the cannons destroy the circle without destroying the surrounding barrier. To ensure the participants understood the scenarios participants were given asked on why each of the scenarios failed or succeeded.

The way the scenarios that were presented differed between the participants. There were 3 conditions that participants could have been placed in. In condition 1, scenarios were received as an animation. In condition 2, scenarios were shown in a low-frame rate animation. Lastly, in condition 3, scenarios were received as a static diagram a verbal explanation of the diagram was given.

The tumor problem was adopted from and served as the target problem in this study. As mentioned above, the tumor problem placed the participant to pretend to be a doctor trying to cure a tumor with a tumor-killing ray without destroying the healthy tissue surrounding it.

Procedure

Participants were first placed into one of the 3 different conditions. They were then placed in front of a computer in a private office space and the experiment gave a brief overview of the experiment. The experimenter said the participant will experience three scenarios, one after the other, and will be answering a set of questions based on the scenarios. Also, the experimenter asked the participants to read the additional instructions on the computer screen and notify the experimenter when the computer program prompts to. The experimenter started the program and left the room.

If participants were placed in condition 1 or 2, they received the scenarios on the computer after the instructions were given. If the participants were placed in condition 3, however, they were prompted to call the experimenter back in the room to administer the
scenarios verbally. For condition 3, participants were given the diagram of the first scenario and were given time to ask any questions if they had any. The experimenter, then, gave the explanation of the first scenario. The same was repeated for the second and third scenario.

Once each participant experienced all three scenarios, the experimenter gave the participants time to explain each scenario on three separate sheets of paper. Along with a diagram of the first scenario, the first sheet of paper had a prompt, which asked the participants to describe the first scenario. The same was repeated for the second and third scenario. After answering the sheets, the participants were given a distractor task, which was a condensed version of the Autism-Spectrum Quotient (AQ).

Following the distractor task the participants were given the tumor problem without the hint. Regardless if they solved the tumor problem the participants, then, received the tumor problem with the hint. Participants were told to write any new solutions, if any, with the given hint; however, they were also told that they could repeat a solution. The experiment concluded by asking the participants two questions, which asked them if they have seen the tumor problem before and if they noticed the analogy between the scenarios and the tumor problem.

**Results and Discussion**

So far, there are 11 participants in each condition. For the first condition 10 participants solved the tumor problem without the hint. When given the hint, 10 participants solved the tumor problem. One participant failed to solve the tumor problem with and without the hint. For the second condition 7 participants solved the problem without the hint. When given the hint, 10 participants solved the tumor problem. One participant failed to solve the tumor problem with and without the hint. For the third condition 6 participants solved the tumor problem without the hint. When given the hint, 8 participants solved the tumor problem. Three participants failed to
solve the tumor problem with and without the hint. It looks as if the participants in the first condition had better transfer rates, but none these results were significantly different. Thus, we cannot confirm the given predictions.

It looks like the trend that might turn into something significant when there are more participants, which the study has recruited. If results do become significant with the analysis of the additional participants, it may be safely concluded that people are creating an understanding of the convergence schema. More importantly, it would show that animations are effective as source analog. This is important because it shows that people can be aware of surroundings and attend to relevant information to solve problems. Although people may not be aware of it in the moment, when they are solving problems they can refer to a relevant previous situation.

Even though the results may end up being significant, there are still some shortcomings within the experiment. The results show that participants in the low frame rate condition nearly solve as much as the full animation condition. This similarity may have been caused by the fluency of the low frame rate animation. Compared to the regular animation, there may have not been much of a difference, which makes both conditions the same. To make a comparison the frame rate in the low frame rate condition can be lowered making the animation look more scattered. Another improvement this design can do is elongate the time of the distraction task. Given that people are not presented with a problem soon after receiving the source problem in real life, future studies that adopt the method used in this study can implement a delay.

Overall, the study is still going on, and results still need to be finalized. With the analysis of data it is still expected that participants that receive a full animation will solve the tumor problem more than the other groups.
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 cannons are attacking an enemy circle. The enemy circle is surrounded by a friendly barrier. The object of the cannon(s) is to defeat the enemy circle without critically damaging the surrounding friendly barrier.

As the enemy circle is damaged by the friendly cannon, 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 cannon(s) cannot damage the enemy circle without also increasing the temperature of the friendly barrier. The objective of the cannon(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 cannon(s) fail their objective. The objective of the cannon(s) is to fill the ‘Circle Damage’ bar before the ‘Barrier Temperature’ bar fills up.
GPS Data and Autobiographical Memory

Laryssa Storozuk

This quarter, I have continued to work in the Rissman Memory Lab, aiding graduate student, Tiffany Chow, with her “Photographic Life Logging” study. This study has focused on the use of innovative cameras, called Autographers, that participants wear around their necks for a duration of three weeks. The cameras capture events from participants’ daily lives, utilizing a sensor that automatically takes pictures each time it detects changes in ambient lighting, motion, and temperature, as well as encodes GPS data for the participants’ current locations. An overarching goal of this study is to create distinct event sequences from the collected photographs and to observe fMRI brain activity associated with these autobiographical events. These sequences are comprised of eight photographs, and depict an event that occurred within a fifteen-minute timespan. The three conditions we have focused on thus far are to analyze how brain activity is affected by viewing images from one’s own life versus someone else’s life (Self vs. Other condition), how it is affected by any temporal alterations made to a sequence of images (Match vs. Mismatch condition), and how previewing image sequences affects memory retrieval (Previewed vs. Non-previewed condition).

An additional focus of this quarter has been to begin the process of extracting the GPS data encoded in these photographs with the ultimate goal of utilizing the data to compare brain activity for “long” and “short” distances within these event sequences. Thus far, to the best of our knowledge, there have been no studies that utilize GPS data from autobiographical photographs in the context of memory research. We are thus intrigued to explore this untouched realm of memory research, in the hopes of exploiting the potential uses of the cutting edge technology embodied in the Autographer.
However, as is often the case in the first generation of a particular technology, the GPS sensor on the Autographer is rather inconsistent. In order to geotag the images, the sensor must have a satellite lock (displayed by the word ‘Locked’ on the camera’s LCD screen). The ability of the camera to achieve this ‘lock’ is dependent on a variety of factors, and although it is recommended that one holds the “Autographer still, away from buildings, with a clear view of the sky,” we do not intend to request that our participants strive to achieve these conditions, for we want to minimize any inconvenience or unnatural behavior imposed by their wearing of the camera (Autographer, 2012).

Due to the great variability in the sensor’s encoding of GPS data, our first topic of interest was to quantify exactly how many images contained latitudinal and longitudinal coordinates from the event sequences we created for each participant. I worked with a fellow Research Assistant, Brittany Medina, to create a MATLAB algorithm that could extract the data, and then convert the coordinates from the format “Degrees Minutes Seconds” into decimal degrees (e.g. “34 4 13” to “34.0703”). Additionally, for each participant, we organized the GPS data into a matrix, in order to calculate the number of sequences, as well as the number of images, that contained GPS data. Thus far, we have implemented the program for the initial two participants. For the first participant’s event sequences, we determined that 254 of the 960 images contained GPS data (~26.46%), and 46 of the 120 sequences contained GPS data (~38.33%). For the second participant’s event sequences, 109 of the 960 images contained GPS data (~11.35%), and 21 of the 120 sequences contained GPS data (~17.5%). We also created an algorithm to determine the mean number of images per sequence that contained GPS data (~5.52 for the first participant, and ~5.19 for the second participant). We currently have twelve participants (6 male, 6 female; average age: 20.4 years), all of whom are UCLA students. We
plan on running the program for the remaining ten participants in order to get an accurate sense of the average amount of data encoded per participant with which we have to work.

Our next goal was to determine the distance traveled within any particular event sequence. To begin with, we took the latitudinal and longitudinal coordinates of one of the first subject’s event sequences (s01_sequence_007), and utilized an online algorithm to determine the distance between the coordinates of the first and seventh images (because the eighth – and final – image did not contain GPS data). We chose this particular sequence because the participant traveled in a relatively linear path, so the distance between the first and final images appeared to be reasonably representative of the total distance the participant traveled. However, there are a substantial number of event sequences, from each of the twelve participants, in which the participant travels to a number of different locations, only to return to the original location depicted in the first image of that particular sequence. In such cases, utilizing GPS data from only the first and last images would not be indicative of the total distance travelled within the sequence. We are currently working to create a paradigm that will overcome this complication. For instance, we may decide to code an algorithm that will determine which images contain GPS locations that are the farthest apart, within a sequence, and then use the data from those two images as the construct for our variable of distance.

We are particularly interested in using this GPS data to explore notions such as how distance information could affect memory encoding of events, the role that medial temporal lobe regions have in spatial navigation, and whether cognitive maps preserve real-world distance relationships. In an article titled, “Distances between real-world locations are represented in the human hippocampus,” Morgan et al. (2011) explored the latter of these topics. To summarize its methodology, fifteen University of Pennsylvania students, who each had at least one year of
experience with the University of Pennsylvania campus, were shown color photographs of ten prominent landmarks from the campus (Morgan et al., 2011). They were shown 22 images for each landmark, viewing a total of 220 images (Morgan et al., 2011). In the scanner, the subjects were asked to “covertly identify each campus landmark,” and would make a corresponding button press once they had done so (Morgan et al., 2011). Of crucial importance to this task was that the subjects were not told to “physically or mentally navigate between landmarks,” and were solely performing an identification task, at least on a conscious level (Morgan et al., 2011). As a result, Morgan et al. suggested that the “entorhinal cortex encodes metric information about the spatial relationships between landmarks,” as opposed to fMRI activity in the hippocampus that appears to be modulated by “the extent to which the current stimulus is consistent or inconsistent with these spatial relationships” (2011).

Morgan et al. posited several hypotheses as to the underlying mechanisms that are responsible for this “distance-related signal” in the hippocampus (Morgan et al., 2011). In my opinion, their most convincing argument was what I will refer to as the ‘context hypothesis.’ In this ‘context hypothesis’, they postulated that each image of a familiar campus landmark may have “established a ‘context’” and that the “hippocampal response on the immediately subsequent stimulus might then reflect the degree to which the new landmark violated this context” (Morgan et al., 2011). Therefore, they attributed decreased activity in the left anterior hippocampus - for successively viewed proximal landmarks - to lesser contextual changes, and, conversely, they attributed increased activity in the left anterior hippocampus - for successively viewed distal landmarks - to greater contextual changes (Morgan et al., 2011).

For our study, we could potentially implement this ‘context hypothesis’ to observe the effects of left anterior hippocampal activity dependent on whether participants viewed successive
event sequences that contained more similar or distinct contextual characteristics. Since our participants are all UCLA students, the majority of the images collected from their time wearing the camera depict similar views of quintessential campus locations, such as the Bruin Bear, Ackerman Student Union, and Royce Hall. Therefore, we could analyze the fMRI activity in the left anterior hippocampus when showing a subject an event sequence that occurred near the Bruin Bear immediately followed by an event sequence that occurred near Ackerman Student Union (which would be defined as proximal landmarks that are contextually similar). We could compare this brain activity to that elicited when showing a subject an event sequence that occurred near the Bruin Bear immediately followed by an event sequence that occurred near Royce Hall (which would be defined as distal landmarks that are contextually distinct). According to the ‘context hypothesis,’ we would expect to find less left anterior hippocampal activity for the Bruin Bear-Ackerman Student Union condition, and greater activity for the Bruin Bear-Royce Hall condition.

One potential pitfall of this study is that Morgan et al. (2011) failed to define their criteria for determining a landmark beyond the statement that their stimuli consisted of photographs of “buildings and statues from the University of Pennsylvania campus.” A study by Chan et al. (2012) chose to focus on defining what features are most indicative of an object or location gaining status as a landmark, and 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. Thus, if we were to use the idea of landmarks and the effects of their contextual similarities or differences on hippocampal activity, it would be imperative that we provide a formalized criterion for what could be considered a landmark.
Finally, a study by 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). In relation to this study, we could analyze changes in fMRI activity in the right hippocampus dependent on how long a particular participant has attended UCLA, as well as how familiar an individual participant felt with specific areas of campus, such as Royce Quad, Wilson Plaza, Bruin Plaza, etc.

This next quarter, I look forward to continuing to work with the GPS data we have collected to explore ways in which we can incorporate it into our study - possibly using the aforementioned ideas. I also look forward to beginning to work with the fMRI data to ensure that all analyses are valid and accurate.
References


SGLT2 Activity in the Rat Brain

Liseth A. Magana

Abstract

Sodium-dependent glucose transporters (SGLTs) are energy-dependent sugar transporter with the important function of absorbing sugar in the intestine and in the kidney, thus exerting an essential role in glucose homeostasis. More recently, the presence of functional SGLTs have been shown in the hippocampus and other regions in the rat brain including the forebrain and midbrain. The hippocampus is the brain structure primarily affected by Alzheimer’s disease and other forms of dementia. As of 2014, one in nine people aged 65 or older were diagnosed with Alzheimer’s disease. This number is projected to increase as the population ages and life expectancy increases. Therefore, research on the possible role SGLTs in neuronal health is essential. In this study, we report that Na⁺-glucose cotransporter (SGLT2) may be expressed in the functionally important areas of the brain including the forebrain, thalamus and hippocampus.
**Background**

Facilitated diffusion and active transport represent two mechanisms for glucose transport across the plasma membrane and into the cell. GLUTs are facilitated transporters that move glucose molecules across their natural concentration gradient. SGLTs are sodium dependent active transporters that utilize energy derived from the sodium gradient across cell membranes and use it to transport glucose into cells (Wright, 2011).

The functionality of both GLUTs and SGLTs is understood in glucose reabsorption in the kidney. Deficiency of SGLT2 has been linked to familial renal glycosuria, while patients affected by the rare genetic disorder Fanconi-Bickel syndrome (FBS), which is characterized by excess glucose in the urine, are known to have a premature stop codon for the GLUT2 gene. SGLT2 is required for luminal uptake, while GLUT2 is essential for basolateral glucose transport in the kidney tubules (Wright, 2011).

The nephron is the functional unit of the kidney in which reabsorption and secretion of micronutrients such as glucose and amino acids take place. SGLT2s specifically were localized to the early proximal tubule of the nephron (Wright, 2011). Localization of SGLT2s eventually led to their pharmacological development. SGLT2 inhibitors were proven to reduce renal glucose reabsorption and increase urinary glucose excretion. Soon after SGLT2 inhibitors started to be investigated as a new treatment for type II diabetes.

Unlike in the kidney, until recently the role of SGLTs in the brain were unknown. Glucose transportation to the brain was only known to be performed by GLUTs. Studies coupling positron emission tomography (PET) with the glucose tracer 2 deoxy 2 [F18]fluoro D glucose (2 FDG) demonstrate the distribution of glucose via GLUTs used to maintain regular brain function (Yu, 2010). However, since 2 FDG is only a substrate of GLUTs, the contribution
of SGLTs in the brain were concealed until the synthesis of α methyl 4 deoxy 4 [F18]flouro D glucopyranoside (Me 4 FDG), a substrate specific to only SGLT2 by Yu et al. In an ex vivo experiment Me 4 FDG was visualized in areas of the hippocampus, amygdala, hypothalamus, and cerebral cortices by autoradiography. An immunohistochemical assay using a SGLT1 specific antibody showed that the distribution of SGLT protein matched the functional activity as revealed by the autoradiography (Yu, 2010). In a follow up study Yu et al. were able to replicate results in vivo. Me 4 FDG accumulation was tracked in the brain after opening the blood brain barrier via osmotic shock (Yu, 2013). Once again the group showed immunohistochemical correlation in the same regions outlined by the Me 4 FDG uptake. Results also showed that in one brain region, the thalamus, there was Me-4-FDG uptake but no expression of SGLT1.

This raised three issues. The first implicated the possible use of SGLT2s for glucose transport in the thalamus. The second was to further understand why the brain would need active glucose transport when facilitated transport via GLUTs was available. And finally, the possibility that there could be a presently unknown functional link between the role of the thalamus in sensory and motor pathways as well as in the regulation of sleep and wake states to the unique glucose transporter that may be present there. In order to address the first question, we performed immunohistochemistry in different regions of rat brain with an antibody specific for SGLT2.

Methods

Rat brain was fixed in formalin and cryopreserved in a 25% sucrose solution to protect from the damage caused by freezing. The brains were cryosectioned into 7mm coronal sections. Two experiments followed. The first was the visualization of the rat brain using a known staining technique via 0.5% Cresyl Violet stain. The second used immunohistochemical assay to visualize
the SGLT2 antigen. For both experiments 7mm slices of human kidney were used as a positive control and 7mm slices of rat intestine were used as a negative control.

**Visualizing the rat brain**

Firstly, a well established technique to visualize the rat brain via a 0.5% Cresyl Violet staining was used. Cresyl Violet is known to stain the Endoplasmic Reticulum of cells, rendering the cell bodies blue while leaving fiber tracts white. The preserved slides were defrosted and exposed to 16 baths designed to prepare the tissue for staining by a series of dehydrations and rehydrations. Dehydration is achieved by replacing the water in the tissue with alcohol by soaking samples in baths with increased concentration of alcohol. Rehydration is achieved by replacing the alcohol in the tissue with water by sequential bathing in decreasing concentration of alcohol.

First the tissue were dehydrated. Bath 1: 70% ethanol for 2 minutes, bath 2: 95% ethanol for 2 minutes, and bath 3: 100% ethanol for 2 minutes. Then Xylene were used to clear the tissue. Bath 4: Xylene 10 minutes and bath 5: Xylene 10 minutes. The tissue were rehydrated: Bath 6: 100% ethanol for 2 minutes, bath 7: 95% ethanol for 2 minutes, bath 8: 70% ethanol for 2 minutes, and bath 9: 50% ethanol for 2 minutes. To complete the hydration the tissue were soaked in distilled water. Bath 10: DH20 for 2 minutes and bath 11: DH20 for 2 minutes. To stain the tissue the samples were soaked in 0.5% Cresyl Violet solution for 10 minutes and then cleared of excess stain by dipping them a few times in DH20. The tissue were again dehydrated by decreasing alcohol concentration. Bath 14: 95% ethanol for 2 minutes and bath 15: 100% ethanol for 2 minutes. A final bath in Xylene cleared all unstained tissue. Bath 16: Xylene for 2 minutes.
Immunohistochemistry

The Immunohistochemistry experiment is ongoing. The protocol used lasted two days. During the first day, the SGLT2 antigen was prepared and the first antibody was added to the tissue. The slides were soaked in PBS for 10 minutes. To perform antigen retrieval, the samples were incubated in citrate buffer for 20 minutes at 110°C. The samples were washed in PBS four times for three minutes, five minutes, seven minutes and seven minutes respectively. The slides were dried using a small vacuum line to remove the excess buffer then lined with a standard PAP pen to create a hydrophobic circle around the tissue. The slides were again washed in PBS four times for three, five, seven, and seven minutes. To block non-specific binding sites, a blocking buffer with no more than 100 microliters per sample was pooled over the tissue. The blocking buffer was removed using the vacuum line. Finally, the antibody was added to the slides using no more than 100 microliters per slide and incubated overnight at 4°C. This concluded the first day.

During the second day, the antibody solution was removed with a soft stream of PBS and then placed in a PBS-filled coupling jar. Samples were washed in PBS four times for three, five, seven, and seven minutes respectively. The slides were once again dried using the vacuum line. The secondary antibody was added at a 1:250 dilution and incubated for 90 minutes at room temperature in a moist box. The secondary antibody were washed off with another four baths of PBS for three, five, seven, and seven minutes. The Avidin Biotin Complex reagent was added and incubated for ninety minutes at room temperature. Samples were washed with another four baths of PBS for three, five, seven, and seven minutes. The slides were dried with the vacuum line and 3,3-Diaminobenzidine (DAB) was added to each slide. The development of stain was periodically monitored on a microscope to prevent overexposure. The reaction was stopped by rinsing the DAB into a waste bottle with bleach. The samples were rinsed in water for five
minutes. Counter stain with hematoxylin was achieved by soaking for 30 seconds in harris hematoxylin diluted 1:5. Followed by two, two minute washes in D2H2. Tissue was then immersed in NH3OH four times followed again by two, two minute washes in water. The slides were then rehydrated with 2 minute sequential passages in decreasing concentrations of ethanol first 70%, then 90%, and 100% followed by two Xylene baths for clearing. Finally the slides were removed from Xylene and mounted onto slides using a drop of paramount (slide adhesive) and cover slip.

Results

Cresyl Violet staining was successful. As expected the Endoplasmic Reticulum of the brain tissue made the cell bodies blue while leaving fiber tracts white or clear. The slides were observed under the microscope and cross referenced with a stereotaxic atlas. A stereotaxic atlas comprises a collection of pictures of brain sections with accompanying diagrams of structures that belong in each section. The crossreferencing revealed proper identification of key functional areas in our sample including the forebrain, thalamus and hippocampus.

Immunohistochemistry results for this particular sample are pending. However previous work done in the lab showed a positive SGLT1 antibody staining in the key functional areas of the brain examined in the present study suggest an increased probability of also tracking SGLT2 signals.

Discussion

Innovative work done at UCLA has discovered previously unknown SGLT glucose uptake in the brain via the synthesis of a high affinity tracers. There relevance to human pathology is significant as tissue in which SGLTs has been traced include high functioning areas such as the hippocampus which is associated with Alzheimers disease. A second tracer 4-[F-
18]fluoro-4-deoxy d glucose (4 FDG) specific to both GLUTs and SGLTs was also developed and has potential to serve as a new PET imaging technique that could provide a new tools to study brain health and disease (Yu, 2013).

**Conclusion**

The preceding was a small scale study to see if SGLT2s are used for glucose transport in functionally important brain regions. While final results are pending the study was motivated by the recent discovery of SGLT1s in the brain. This was a step to further understand why the brain would need active glucose transport when facilitated transport via GLUTs is available. And finally, a necessary step to address the possibility that there could exist a presently unknown functional link between the role of the thalamus in sensory and motor pathways as well as in the regulation of sleep and wake states to the unique glucose transporter that may be present there.
References

Ernest M. Wright, Donald D. F. Loo, Bruce A. Hirayama

Physiological Reviews Published 1 April 2011 Vol. 91 No. 7 733-794 DOI:

10.1152/physrev.00055.2009


2014 Alzheimer’s Disease Facts and Figures retrieved from:

Multiplicative Reasoning and Rational Numbers.

Marvin Lopez

Abstract

Many times people have difficulty multiplying fractional values. This may be a problem with the way people view the values. If seen as a number than decimals become more suitable for multiplication. However if the values are seem as expressions than the additional relational knowledge expressed by the fraction may prove to be advantageous. In this study we tested to see if people were actually able to pick up the information from one fractional equation and use it in an effective way when a related equation was displayed. It was found that people that were using fractions and not decimals gained an advantage in response time and in some cases accuracy. This suggests that people’s ability to see these relationships and use them in an effective manner is better with fractions than it is with decimals.
Introduction

When learning mathematics as children one of the hardest concepts to grasp is multiplication of rational numbers. This is especially true when the numbers in question are in fractional form. There are many possible reasons for this starting with an incomplete or inadequate knowledge of what it means to multiply. For simplicity’s sake children are often taught multiplication as repeated addition. That is they add the first number in the expression a certain amount of times, with the certain amount being the second number in the expression. This is not necessarily incorrect and in many cases it will lead you to the correct answer but that is not all there is to it. Multiplication is actually a relation of how many times larger one value is to another. This is important to know once multiplication is used in something other than whole numbers.

Another potential issue when multiplying is our understanding of rational numbers. There are many different ways that a person can perceive as well as express rational numbers. It is important to choose which notation to use in different situations because the form that is chosen will have an effect on the way the number is perceived as well as the way the data is processed. For example one way to look at a fraction is as a number. It has been shown that when asked to plot a fraction on a number line they automatically think of it as a number and look for the spot between two other digits that the value will fall into. This can also be done with decimals but in this case the value is much more apparent. In terms of multiplication seeing a fraction as a number does not work. When using decimals people are taught multiplication as an extension of multiplication with whole numbers. It is the same except there is the added step of manipulating the placement of the decimal place. In terms of fractions repeated addition does not work. It is very important to know the actual relational meaning of multiplication in this case. It’s true that
children might be able to easily add 6 to itself 4 times to get 24 but when the problem is turned into one of fractions this method becomes inconceivable. It would take someone with a lot of background in math to be able to multiply $\frac{7}{5}$ by $\frac{3}{7}$ in terms of repeated addition.

A more effective way might be to view fractions as an expression between their two terms. In previous research it has been shown that fractions are better than decimals at expressing relations between values. If people understand the reciprocal relationships that fractions contain then they may be able to use that understanding to improve performance on fractional multiplication.

Due to the fact that people learn multiplication with whole numbers first and then decimal multiplication as an extension of that they should be better at multiplying with decimals over fractions because they are more familiar with the method. However if people develop a better understanding of fractions then they might be able to use the expressed relationships during multiplication to their advantage. We believe that when people will be able to use the relationships in order to outperform people that multiply with decimals.

We looked at this by the use of priming. There were two conditions in our experiment. In one condition people only saw decimal expression. Participants were presented with paired trials that used numbers that had the same relationships but in their reciprocal forms as either fractions or decimals depending on their condition. If people can more readily see relationships in the fractions then they should be able to perform better than they would in the decimal condition.
Methods

Experiment 1

Participants
The participants were 13 male and 47 female undergraduates from the University of California Los Angeles. Their mean age was 20 years and their placement into the two conditions was randomly assigned.

Design
The study was a 2 (number type: fractions vs. decimals) x 2 (trial type: first trial vs. primed trial) x 2 (Problem type: true vs. false) design. Number type was a between subjects factor while trial type and problem type were within subjects factors.

Procedure
The subjects read the instruction sheet from a screen before the practice and test trials started. They were instructed to answer as quickly as accurately as possible pushing the “a” key for correct trials and the “l” key for incorrect trials. Their response time as well as their accuracy was recorded.

Results
In terms of accuracy there was a significant difference between participants in the fractions condition and the participants in the decimals condition with the people in the fraction condition being more accurate (90% vs. 78%; t(58) = 5.30; p < .001). However there was no difference between trial 1 and 2 for decimals or fractions whether they were true or false equations.

In terms of response time the true fraction trials were significantly faster than all of the decimal trials as well as the false fraction trials (2.76s vs. 4.03s; t(58) = 3.93, p < .001). For the
true prime trials, the response time for the 2 trial was significantly faster (2.21s vs. 2.58; t(1,29) = 3.08, p = .004). No other condition showed this improvement and in the case of true decimal trials the second trial actually seemed to slow down response time.

Experiment 2

Participants

For this experiment a total of 87 UCLA undergraduates participated. Their mean age was 20 with 70 females and 17 males. The condition that 29 of the participants were placed in was randomly assigned.

Design

This experiment was a 3 (number type: fractions “matching” vs. fractions “non-matching” vs. decimals) x 2 (trial type” first trial vs. primed trial) x 2 (Problem type: true vs. false) design. The non-matching group was added to this experiment. In this case problems that are the same as the original fractions are displayed with the exception that the numbers in the fractions did not match one to one with the other numbers displayed in the equation. For example, in the original problem, a participant might have seen the equation 4 x ¾ = 3. In the new condition they would see problems like 2 x 12/3 = 8

Procedure

In this experiment the procedure was identical to the one used in experiment 1

Results

For the true primed pairs, the matching fractions condition showed the best accuracy for trial one. Non matching fractions and decimals showed the same accuracy for trial 1 but for the nonmatching decimals accuracy improved during trial 2. In the case of false prime pairs the
matching and non-matching fractions performed similarly on both trials and were more accurate than the decimals.

In terms of response time, for the conditions that already existed, were the same as experiment 1. For the non-matching true prime fraction trials response time for trial 1

Discussion

Overall people were able to multiply with fractions much better than they did with decimals. In experiment 1 this was shown in terms of accuracy and reaction time. The improvement in response time for the second trial demonstrates that people are able to recognize the relationships between the numbers and use this information to their advantage. It could be argued that this advantage came from the fact that only the same two numbers were in the fraction so people might be able to visually recognize it sooner. The additional condition in experiment two is a good way to check this concern. Due to the fact that in the non-matching condition people have to perform some amount of computation it makes sense that the accuracy was similar to the decimal condition during trial 1. What is interesting is that for the second trial both accuracy and response time improved significantly. This is more evidence that suggests people are using the relational information that they recognize to their advantage in order to improve their performance without sacrificing accuracy or response time.

A way to take this research further would be to turn this into an fMRI study. It would be interesting to see which structures are activated for the different conditions. Findings in this study suggest that brain regions associated with processing relations will be more likely to activate for the fraction conditions. Memory regions will probably also activate during the second trial because the information gained during the first trial will be recalled either implicitly or explicitly.
References

DeWolfe, M & Holyoak, K.J. (2014) Reciprocal and Multiplicative Relational Reasoning with Rational Numbers
Metacognitive Judgments and Memory

Monica Vu

Finding truly effective and efficient ways to study are difficulties that most students face. The Bjork Learning and Memory lab looks at a multitude of mechanisms that can hopefully optimize the way students acquire information. Many students have a tendency to read and reread material. Though this seems like an effective process, the now familiarity of the subject matter can actually just be considered low-level priming (Bjork 2011). It is now considered that learning is only acquired when there is some level of difficulty in retrieving the material. Improvements in performance can produce the illusion of understanding and consolidating the material. This is an example of retrieval strength. While this retrieval strength is significant to performance such as taking exams, it is storage strength that determines how integrated the memory representation is. To achieve the desirable difficulties that one needs to solidify material, there are strategies such as varying learning conditions, spacing, and testing. Varying learning conditions is contradictory to what some may advise. Learning under novel conditions proved to be much more effective than consistent conditions, much of which may be attributed to adjusting oneself in multiple ways to provide a better holistic view.

Spacing has been discussed as a well-established means of consolidating information over time (Bjork & Kornell, 2008). Interleaving however is somewhat less intuitive. Interleaving is opposite to grouping, which means information is displayed in a manner that does not bunch relative information together but instead mixes content within one another. Most people who utilize massing may prefer it because they are able to notice similarities easily and attribute this feeling of familiarity with learning. The experiment in this article proves that interleaving is actually more efficient despite these intuitive feelings. The subjects in the
experiment had to do a memory recall test of different painting styles attributed to specific painters. Although the participants did better when the information was presented in an interleaved way, they incorrectly felt that massing produced a better effect.

Testing is an extremely effective way of determining if information has been processed. The way in which questions are presented however is debatable my many, especially in regards to multiple choice. Little et al. (2012) examines and attempts to debunk the criticisms against multiple choice testing in a very recent article. One of the criticisms is that multiple choice tests do not engage the student in retrieving information and thus does not support remembering it in the long run. It was then proven that if given plausible incorrect choices, the multiple choice test could actually be helpful in evaluating the information that they know or do not know and also assist in the recall of relevant material. In our study, we mirror their design to hopefully find more supporting evidence for the effectiveness of multiple choice testing.

A different approach to multiple choice testing is seen through an alternative scoring method called Modified Confidence Weighted – Admissible Probability Measurement, or MCW-APM. The experiment to test this method was used on students in an economics class. It was done as a possible solution to the concern that students might not have been confident with their retention of the material or if the material was even effectively presented. The design of the questions is much like the one we use in our study. It is still a multiple choice question in essence that there are a finite number of possible answers. The presentation is that of a triangle with the extreme options on each of the points. Other options included a midpoint between two options, an option that is between two options but one slightly higher than the other, and finally a dead center option that meant no confidence in any direction. The terms for describing these options were “near misinformed”, “misinformed”, and “uninformed”. Choosing an answer
closer to one of the points would result in a higher score if it were right or a lower score if the option was incorrect. In the experiment, a reward system was given to provide strong incentives to be honest. Despite this, many students seldom chose to be 100% sure with their answers. The results of the experiment showed that when students do exams that have the MCW-APM scoring, they are more engaged with the material than with just a right-wrong (R-W) approach. The students can then think more critically about the subject matter and understand the material better.

Our experiment uses this MCW-APM approach to look how metacognition affects memory for comprehension. In only its first run through, the experiment is not ready to produce our final results though the data is going towards the expected path. We hypothesize those students who are given the MCW-APM criteria to pre-test their knowledge will do best on a finalized test than those who did not.

Method

Participants

So far, 69 participants have participated in the experiment. All participants were students of the University of California, Los Angeles and were given class credit for their participation. They were all recruited from the SONA online system. An additional 51 students will be needed in the future to participate in order to see an expected effect at a level that is statistically significant.

Design and Materials

The experiment was a one way between subjects design. The independent variable was how the experiment was presented to the subject. They were in one of three conditions. The participant either had the a condition in which they were presented with a triangle that was
similar to the MCW-APM format and given a pretest, given a pretest with multiple choice answers, or given no pretest at all before answering the final questions. The dependent variable is how well the participants do on the final exam. The experiment was done on an online form. All of the participants were presented with the same two articles, one on Saturn and one on Yellowstone National Park.

**Procedure**

Each of the participants did the experiment individually on separate computers. A unique link was used to host the experiment. Instructions displayed for what was to occur during the study. If the participant had the condition that included the MCW-APM format, the scoring pattern of the triangle was explained and a series of questions were asked. Only when all the questions were answered correctly could those in that specific condition continue on. In all conditions, a passage was presented. The participant was given 9 minutes to read the passage, following the 9 minutes they were encouraged to continue studying it. In the condition with no testing, the participant moved on to the next passage. In the condition that had regular multiple choice questions, the participant went through a series of questions to test their knowledge of the passage that just followed. The subsequent passage followed and the same procedure with multiple choice occurred. The MCW-APM format was the same as the multiple choice but now instead of having the regular choices, the participant was given the triangle to try and answer their questions. Although feedback on correct answers was not given, the participants who had practice questions to answer were given a score relative to their responses. The game Tetris was then used as a distractor. Following the game the participants had their final test that was done by filling in the blank. Finally, at the end the students were asked to provide any comments or feedback about the experiment.
Results / Discussion

Since the experiment does not have an adequate amount of participants to produce statistically significant data, there are no results yet. This difficulty may be due to the low number of questions that we are asking the participants. Acquisition of more participants will be done next quarter to obtain additional data. Despite this limitability at the moment, the data that is currently available leads to a trend in the “correct” direction, one that is concurrent with our hypothesis.

We expect the results of the experiment to mirror that of the experiments previously mentioned. As a whole, we expect that the conditions where the students are being tested will do better than the condition where the student only reads the passage. This testing that is done reflects the desirable difficulty that was mentioned in Bjork (2011). It is proved to be more engaging than just rereading the passage since the student will have to retrieve memory that they had processed previously. We also expect in our results that the students who are in the condition where they use the MCW-AMP form will do better than the ones who do multiple choice alone. This mirrors the results from Bruno (1989) that showed students who dealt with this design became more interactive with the material and consequently thought more critically about what they had truly learned.

There are a number of ways to improve the experiment. The first one of course would be to have more participants. A second improvement would be to have more questions. A higher number of questions would provide more accurate data on the participant’s abilities. Through these improvements, we hope that the experiment can provide us with information that can help us encode information better. Possible implications of the experiment would be for professors to change their typical multiple choice questions to the triangle form that has been illustrated here.
Using this method could produce a very significant change in how professors test their students and how students encode the information that they learn.
References


Perceptual Learning in the Interpretation of Electrocardiograms

Olga D. Mercer

Perceptual learning (PL) can be defined as the process of improving perception (e.g., Gibson, 1969). Essentially, it is getting better at simple sensory discriminations as well as complex categorizations of spatial and temporal patterns that apply to real-world expertise. In our current research, we use perceptual learning principles to teach a variety of skills in domains ranging from mathematics to medical education. This research paper focuses on the application of perceptual learning to cardiology, more specifically to the interpretation of electrocardiograms (ECG).

Perceptual Learning

Eleanor Gibson set the stage for research on perceptual learning. She was interested in this “increase in the ability of an organism to get information from its environment as a result of practice”, and she attributed it to the detection of properties, patterns and distinctive features (Gibson, 1969). To illustrate this idea, Chase and Simon (1973) looked at the way perceptual learning operates in chess playing. They found that a master chess player was able to recognize much more advanced and complex patterns on a chessboard compared to novices. Even more surprisingly, a master has a sense of an impending chess mate ahead of time. Another example of the powerful effects of perceptual learning is the fact that a radiologist can spot a tumor on an x-ray quite rapidly (Snowden, Davies, & Roling, 2000).
As this phenomenon became more widely studied, Kellman (2002) characterized two kinds of effects that occur in perceptual learning: discovery effects, which have to do with the discovery of features and relations that are important to make a classification, and fluency effects, which are determined by the ease and automaticity with which we extract these relevant features and relations. Usually, the improvement in one type leads to improvement in the other. In order to study these effects more closely, Kellman and colleagues developed a computer-based learning technology called perceptual learning modules (PLMs). In PLMs, students are not asked to solve problems per se. Rather, they need to make a classification or a mapping of representations in a given domain, based on their underlying structure. For example, the task could be to match an equation with its graphical interpretation instead of solving the equation (Kellman, Massey & Son, 2009). Students are presented with many learning trials in which they learn to see the key structures and relationships.

Perceptual learning modules have many advantages. First, they employ 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). This means that they are based on a priority score system, where the priority for each item or learning category is updated after each learning trial as a function of accuracy, speed, and the trials elapsed since the last presentation of said item or category. Thus, students spend most of their practice time with the classifications that they are having trouble with. As a result, PLMs improve the efficiency of learning because it adapts to the learning faculty of individual students. Another central aspect of PLMs is that practice is done with mostly novel instances. This leads to pattern recognition in a wide range of cases, a task that is notoriously difficult to teach in regular instruction. Previous research has shown that, when learning a concept, being exposed to practice problems with high-
variability leads to better performance than practice problems with low-variability (Paas & Merrienboer, 1994). In general, we can say that PLMs are a very innovative and promising learning tool for modern instruction.

Electrocardiography

Electrocardiograms (ECGs) produce patterns reflecting the electrical activity of the heart, measured via electrodes placed on the skin. They usually require a trained clinician to interpret it in the context of the signs and symptoms the patient presents with. ECGs are not only one of the key tests performed when a heart attack is suspected but they are also widely used in hospitals. Although modern ECG machines often include analysis software, the diagnosis this produces is not always accurate. Therefore, knowing how to interpret ECGs is a very useful and helpful skill.

The issue is that such a skill is extremely difficult to obtain because of the complexity and variety of ECG patterns that exist. Medical students struggle with this and there is even a “trend toward subspecialization and dilution of ECG expertise among younger faculty” in medical schools (Auseon, Kolibash, Lewis, Lucey, Nagel & Schaal, 2009). Another alarming finding is that there that junior doctors have trouble diagnosing a heart condition from reading an ECG (Montgomery et al., 1994; Morrison & Swann, 1990). In fact, a number of studies have shown that “physicians of all specialties and levels of training, as well as computer software for interpreting ECGs, frequently made errors in interpreting ECGs when compared to expert electrocardiographers” (Salerno, Alguire, & Waxman, 2003). Since we cannot rely on computerized interpretations, our only hope is for “clinicians to re-examine their own skills in ECG interpretation” (Hongo & Goldschlager, 2004).

For our study, we developed a computerized training module in collaboration with faculty from the David Geffen School of Medicine to teach students how to interpret nine basic
heart patterns from ECGs. Typically, in medical school, students only practice with a few cases for each diagnosis. Our perceptual learning module utilizes over 300 unique and real ECGs. This means that the participants are exposed to many different cases, even for a single diagnosis. Since no two ECGs look alike in real life, practice with an extensive database in the PLM will encourage participants not to memorize any particular instances, but enabling them to extract relevant heart patterns across ECGs to later interpret a wide range of ECGs both with high accuracy and fluency. If this proves to be successful, we intend on making the use of these PLMs a permanent component to medical instruction.

Pilot Study

Participants

Undergraduate students of the University of California, Los Angeles took part in this study and received course credit for their participation. The participants did not have any prior experience reading electrocardiograms. However, a good amount of the participants took an interest in medicine and many were students of pre-health sciences.

Materials and Apparatus

For this experiment, we created a perceptual learning module that pulled images from a database of real 12-lead ECG recordings. We also put together a brief instructional primer of some basic knowledge about ECGs. It went over the 12 leads on an ECG tracing, how to measure the widths and heights of the waves for an ECG on the grid, and a description of the distinctive features for each of the nine heart diagnoses in the PLM. Finally we had a quiz that tested participants’ knowledge of the primer, asking them to match the name for each heart pattern to its corresponding description. In total, we trained participants to diagnose nine distinct heart patterns: Normal, Acute Anterior ST Elevation Myocardial Infarction (Anterior STEMI),
Acute Inferior ST Elevation Myocardial Infarction (Inferior STEMI), Left Bundle Branch Block, Right Bundle Branch Block, Left Axis Deviation, Right Axis Deviation, Old Anterior Myocardial Infarction and Old Inferior Myocardial Infarction.

The study consisted of a pretest, a training phase, a posttest and finally a delayed posttest that was scheduled one week after the first session. For each part of the study, we used a different set of images to prevent rehearsal effects.

Design and Procedure

Participants were assigned randomly to one of two training conditions. There was the active classification condition, in which, for each trial in the training phase, participants had to choose the correct heart diagnosis depicted on the ECG recording out of four options. The accuracy and response time (RT) were recorded for each trial. An example for a typical active trial is shown in Figure 1a. After each response, or after a 30 seconds timeout, the correct diagnosis for the ECG was highlighted in green, with a brief description of the diagnosis below. In addition, the diagnostic features were shown on the tracing (arrows and bars). The participants had to learn from this feedback. An example for a feedback screen can be seen in Figure 1b.

There was also a summary feedback screen of the participants’ accuracy and response time after every 20 trials for that block so that the participants could track their progress. The PLM used the adaptive sequencing algorithm well-known in the literature and optimal for learning to determine the sequence of presentation of ECG traces (Mettler, Massey & Kellman, 2011). In order to complete the module, the participants had to go through nine mastery levels, meaning they had to steadily improve on their accuracy but also in their response time for diagnosing each heart pattern.
In contrast, in the passive classification condition, for each trial in the training phase, participants were shown the ECGs directly with all the relevant features marked up and with a brief description of the corresponding diagnosis. Figure 2a is an example of a typical passive classification trial. Each trial was shown for 12 seconds. This duration was determined based on the average response time per trial of the first 10 participants in the active condition. After 12 seconds, the inscription ‘Time is up’ appeared at the bottom of the screen as well as a next button. This is shown in Figure 2b. This means that the participants in the passive condition simply had to pay attention to the relevant features and they completed the module when they had viewed about 285 ECGs. This number was assigned to be the number of trials in the passive version because it was the average number of trials that the first 10 participants in the active version took to complete module.

The pretest, posttest and delayed posttest were the same for the active and passive condition, meaning each was a set of 18 trials were the participants had to select the correct diagnosis for the ECG tracing. During those trials, they were not given any feedback. We recorded the accuracy and response time (RT) for each trial and calculated an overall score (percentage of responses answered correctly and under 15 seconds) and average response time on correct responses across all trials for each part of the study (pretest, posttest, delayed posttest) and for each participant.

The participants had 20 minutes to study the primer. Afterwards, they took the quiz and were asked to self-correct their responses using the primer. This lasted about 15 minutes. Then, they completed the pretest (18 questions) the training phase, which was meant to last about an hour in both conditions and finally the posttest (18 questions). They came back a week later for a delayed posttest (18 questions) to see how much they remembered.
**Expected Results**

Through this ECG PLM, we expect to teach heart patterns effectively and efficiently such that the overall response time and accuracy will have significantly improved from pretest to posttest. Also, consistent with prior PLM studies (e.g., Kellman, Massey & Son, 2009; Krasne et al., 2013) we expect that there would be virtually no decay over time, meaning that the percentage accuracy will not have significantly decreased from posttest to delayed posttest. Further, we hypothesize that the participants in both conditions will have the same average accuracy. This is due to the expectation that perceptual learning should happen in both the active and passive conditions. In fact, in real life, perceptual learning usually happens in a passive manner.

However, we hypothesize that the participants in the active classification condition will be faster (have lower response times) than those in the passive condition since they are more actively engaged in the classification task in the module. Research has demonstrated in a number of studies that when learners are actively engaged with the learning material (e.g., answering test questions), its representation in memory is changed such that the material becomes easier to recall in the future (e.g., Bjork, 1975). Thus, in our study, the participants in the active condition should be more fluent in interpreting ECG.

We are still in the process of running the pilot study and we are looking forward to see the results of our data analysis.
References


Figure 1a. Sample active classification trial in the perceptual learning module. This is an 12-lead ECG tracing with 4 answer choices below.

Figure 1b. Sample feedback screen after an active classification trial in the perceptual learning module. The correct answer is highlighted in green, the relevant features are marked on the ECG trace and there is a brief description of those features below it.
Figure 2a. Sample passive classification trial in the perceptual learning module. This is a 12-lead ECG trace. The correct answer is shown below the tracing, the relevant features are marked on the ECG trace and there is a brief description of those features below it.

Figure 2b. Sample screen after a passive classification trial is over. The inscription “Time is up” and a “Next Question” button appears.
Causal Parsimony and the Appreciation of Fractal Artwork

Priyanka Mehta

This quarter, we conducted a pilot study for an experiment exploring the relationship between the usage of parsimony in making causal decisions and rating the aesthetical appeal of fractal artwork. Our study was based on the central hypothesis that appreciation of visual art in human society evolved at least in part through an appreciation for a general level of complexity that can be applied to both causal and visual appreciation decision making. That is, we predicted that the criteria we use to create causal explanations for events might be related to the criteria we use to decide whether we find something visually appealing. Specifically, the criterion we focused on was parsimony.

In causal reasoning, parsimony is the idea of striving for the simplest explanation that satisfies the most requirements. Previous research has already indicated that we have evolved to make casual decisions not based on the mere fact of covariation of events, but based on actual causal reasoning—an attempt to arrive at a logically sound solution that accounts for all factors present in the event (Liljeholm & Cheng, 2007). But because there is an infinite number of possible causal explanations for any one event, it is important for us to be able to cut out all the far-fetched possibilities and focus in on the most efficient one we can find—and indeed, we do this when we make causal decisions (Liljeholm & Cheng, 2007). According to Hawking and Mlodinow (2010), parsimony is what drives us to create and improve scientific explanations. Their example illustrating this concept was that of the geocentric-versus-heliocentric theory of the solar system. Hawking and Mlodinow (2010) pointed out that most people believe the heliocentric theory (the sun is the center of the solar system) because they believe it is true, while the geocentric theory (earth is the center of the solar system) is just plain false, because of
countless pieces of evidence in mathematics, physics, and astronomy. However, this is not the case—that is, neither theory is actually “false.” This is because both theories can be explained through mathematics and physics and astronomy; however, the heliocentric theory makes these explanations considerably simpler (Hawking & Mlodinow, 2010). While explaining the geocentric theory scientifically would be extremely complicated, the heliocentric theory explains our observations just as well, with fewer rules (Hawking & Mlodinow, 2010). This is the essence of parsimony—people select between explanations of equal validity based on their efficiency and simplicity. It is important to note, however, that an explanation can be simple without being parsimonious, and an explanation can satisfy the most requirements and not be parsimonious. Clearly then, although the idea of parsimony is quite specific, one can see a lot of room for subjectivity in deciding what exactly “simpler” means, and what balance between simplicity and thoroughness is the greatest level of efficiency. The way a question is written can have a large impact on how the participant views his task (Liljeholm & Cheng, 2007), so it is important for the future to make sure our experiment does not allow for ambiguity in places where we do not want it. We encountered some difficulties with this later on during the experiment.

Our proposed relationship between parsimony and fractal artwork also deals with the idea of people choosing options with a preferential level of parsimony. While in causal reasoning, we expect people to choose the most parsimonious explanation to explain something, in art, we expect people to prefer artwork with a pleasing level of parsimony. But how does one measure parsimony in artwork? The way we chose to define this idea is through fractals. Fractals are patterns characterized by “self-similarity”—the idea that different sections or levels of the pattern are highly similar to other sections or levels of the same pattern (Taylor, 2002). For example, a fractal pattern of triangles might begin with one triangle with three more triangles
inscribed in that triangle, and three triangles inscribed in each of those triangles, and so on.
Fractals are not limited to artificial geometric patterns, though (Taylor, 2002). Fractals are actually very common in nature. An example of a fractal in nature might be a certain tree, whose large branching structure from the trunk follows a pattern which is more or less repeated in the branching pattern of small branches from each of the large branches, which might be seen again in the branching pattern of twigs from the small branches (Taylor, 2002).

There are two relevant facts about fractals to this study: The first is that they represent order in chaos. This is important because visual art, like paintings, are in a similar situation. Usually (unless we are dealing with geometric art) paintings do not follow geometric formulas; they represent organic forms that do not have a clear order. But, as discussed earlier, natural forms do often contain and underlying fractal order whether we notice it or not (Taylor, 2002). One of our goals in this study was to explore how fractal order in art contributes to how much we appreciate it. But, to appropriately measure such a construct, we needed an art form that was not obviously fractal, but at the same time able to be quantified in fractal terms, which is quite difficult to do with say, the Mona Lisa.

For this reason, we looked at the work of abstract artist Jackson Pollock. Pollock was famous for drip paintings—he would lay out a canvas on the ground and splatter paint all over it in a seemingly random fashion. Although this sounds completely visually unappealing, people were inexplicably drawn towards his artwork, agreeing that there was something aesthetically pleasing about it (Taylor, 2002). Then, a possible explanation for this visual phenomenon appeared: Taylor (2002) found that Pollock’s paintings actually contained fractal patterns. In this way, fractals were linked to modern artwork, bringing us one step closer to understanding why chaotic visual art might be aesthetically pleasing.
The second important fact about fractals is that they have a quality called dimensionality (Taylor, 2002). A fractal’s dimensionality has to do with the amount of repeating patterns present in the particular fractal. It is calculated mathematically and is expressed as decimal values between 1 and 2. The greater dimensionality, the more levels of repeating pattern, and the higher the value. So, a pattern with an intricate, dense construction of multiple iterations might have a dimensionality of 1.8, while a simpler, less dense construction of a few self-similar levels might have a dimensionality of 1.2. Because dimensionality is related to complexity, one could say that a low dimensionality is comparable to a high level of simplicity, or parsimony. A high dimensionality is comparable to high complexity, or low parsimony. In this way, we are linking the concepts of dimensionality and parsimony.

We put these facts about fractals and the idea of parsimony in causal reasoning to hypothesize that there is a positive correlation between the way in which we reason causally and the way in which we judge artwork. This means that we predicted that not only would people distinguish between the more parsimonious possible explanation for an event and a less parsimonious one, but that they would also employ the same process in making decisions about how appealing an artwork is. That is, if they preferred extremely parsimonious explanations, they would also prefer parsimonious artworks.

The study this quarter was a pilot to fine tune the experimental materials, and hopefully eventually expand further into a study involving MRI scans during the experimental process to see the underlying brain regions involved in these processes.

**Methods**

**Participants**

Our participants were 48 students at UCLA participating for course credit.
Design

We used a within-subjects design to measure the effect of the independent variable of parsimony level on two dependent variables: one was causal judgment and one was art preference—these were defined and measured as the participants’ self-report responses when asked which explanation they thought was better, or which artwork they preferred. We were looking for a correlation between the two. We operationally defined parsimony for causal explanations as the amount of rules and premises needed to account for a result. For example, if there was a choice between an explanation that had three rules and one that had one rule, as long as they both explained all the circumstances, the latter would be the more parsimonious.

Materials

This experiment required a computer program that displayed to the participants three different conditions: the causal condition, the causal control condition, and the art condition. These three conditions were presented in random order to each participant, but the control condition never came before the causal condition.

The art condition was a series of 18 sets of paired artwork. Each set consisted of a slide with one artwork, a second slide with a second artwork, and a third slide with them both side-by-side, and an instruction to the participant to compare them and select which one they found more “visually appealing.” The artworks were black-and-white computer-generated fractal images. One third of them had dimensionality level 1.2, one third had 1.5, and the rest had 1.8. Each image pair was one of three types: low-high (one 1.2 image, one 1.8 image), medium-high (one 1.5 image, one 1.8 image), or low-medium (one 1.2 image and one 1.5 image). The purpose of this section was to determine what level of dimensionality (or parsimony) the participant appeared to favor.
The causal reasoning section consisted of a series sets of slides each depicting a short chain of events. The first slide of each set was labelled “Day 1” and contained a picture of a fruit and a picture of a man, who either had or did not have a rash on his face. The next slide was “Day 2” which had either a picture of a different fruit, both the first fruit and a different fruit, or no fruit, and the man who either again did or did not have a rash. For example, the episodes might have occurred thusly: on Day 1, he ate Fruit A and had no rash. On Day 2, he ate both Fruits A and B, and did have the rash. The next slide reproduced the two episodes and then offered one explanation for the events (i.e. Fruit A does not cause the rash, but Fruit B does). The following slide had a different explanation (i.e. neither fruit causes the rash alone, but something special about the combination of the two of them causes the rash). For each account, the slide asked the participant to answer if that account explained the circumstances (yes or no). The final slide of the set asked the participant to choose which of the two was the “better explanation.”

The purpose of this section was to assess the participant’s use of parsimony in causal reasoning—whether they would choose the more parsimonious explanation (in the example given, it would be the former, not the latter.) The specific manipulation in this section of the experiment was the level of parsimony; all other aspects of the explanations (like syntactic simplicity) were held constant. Each explanation pair was ranked on a scale of low, medium, or high parsimony, and paired in the same ways as the art section: there was an equal number of low-high, low-medium, and medium-high pairings. There was also an additional condition—a control, where an explanation that did explain the circumstances was compared against one that did not (in the previous example, this might say “Fruit A causes the rash and Fruit B does not.”) This was to ensure that participants were actually taking into account the validity of the causal explanations.
The causal control section was the same as the causal reasoning section, save for two main differences. The first was that the question asked was not “which account is the better explanation,” but “which account do you prefer?” This was to eliminate any explicit causal thought process in the participant so that they could focus completely on the structuring of the accounts put in front of them. The second difference was that this time, instead of holding simplicity constant and varying the parsimony levels, there were five separate conditions with different variations between accounts. The five conditions of this were: simplicity change, parsimony change, congruent, incongruent, and a control. Simplicity change meant the parsimony level of the two conditions were the same, but the syntactical simplicity was different. We defined syntactic simplicity as the use of active voice as opposed to passive voice to say the same thing. For example, saying “Fruit A causes the rash” would be syntactically simpler than saying “The rash was caused by Fruit A.” Parsimony change meant simplicity remained constant, but parsimony level changed. For example, both sentences might be in active voice, but one would be the low parsimony level, while the other would be high. The congruent condition was when the parsimony level and simplicity level both changed, but they agreed for each explanation. For example, one account would be high parsimony in active voice, while the other would be low parsimony (high complexity) in passive voice. The incongruent condition was the opposite of this: parsimony and simplicity disagreed with each other. That is, one explanation could be active voice with low parsimony, while the other would be passive voice with high parsimony. Finally, the control was like the control for the causal section: both accounts were in active voice, but one explained the situation and one did not. The purpose of all these conditions was to detect patterns in the types of answers that participants were giving, and to see what they were taking into account besides parsimony level to make their decisions.
Procedure

The participants came into the lab one at a time. We handed them the consent form and asked them to read it. Then we led them to the computer and instructed them that all the instructions were on the computer. They then sat through the experiment, pressing buttons labelled “L” and “R” (for “right” and “left”), or “Y” and “N” (for “yes” and “no”) depending on what the question asked (when choosing between explanations, it was R or L, and when being asked if the account explained the episodes, it was Y and N). When the three conditions were over, the participants were asked to fill out a post-experiment questionnaire about their thought process during the experiment before they left. We used this questionnaire to better our understanding of how well out manipulations were working.

Results

We have not completed the statistical analysis for this project yet. From preliminary observations however, we can make a few guesses. First of all, from the questionnaire it is apparent that there is some division between participants; while some felt that a more general explanation and a lower dimensionality picture was better, others felt that a more specific explanation and a higher dimensionality picture was better. On a positive note, however, these preferences did seem to correlate—if the participant said they made decisions based on greater causal detail, they also said they chose paintings with greater detail (high dimensionality). However, this difference in preference may lead to some ambiguous results for the tests in comparison of levels of parsimony, if participants had different ideas of what parsimony is. In any case, the results of this pilot study will be of considerable help in developing the next stage.
References


Animation Understanding and Problem Solving

Sana Tauqir

Abstract

Scientists have been trying to understand the human use of analogical reasoning in everyday life. There is obvious evidence of it however the mechanisms involved in the transfer of information from a prior situation to its application to another seemingly irrelevant situation is still undergoing wide research. In this experiment, we hope to determine how successful participants are at transferring information from a source analog onto a target problem. Specifically, we manipulated the methods of data presentation into three conditions: 1) Full Motion, 2) Implied Motion and 3) Verbal Only. The preliminary results are in line with our hypothesis. Participants in the full motion condition were more successful at the analogical transfer task whereas those receiving a verbal only explanation were less successful. Eventually the hope is to create a computational model such that it enables the teaching of such an efficient method of learning to computers.
Analogical reasoning is one of the most interesting induction techniques that humans employ in problem solving. It is a very efficient shortcut method however individuals do it almost unconsciously. Many teaching techniques are built from this basis, especially in complicated fields such as physics. Research into this topic has been proceeding for quite some time. A better understanding of this process in humans would lead to the creation of computational models for analogical reasoning which could then be taught to computers to make them more efficient. What is about a source analog, is it surface similarities or the method of its presentation, that helps individuals form a mental representation which can then be called upon when presented with a novel situation.

One of the earlier studies in this topic was done by Holyoak and Koh on whether surface and structural similarities between the source and target analogues influence analogical transfer. In the first experiment they demonstrated that even with a delays of several days in between, participants experienced spontaneous analogical transfer. The second experiment showed that similar structural features between the target and source were a more reliable determinant of successful trials than similar surface features. (Holyoak and Koh 1987) Our target and source problems were not very structurally similar, so according to the Holyoak and Koh study we should not expect a high transference. The target and source analogue used in this experiment were created considering these aspects of the Holyoak study.

Additional studies on analogical transfer influenced the creation of the three different conditions presented to the participants. Our study is most closely related to the study done by Pedone and colleagues on the use of diagrams in problem solving. Initially they presented participants with static diagrams showing either divergence or convergence which participants later found difficult to relate to the target problem. (Pedone, Hummel and Holyoak, 2001) The
additional manipulation of adding a verbal explanation or a convergence animation lead to easy recognition of the analogy between the source and the target. In our experiment we used this verbal explanation, which as expected from the Pedone study did influence recall but not as effectively as the animation condition.

The purpose of this study was to gain a deeper understanding of the use of analogical reasoning on human problem solving abilities. The participants were placed into one of three conditions: 1) Full Motion, 2) Implied Motion or 3) Verbal Only. Instead of simply using diagrams as in the Pedone study the participants were presented with either a full animation of a cannon shooting an enemy circle through a friendly barrier or a low frame-rate animation or pictures with a verbal explanation. They were then asked to solve the tumor problem, which was built off the Duncker radiation problem used in the Koh and Holyoak study. Successful transfer from the source analog to the target analog was defined as providing the correct solution to the tumor problem, which was to point lower-intensity rays from positions around the tumor and thereby minimize damage to the healthy tissue while eliminating the tumor. The full motion condition was expected to be the most successful since motion has been found to make features of a problem more salient than a simple verbal explanation, where motion would have to be imagined.

Experiment 1

Method.

Participants

Thirty-three undergraduate students participated in this study for course credit through the University of California, Los Angeles’ (UCLA) Psychology Department's SONA Experiment
Management System. The participants had normal vision and were able to understand, read and write English.

*Materials*

A convergence analogue was written for the purpose of this experiment. It involved a cannon shooting at an enemy circle through a friendly barrier. Figure 1 shows this set up as it was presented to the participants. There were three scenarios based upon this source problem. The aim was for the circle damage to reach its maximum before the friendly barrier reached its critical damage point. This is demonstrated in Figure 2. In the first scenario, the cannon is unable to defeat the enemy circle because the friendly barrier overheats before the circle has suffered enough damage. In the second scenario, the cannon is once again unable to defeat the enemy circle because the small cannonballs it is now attempting to use, although causing less concentrated damage to the friendly barrier, overheat the friendly barrier and do not cause sufficient damage to the enemy circle. In the third scenario the eight cannons aimed at the enemy circle from different directions succeed in damaging the circle before the friendly barrier overheats.

The target problem assigned to the participants to solve was very similar to Duncker's (1945) radiation problem. In this experiment it is referred to as the tumor problem. You are a doctor who must destroy a malignant tumor in your patient using rays that can vary in intensity. High intensity rays will kill the tumor however the problem arises here from the fact that these rays will then also damage the surrounding healthy tissue. Therefore in the source analog, the canon is a substitute for the ray generator, the cannonballs are the rays, the enemy circle is the tumor and the friendly barrier is the healthy tissue.
One of the primary materials used in this experiment was a Dell computer for the presentation of instructions and animations. The experimenters recorded relevant characteristics of the participants into a simple table which also indicated the condition which they would participate in. Additionally, there was a packet of papers for each run of the experiment. The first three pages of this packet questioned the participants on the scenarios they had just been presented with, the fourth and fifth pages were the tumor problem presented without and with the hint, respectively. The last page was a questionnaire.

**Design**

The participants were assigned into one of three conditions. Condition 1 was the “Full Animation” where they saw each of the three scenarios as a high frame-rate animation. In the “Implied Motion” in condition 2 participants saw each of the three scenarios as a low frame rate animation. And finally the participants in the “Verbal Only” condition 3 were provided a verbal explanation of each of the three scenarios with pictures as their only visual. By manipulating the presentation of the source analog we wanted to see the effect on the rate of analogical transfer.

**Procedure**

When participants arrived for the experiment they were asked their major, this was recorded along with their gender into a table. They were assigned in serial order to one of three conditions. Each of these began with the same step, the participants were presented with instructions and the source problem on the computer screen. After this point, they proceeded to watch the animation or listen to an explanation, corresponding to their condition number. Upon completion of this step, the participants were given a modified version of the AQ test as a filler task. They were then questioned about what occurred in each of the scenarios in order to gauge to what extent they understood the source problem alone. While filling out these questions they
were presented with pictures of each scenario. The participants were then presented with the tumor problem (the target problem), first without the hint and then with the hint. The hint simply stated to think back to the source problem. The final step was filling out a questionnaire asking whether they had seen the problem before and whether they successfully picked up on and applied the source to target analogy in their problem solving.

Results and Discussion

The results for this experiment have not been analyzed yet since experimenters are still generating a proper grading schema.

Experiment 2

This experiment was very similar to Experiment 1 in many respects. The first experiment demonstrated that the participants were able to carry over their knowledge about the three scenarios through the AQ test filler task. In Experiment 2, by questioning participants on the scenarios right after their presentation we hoped to gain some idea of how successful they had been at understanding the target problem.

Method.

Participants

Sixty-one UCLA undergraduates took part in this experiment for psychology course credit through the UCLA Psychology Department's SONA experiment management system.

Materials

The materials used in this experiment were the same as those used in Experiment 1.

Design and Procedure

The design of the experiment was very similar to Experiment 1 as well, the only difference was that the participants were questioned on their understanding of the source
problem before, rather than after, the presentation of the AQ test. This was a key difference

BECAUSE

Results and Discussion

At this point there are only preliminary results available. However a rough analysis is
proving to be in sync with our hypotheses.

The preliminary results from the first thirty-three participants show some promising data.
These results are shown in detail in Table 1. Initially twenty-three participants successfully
transferred, and with the addition of the hint, this number was increased to twenty-eight and only
five participants failed to see the analogy. These are quite promising results but since only a
small sample size has been graded it is necessary to analyze the remaining data to determine a
proper pattern.

General Discussion

The results of a large part of the experiment still have to be analyzed. However the trend
appears to be in line with the preliminary results from Experiment 2. Once again, what was
interesting about Experiment 1 in comparison to Experiment 2 was that the participants were
able to maintain their knowledge of the scenarios through the filler AQ task and then still see the
analogy between the two situations with relative success. It is also possible that by questioning
participants on the scenarios right before the presentation of the tumor problem, the information
was more readily available to these earlier participants than those in Experiment 2. A deeper
analysis of the data should include some way to measure the extent to which they were able to
pick up on the salient features of the source analog and apply it to the target problem.

The preliminary data from Experiment 2 showed that participants in the Full Motion
condition were most efficient at solving the tumor problem even before the hint was provided. It
appears that by providing a full feature animation, the convergence motion is very easy for the participants to pick up on. The low frame rate condition was as unsuccessful as the verbal in initial transfer, however the hint seemed to put it on par with the full animation. The verbal explanation was least efficient in the analogical transfer task because the . The hint was successful in reminding as can be seen in the increased number of participants who successfully transferred in condition 2 and 3. In condition 1 the hint appears to have made no difference since the motion alone was enough for transference, the one participant who failed simply did not make the analogy.

This study is incomplete in that most of the data has yet to be thoroughly analyzed. The position of the Implied Motion condition as first seeming close to Verbal Only and then the rate of transfer being same as that of the Full Animation condition is worth doing additional research on. It would be interesting to see what was the reason for this sudden increase, the verbal condition shows that the hint was not that informative alone. Perhaps it was the level of animation played a role or there was some interaction with working memory abilities. In addition to the raw data, there are some auxiliary data to analyze as well. We hope to analyze whether the gender made a difference or the declared major of the participants had an effect on the rate of problem solving. Further experiments could then arise from these analyses as well.
References


**Tables and Figures**

**Figure 1:** The source analog presented to each participant and explained as part of the instructions.

<table>
<thead>
<tr>
<th>Barrier Damage:</th>
<th>Circle Damage:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2:** The damage bars shown in each condition, the aim was for the circle damage to reach maximum before the barrier damage.
Table 1: The preliminary data from Experiment 2. 11 participants in each condition. There was an increase in transference with the hint and participants were less successful if the condition presented less animation.

<table>
<thead>
<tr>
<th></th>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Condition 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>W/O Hint</td>
<td>10</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>W/ Hint</td>
<td>10</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Failed</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
Reward and Negative Affect Study

Sonia Bhatia

I have been a research assistant at the Social and Affective Neuroscience Lab since the summer of 2013. During the summer, I began working on a study investigating the relationship between reward and negative affect, which I continued working on during Fall and Winter quarter, along with other small projects.

As a research assistant, I’ve conducted many tasks, ranging from running experiments on subjects to data entry. One of the main projects that I began working on during the summer investigated the relationship between reward and negative study. I continued working on this study in the fall and during this past quarter as well.

My tasks during the Reward and Negative Affect study included running the study on participants, and conducting data entry and analysis. Another Reward and Negative Affect study had also previously been conducted in my lab, but it was slightly different than the study I worked on. The first study consisted of two conditions, one where people concurrently consumed a reward (in this case, chocolate) and did the stop signal task, and another where people were not consuming a reward while doing the stop signal task. The study that I worked on, however, consisted of one condition where subjects anticipated a reward during the stop signal task, and a control condition where subjects did not anticipate a reward.

The results from both of these studies combined were that when people were concurrently consuming a reward (in this case it was chocolate), they expressed less frustration and negative affect to the stop signal task in comparison to when they were not consuming a reward. But, when subjects were anticipating a reward (instead of concurrently consuming a reward), participants showed no difference in comparison to the control condition (where they
were not anticipating a reward). In order to understand more about these results, I was asked to find any articles that may contribute information pertaining to these results.

I used three main databases to find articles pertaining to the results- Google Scholar, PubMed, and Psych Info. Because I had taken Psychology 100B: Research Methods, using these databases were very easy for me. The goal that I had to achieve was to identify what the neuroscience, medicine, and psychology literature was saying about the relationship between reward processing and stress/negative affect. I knew that finding information about the first part of the study would be easy. All I needed to do was find information that supported our conclusion, which was fairly straightforward and predictable. I therefore assumed that there would be many other studies that would have similar results. The second study, however, had a somewhat strange and unpredictable result. I knew that it would be a challenge to understand why people were not affected by the anticipation of a reward.

As I predicted, finding articles for the first part of the study was fairly easy. What I did was that I searched for certain keywords in the databases. The keywords I used were: reward, stress, negative emotion, negative affect, positive emotion, positive affect, ventral striatum, and threat. Finding articles that pertained to the second part of the study was significantly more difficult. I had to alter many keywords, and as I suspected, there were extremely few articles published pertaining to this research.

After reading so many articles throughout the quarter, I furthered my knowledge on the study that I had worked on for six months. It was very interesting to see how the research we were conducting pertained to other research. It was also interesting to note how many people had done similar experiments on animals. I think that doing this type of research improved my skills in using a database vastly. By the end of the quarter, I read more than 300 articles, which
extremely helped me in my knowledge of social psychology, neuroscience, and research in
general. Additionally, by the end of the quarter, I was able to instantaneously determine whether
an article would be helpful to our research. This was very gratifying, because I definitely
struggled at the beginning of the quarter, as the vast amount of information in these articles was
extremely overwhelming.

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 am going to be
continuing research in the Social and Affective Research Lab on a volunteer basis. I really
appreciate the fact that 186B is a course that is mandatory for all cognitive science majors,
because I definitely feel as though hands- on research is a significant part of the learning process
in this major. I think that through research, I have learned just as much about psychology than I
have from my actual psychology classes.
Perceptual Learning

Victoria Groysberg

When we learn new perceptual classifications, what principles govern successful learning? Are there ways of organizing the order of presentation of material such that learning is enhanced? The principles that govern successful learning when people learn new perceptual classifications and the ways in which the order of presentation of material can be optimized remain questions that require further scrutiny. This is key because appropriately designed Perceptual learning (PL) technology can produce rapid and enduring advances in learning (Kellman et. al 2009). PL is the process of learning improved skills of perception. PL differs from other types of learning in that it can occur over a wide scale of time, like weeks, and it can occur for a wide variety of perceptual tasks, ranging from sensory discriminations to very tactile acuity tasks. PL is also restricted to the specifications of the stimuli and task where training is occurred, and the learners do not usually require feedback in order to exhibit PL effects. The PL often does not transfer to other tasks, stimuli, or sensory locations. On the other hand, adaptive learning is when computers, used as interactive teaching devices, adapt the presentation of educational material according to students’ learning needs, as indicated by their responses to questions and tasks.

The 2009 study conducted by Phillip J. Kellman, Christine M. Massey and Ji Y. Son, “PL Modules in Mathematics: Enhancing Students’ Pattern Recognition, Structure Extraction, and Fluency” tested PL modules (PLM). Three Perceptual Modules were tested, each emphasizing different aspects of complex task performance, in middle and high school mathematics: MultiRep, Algebraic Transformations, and Linear Measurement. In the MultiRep PLM, participants were presented with an equation, graph, or word problem of a function and asked to pick the equivalent function across multiple representations. Results demonstrated that the PLM improved students’ abilities to generate correct
graphs and equations from word problems. In the Algebraic Transformations PLM, participants were asked to choose a valid equation that was a correct transformation of another equation. This practice in seeing equation structure across transformations, but not solving the equations, led to dramatic improvements in the speed of actually solving the equation and obtaining a solution. In the Linear Measurement PLM, participants were presented with a graphic display of a ball on top of a ruler with a billiard cue aimed at it. They were asked to extract information about units and lengths to produce successful transfer to novel measurement problems and fraction problem solving and the results demonstrated that these successful transfer were indeed produced. It has been suggested that these PL effects can be divided into discovery and fluency effects.

Both discovery and fluency differences between experts and novices have since been found to be crucial to expertise in a variety of domains, such as science problem solving, radiology, electronics, and mathematics. Therefore, a discussion of discovery and fluency and the differences and similarities between the two processes are vital in order to gain a greater understanding of what it means to gain expertise in an area. Discovery pertains to finding the features or relations relevant to learning some classification, whereas fluency refers to extracting information more quickly and automatically with practice. The interplay between discovery and fluency is incredibly complex and in order to disentangle or even begin to understand the two, greater understanding of learning needs to be obtained.

It has been demonstrated that there are effects of varying either the temporal distribution of practice mathematics problems or the order in which mathematics problems are solved (Kellman et. al 2009). Neither manipulation required an increase in the total number of practice problems, yet both experiments showed increases in subsequent test performance. That showed that altering the timing of practice led to large gains in test performance. A different study used an adaptive learning system (Adaptive Response Time Based Sequencing—ARTS) that uses both accuracy and response time (RT) as direct inputs into sequencing (Mettler & Kellman
Response times were used to assess learning strength and to determine mastery, making both fluency and accuracy goals for the participants. The study used adaptive learning and created the spacing between items presented by expanding item recurrence intervals as an inverse function of RT. Third graders in an online school learned basic multiplication facts in about two hours using ARTS and outperformed a control group using standard instruction. Results of experiments in a dissertation study showed that an adaptive scheduling algorithm produced greater learning gains than fixed schedules. This was true for when the total number of presentations was limited the gains were measured after a one-week delay. Adaptive schedules still outperformed fixed condition schedules, expanding schedules, in terms of learning gains at immediate and delayed tests (Mettler 2014). Recent pilot data from the Learning Science Facts study, a study of adaptive learning of facts, also appears to support the claims that presentation of information using an adaptive learning algorithm that determines presentation of items based on previous performance has been found to be more effective than when information is presented using a fixed, expanding schedule (Mettler 2014). These results suggested that response time-based adaptive learning has remarkable potential to enhance learning in many domains.

Sometimes factual and procedural learning are types of learning that are necessary complements to the perceptual learning that occurs. The main question posed is how do we use a learning intervention that targets learning? It was hypothesized that the use of an adaptive system would lead to larger gain in learning rather than using a fixed, expanding schedule that was matched to the adaptive schedule. The fixed schedule was a schedule in which time elapsed between the presentation of the item continuously increases throughout the duration of the study.

It has been demonstrated that when radiologists fixate and scrutinize potential abnormal features during a search for lung tumours in chest images, a high proportion of those features which the radiologist has detected visually, but dismissed cognitively, turn out to be true abnormalities, thus indicating that
further research into how the information that is learned perceptually by humans should be used, understood, and weighed against other sources of the information the person is aware of (Gold 2002). A study found that luminance contrast-specific improvements in sensitivity to low-contrast dots in X-rays in expert radiologists was better than novices' (Sowden et. al 2000). The study also found a direction of luminance contrast-specific improvement in novices' detection of low-contrast dots in X-rays as a result of practice. Therefore, perceptual learning occurs when radiologists learn to read X-rays. The experiment tested the hypothesis that if part of the expert's superior performance rested on enhanced sensitivity, then tasks where other components of their expertise (e.g., search skills, knowledge of likely patterns of abnormality, etc.) are irrelevant would still demonstrate this superior performance. To test this, the stimuli were X-ray phantom images to which any enhanced sensitivity experts possess should transfer, whereas their other task-specific skills and knowledge would not.

The expanding schedule has shown to be effective for adaptive learning, but there is sufficient support to think that the expanding schedule could also be highly effective for perceptual learning. Recent analyses suggest that PL in a variety of tasks depends on processes that discover and select information relevant to classifications being learned. Discovery involves finding structural invariants amidst task-irrelevant variation (Gibson, 1969), allowing learners to correctly classify new stimuli. In an investigation of whether an adaptive, response-time-based, category sequencing algorithm could enhance perceptual category learning and transfer to novel cases, it was shown that effects of adaptive learning existed for perceptual categories. These results suggest that adaptive category sequencing increases the efficiency of PL and enhances generalization of PL to novel stimuli, key components of high-level PL and fundamental requirements of learning in many domains. The hypothesis was accepted that the use of an adaptive system would lead to larger gain in learning rather than using a fixed, expanding schedule that was matched to the adaptive schedule. The purpose of the fixed schedules are not to give insight into perceptual learning, discovery or fluency, but the fixed schedules were used to test principles of adaptive scheduling in fact learning. The potential of changing training programs and educational systems using
findings that can improve learning fuels further research into the use of adaptive schedule in different types of learning.
References


doi:10.1111/j.1756-8765.2009.01053.x


Desirable Difficulty as a Product of Audio Disfluency

Zach McFadden

Perceptual disfluency is a central component to desirable difficulty. Given this context, perceptual disfluency is defined as perceptual stimuli which are strenuous to process. Desirable difficulty occurs due to perceptual disfluency enhancing learning and memory, and is thus “desirable”. Examples of perceptual disfluency include: blurred words (Yue, Castel, Bjork 2013) and difficult to read texts (Diemand-Yauman, Oppenheimer, Vaughan 2011). Otherwise known as the perceptual-interference effect, given the correct circumstances, memory can be improved due to perceptual disfluency. Note however, that desirable difficulty does not always produce a positive effect on memory, which we will discuss later (Bjork 2013). Thus the question stands: how does perceptual disfluency act on memorial processes?

Previous literature has revealed that disfluency leads to deeper analytic processing of data (Oppenheimer 2013). Furthermore, Oppenheimer found that due to the rallying of cognitive resources as triggered by perceptual disfluency, memory performance increases. Therefore, by actually making tasks more perceptually difficult to interpret, individuals will put forth more effort to encode information in order to derive an answer. Thus, disfluent items are encoded differently in memory than fluent items (normal items). Fluent items do not trigger the same amount of deep analytical processing that disfluent items trigger. However, if a task is of a certain difficulty, perceptual disfluency has no noticeable positive effect on accuracy of participant responses. In some cases even, if a task is too difficult, disfluent conditions can do worse than fluent conditions in learning and memory tasks. It appears then, that desirable difficulty does not always supplement memory. Therefore, it is pivotal to know when disfluency promotes learning, and when it does not.
Bjork (2013) inspected various situations in which difficulty was desirable and when it was not. Blurred word lists comprised the disfluent conditions for this study, whereas normal words were the fluent condition. It appeared that when processing time for words was a short three seconds, disfluency actually induced a negative effect on the memory. A replication of this study with five seconds of processing time produced equivalent recall in disfluent and fluent conditions. This suggests that the disfluent condition was just too difficult given a short processing time, preventing a perceptual-interference effect. This expresses a concern for disfluency as mentioned above; if a task is too difficult, disfluency can produce a negative effect. Increased processing time may indicate that disfluent conditions induce a rehearsal of items, which could be a component of the deeper analytical thought occurring for disfluent items. Other possible explanations for the lack of an effect of disfluency could be in part due to participant beliefs that disfluent items are less important than fluent items. When participants are given disfluent stimuli such as blurred words, they may perceive that word as less important than words which are not blurred. Thus, when constructing materials with the intent of creating a desirable difficulty one must be careful in making sure this does not arise.

A similar research paradigm regarding disfluency was executed by Oppenheimer and colleagues (2011). A variety of difficult to read fonts were used as the disfluent condition, versus normal fonts. It appeared that, as predicted, difficult to read words produced greater memory performance than easier to read fonts. Again, greater cognitive load (disfluency) produced enhanced task performance. It is important to note that the type of disfluent font did not matter in the given study. The mere fact that fonts were disfluent meant that an effect would be produced due to the text’s perceptual disfluency. Interestingly enough, metacognitive judgements of individuals were in favor of fluent word lists. This can be explained possibly, by cognitive load.
As cognitive load increases the subjective difficulty of a task also increases, causing deeper thought and processing. However, this cognitive load makes individuals less confident in their ability to execute the task due to its difficulty. In addition, Oppenheimer extended the results to high school classrooms, revealing a practical application of desirable difficulty.

The root of inspiration for the given study was Castel and colleagues (2011). In this case, desired difficulty was found by inverting words in a list. Inversion of words produced greater recall than regular upright words, even though judgments of learning (JOLs) of participants supported upright words. It is apparent in all of these studies that the “ease of processing” guides participants JOLs, in which case results typically contradict participant JOLs. When JOLs are low, it seems that the task is perceived to be more difficult, causing an increase in cognitive load, which produces the desirable difficulty.

Our study wanted to investigate disfluency in a modality other than vision. Therefore we chose audio stimuli, to see if a disfluency effect could be produced with word lists presented through sound. Rhodes and Castel (2008) produced findings related to memory and perceptual disfluency in sound which are important to note. During the study participants listened to words of either normal (fluent) volume, or a very soft (disfluent) volume. Free recall tests and JOLs were administered post word list. Although higher JOLs were found for loud words over quiet words, volume actually had no effect on recall. Although this is still a metacognitive misconception, it does not correlate with the mentioned study on inverted words. It could have been that volume is not sufficiently disfluent to induce a desirable difficulty. This illustrates yet again, that desirable difficulty does not always arise as a result of disfluency. So, does disfluency in audio information increase cognitive load, and consequently enhance recall? Answering this question is obscenely difficult, as it is still a huge grey area in terms of when a perceptual
disfluency is desirable. Thus, because testing of desirable difficulty is contingent on a perceptual disfluency, material development in regards to desirable difficulty is pivotal. It is apparent from current literature in regards to desirable difficulty that it is found in visual stimuli, but I am interested in producing findings which reveal a desirable difficulty while processing disfluent audio information. Aspects like the amount of cognitive load the perceptual disfluency places on the individual is an important aspect to take into account, if cognitive load is too high a desirable difficulty is not likely to arise. On the other hand however, if a condition is not disfluent enough in comparison the fluent condition, then a desirable difficulty is not likely to be found. Thus this topic stands on a very slippery slope, in terms of what will bring about proper results. Also, as the Bjork paper discussed, a rehearsal mechanism may be occurring during the processing of disfluent information, revealing that it requires more processing time in order to be properly integrated.

The goal of the current study is to prove that desirable difficulty can be found in the memorization of audio information. It is predicted then, that auditory word lists will enhance memory as a result of perceptual disfluency. The relevance of these findings can apply to schools with large lecture halls and sound systems. It is not atypical to have lecturers with bad microphones etc., and if difficulty is or is not desirable could provide insight into whether or not universities should make a point of investing in better sound equipment for their lecture halls. However, this does not totally apply to a classroom setting as we do not have visual aids. On the other hand, it could be applied to audio podcasts.
Method

Participants

50 Participants would be collected using amazon’s Mechanical Turk.

Design

Independent variables manipulated for the given study would be fluency of the given word, whereas dependent measures would be the recall rate of fluent vs disfluent words. In addition, JOLs would be measures in terms of what participants thought benefited their learning more efficiently: muffled (disfluent) or clear (fluent) audio recordings.

Materials

Participants will be exposed to two different lists of 30 words via audio. The two lists will differ based on fluency; one will be a muffled audio file list while the other list will be a clear (normal) audio list. Due to a within subject design, distractor tasks (most likely tetris) will be administered between each list exposure. Words used for the given study would most likely be two syllable, high frequency nouns compiled using http://elexicon.wustl.edu/query13/query13.asp.

Words would be recorded using some simple recording software like garage band. All words used will be recorded as a clear recording, and then the same words will be recorded again as muffled words, which will constitute the disfluent condition. Proper piloting of materials will be necessary. If the materials are too difficult to process, or too easy to process, a desirable difficulty effect may not arise. Whether or not a participant is presented with a disfluent word list or fluent word list first would also be counterbalanced. Participants should use headphones with proper sound quality, and participants should also be fluent in English as the recordings will most likely be in English.
Procedure

Participants will most likely open up a computer module online, and be asked to fill out some personal information related to their age, gender, and language fluency. Participants would then be exposed to one of the two lists in its disfluent or fluent condition. Each word would be presented for five seconds, in order to give participants a sufficient amount of time to process the word, especially for disfluent words. After the presentation of the words, participants will partake in a two minute distractor task (most likely tetris). After the distractor period, participants would be given a free recall period, and type out all the words which they remember hearing. This would be repeated again for the second word list. After the second recall, a Judgment of learning (JOL) test would be administered in order to collect metacognitive reports.
References


