## Table of Contents

**Effect of Presence and Immersion on Learning in a Virtual Environment** ........................................3-9
Alana Sanchez-Prak, Faculty Mentor: Jesse Rissman, Ph.D.

**Test of Social Exchange Theory** ........................................................................................................10-17
Ariana Taghaddos, Faculty Mentor: Martin Monti, Ph.D.

**Biological Motion Perception and Correlations with Autistic Traits** .................................18-20
Ashley Vu, Faculty Mentor: Hongjing Lu, Ph.D.

**Rule Based Thought: Logic and Language** .........................................................................................21-29
Brenda Gil, Faculty Mentor: Brenda Gil, Ph.D.

**Report on BRITE Center “Quit Smoking App”: Progress and Suggestions** ..................30-35
Christopher Yeh, Faculty Mentor: Vickie Mays, Ph.D.

**Defining Characteristics of Social Facilitation of Deductive Reasoning** .......................36-43
Elliott Kim, Faculty Mentor: Martin Monti, Ph.D.

**The Effects of Feedback and Length of Retention Interval on the Testing Effect** .................44-53
Hamoon Azizi, Faculty Mentor: Elizabeth Bjork, Ph.D.

**The Effects of Feedback and Length of Retention Interval on the Testing Effect** ..........44-53
Hamoon Azizi, Faculty Mentor: Elizabeth Bjork, Ph.D.

**Virtual Reality Language Learning - A Procedural Summary of AA2 with Scanned Participants** ........................................................................................................61-69
Hugo Shiboski, Faculty Mentor: Jesse Rissman, Ph.D.

**Can the Retrieval Difficulty Effect be Reversed?** .................................................................70-82
Rufei Fan, Faculty Mentor: Robert Bjork, Ph.D.
Perceptual Processing and Computational Lab ................................................................. 83-89
Tania Jarjoura, Faculty Mentor: Zili Liu, Ph.D.

Effects of Aging in Memory and Learning Strategies ......................................................... 90-97
Uyen-Ly Nguyen, Faculty Mentor: Barbara Knowlton, Ph.D.

Memory Measurements ........................................................................................................ 98-106
Uyen-Ly Nguyen, Faculty Mentor: Barbara Knowlton, Ph.D.

The Scientific Study of Human Consciousness Using Psychophysics and
Metacognition ....................................................................................................................... 107-113
Yasha Mouradi, Faculty Mentor: Chris Hak-Wan Lau, Ph.D.

**These research papers are written by undergraduate students as part of the capstone requirement for the Cognitive Science major.**
Effect of Presence and Immersion on Learning in a Virtual Environment

Alana Sanchez-Prak

For the six weeks of summer session A at the University of California, Los Angeles I assisted in the Rissman Memory Lab as a Research Assistant for a project examining the effect of presence/immersion on learning within a virtual environment. The use of virtual environments has become a viable tool in many fields including education, medicine and military use. With its many uses already in place, creating a more immersive environment is still a main goal of many researchers and developers today. However, with the time and money spent on making the latest technology more immersive for uses, its effect on learning outcomes based on the immersion and presence effect remains unclear. In this study, immersion is defined as the sense of oneself being within a virtual environment. Presence is defined as the sense of oneself “being” one’s avatar. Multiple studies have been done with conflicting results on the presence/immersion effect on better recall. Some report findings that supported better recall (e.g., Lin, Duh, Parker, Abi-Rached, & Furness, 2002; Mania & Chalmers, 2001) yet others report findings that presence/immersion did not support better recall (e.g., Bailey et al., 2012). However, both studies examined visuospatial memory with participants tested on recall of the virtual environment in which they interacted in where results can vary based on the construction or quality of the environment and not so much their immersion or presence. This experiment aimed to study the effects of presence/immersion on verbal memory instead. This was done to separate the memory task from the virtual environment upon which presence/immersion measures were based. By
doing this we hypothesized that increase immersion and the feeling of being present within the virtual environment will result in better recall of learned languages within the virtual environment. Although the study is still ongoing, the information below is based upon currently completed data analysis.

**Method**

**Participants**

Participants were 12, right-handed UCLA undergraduates who participated in the study either for cash compensation or study participation credits for their psychology courses. All participants had normal or corrected-to-normal vision, native English speakers, had no extensive experience in virtual environments, not prone to motion sickness, and did not have previous exposure to both languages used within the experiment.

**Procedure**

This experiment consisted of two sessions on location that were 24 hours apart and one phone on the 8th day of the experiment. On the first day, after signing a consent form for the study, participants explored one virtual environment in a mission exploration task, had one round of learning 40 words in either Swahili or Chinyanja, and tested on the language they just learned in round 2 and 3 of testing all in the same environment as the mission exploration task. After the first mission exploration task and the three language rounds, the participants went through a second mission exploration task of a different virtual environment, one round of Swahili or Chinyanja learning (the language not learned in the previous learning round), and 2
subsequent testing rounds of the second language learned all in the second virtual environment that was explored in the second mission exploration task. The order of the language learned as well as the environment it was learned in was counterbalanced. The virtual environments were custom made by multiple contributors within the Rissman Memory Lab.

On day 1, participants explored two different virtual environments. This was termed their “mission exploration task” in which immersion was the main focus of the task. In each virtual environment, participants explored a total of 9 areas of the “world” by approaching a floating marker, rotating in place to immerse themselves into the environment, and clicking on the marker for a path to the next area within the world. The name of each of the nine areas within the world was told to participants for a total of 3 times, once when entering the area, once before rotating, and once when exiting the area.

During the the first encoding sessions for each language, participants approached 40 virtual items placed on 40 virtual pedestals one-by-one and first said aloud the english translation of the corresponding item before clicking to hear the foreign word and then speaking it back for a total of 3 times per word. The second and third encoding rounds for each language were the testing rounds and participants’ responses were recorded. Rounds 2 and 3 were similar to round 1 except now an attempt to recall the foreign word or to say “Pass” after stating the english translation and before clicking on the item for additional encoding was prompted. Participants’ presence score were measured using the Fox, Bailenson, and Binney (2009) survey as well as immersion survey adapted from Slater et al (1998).
On day 2, participants had their 4th testing round for each language in the virtual environment similar to rounds 2 and 3 from the day before. The final test of all 80 words was done through a created program using PsychoPy and was not within the virtual environment but through mental reinstatement of the environments instead. In the PsychoPy task, all instructions were given through the program and participants were first given an imagination task where they were told the world and area to imagine that would correspond to the location of a word they learned in that environment. They were asked to close their eyes, mentally rotate in the instructed environment similar to the their mission exploration tasks while indicating rotational movement through specified keyboard buttons waiting for a beep to indicate the end of the imagination task. The language task then followed with participants being told the language along with the word to recall, asked to push a specified keyboard button indication confidence or not in their knowledge of the word, and then waited for a beep to attempt their recall. Their responses were recorded for later data scoring. Next, participants were asked to rate imagination vividness from very vivid, vivid, not very vivid and unsuccessful then given a distractor task before moving on to the next word.

Finally, the last portion of the experiment was a day 8 phone call in which participants were recorded and asked to recall all 80 of the words they learned from day 1 and day 2.

Results

Verbal recalls from language rounds 2, 3, and 4 of each language as well as from the day 8 phone call were recorded and scored by two different scorers. In the case when scorers
disagreed, the score given was averaged. Correct recall meant pronunciation as well as syllables order were stated correctly. An analysis of the data using an independent samples t-test showed that individuals reporting higher levels of immersion showed significantly greater recall on Day 8 compared to those who reported low immersion. There were no significant differences found between high and low presence groups.

<table>
<thead>
<tr>
<th>Immersion (Test 5)</th>
<th>Presence (Test 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High immersion: $M=.42$, $SD=.18$</td>
<td>High Presence: $M=.27$, $SD=.01$</td>
</tr>
<tr>
<td>Low immersion: $M=.17$, $SD=.15$</td>
<td>Low Presence: $M=.28$, $SD=.23$</td>
</tr>
<tr>
<td>$t(10)=2.597$, $p=.027$</td>
<td>$t(10)=0.56$, $p=.957$</td>
</tr>
</tbody>
</table>

The results displayed above give evidence that an increase of immersion during learning using virtual environments may increase long-term retention. This shows an importance of a sense of oneself being within the virtual environment in contributing to both the learning as well as memory from the virtual environment. The lack of a presence effect is not conclusive due to only two participants having reported high presence possible possible due to the first-person point of view within the environment that gave little exposure to their avatar self. The next stop for this
project is fMRI scans of participants’ brains during the non-virtual environment PsychoPy test to examine the neurological activity during the mental reinstatement of the virtual environments leading to the recall of the learned words. More data is still being collected to study this presence/immersion effect on learning and memory but the potential to increase better recall is far reaching for education use, military use and more.

References


Test of Social Exchange Theory

Ariana Taghaddos

The facilitation of deductive reasoning, specifically, the facilitation in the Wason Card selection task can be explained by the Social Exchange Theory which was created by Leda Cosmides and John Tooby. Cosmides’ argues that the theory plays a role in explaining the deductive reasoning process of humans. Social exchange, at a very basic level, is the process of exchanging something less valuable for something more valuable. This leads to the question of whether or not this intrinsic behavior is the cause of our reasoning mechanisms. Through Cosmides’ studies she found that within every social exchange there is relative importance of intentionality, presence of a cost, and presence of a reward yet this has yet to be tested on a larger more diverse set of participants. The idea that deductive reasoning is an evolutionary distinct modular process will be tested through contemporary online medium. We plan to conduct a large scale test of Social Exchange Theory using Qualtrics for survey questions and pre-screening measures such as an Advanced Raven’s Progressive Matrices (ARPM), which screens for abstract reasoning and participants will be recruited through Amazon's Mechanical Turk system. Our findings will find evidence that either disproves or further proves whether deductive reasoning is a modular and distinctive cognitive process.

In order to see whether human cognition utilizes rules of logical inference when given a set of conditional rules, researches use the Wason selection task. In the Wason selection task, participants are presented with either a modus ponens conditional rule that says “if P then Q” or a modus tollens conditional rule that says “if not P then not Q”. Participants are then shown four cards that either follow or break the rule. For example, if one side of the card has property P then the other side has property Q
and if one side of the card does not have P then the other side does not have Q. The selection task asks participants which cards they should turn over in order to find what scenarios would violate the conditional rule. For example, a participant would be given a modus ponens conditional that says “if there is a G on one side of the card, then there is a three on the other side” and then would be asked which of the following cards would you turn over to test this rule.

Choosing card one would be modus ponens and choosing card two would be modus tollens due to the fact that if the participant turned over the eight card there should not be a G on the other side because if there was a G, then a three should have been on the other side.

Testing modus ponens and modus tollens are, however, abstract theories of deductive reasoning and do not address specific cases and contents. Participants' performance on the Wason selection task has proven to depend on the content of the task. The previous example would be arbitrary for a participant, which basically says “if the card has a specific letter, there is a corresponding number on the other side.” A less arbitrary conditional rule, which would seem more familiar to a participant would say “if you drink alcohol then you must be over twenty-one.” Participants tend to perform better on less arbitrary conditional rules. But why is it that we are suddenly significantly better at deductions when talking about things more socially familiar to us?

There are two main approaches to content specificity:
Cosmides believes that the reason for this better performance on the Wason selection task was due to the fact that the rules are placed in the context of a social rule that is asked to be followed. Under Cosmides’ social contract theory, we humans are good at modus tollens in social settings. This is due to the fact that modus tollens is equivalent to recognizing a cheater. If an individual wants P, then they need to follow rule Q and if they do not follow rule Q then they will not be able to benefit from P.

Within a social exchange, a set of rules, or social contract is formed and when an individual defies these set of rules intentionally in order to gain a benefit for themselves, then this individual is considered a cheater. Humans are social species and cooperation is crucial for our survival and Cosmides says that through time we developed a "cheater detection module" to allow us to detect individuals that disobey social rules. Under this theory, we are a specialized species and the computational complexity we face can be solved using this module. In order to detect cheaters, Cosmides suggests we need “content-specialized logic” or specialized deductive reasoning in adaptive, significant domains of human life, such as social exchange (Cosmides, 1992) and hazardous situations (Cosmides & Tooby, 1997). Cosmides argues that social exchange knowledge is due to “Darwinian algorithms” which provide human beings with “specialized learning mechanisms that organize experiences into adaptively meaningful schemas” (Cosmides 1989, p. 195). This is the reason that performance is better on the Wason selection task when it is related to social settings and social exchange.

Another approach to this content specificity, an opponent of the social exchange theory, is the pragmatic reasoning schema theory proposed by Cheng and Holyoak. Both the social exchange theory and the pragmatic reasoning schema theory agree that there is a content specificity to reasoning and that “this knowledge is more general than specific remembered experiences. The difference between these theories comes when dissection the origins of reasoning schemas. While Cosmides attributes social
exchange knowledge to evolution, Cheng and Holyoak argue that pragmatic schemas like “abstract knowledge structures are induced from ordinary life experiences” (Cheng & Holyoak, p. 395). This theory focuses on the fact you learn schemas that you slowly learn as you grow up. As you grow up you learn about permissions of things you can and cannot do. A pragmatic reasoning schema consists of general (non-abstract) rules that relate to particular kinds of goals such as permissions, obligations, and causation (p.2). A permission rule is a conditional that satisfies the schema “in order to take action A, a precondition X must be satisfied. Therefore, all social contract rules are permission rules but under the pragmatic reasoning schema, many permission rules are not social contracts. The theory says reasoning is not all about social contracts, we as human beings are good at permissions because we learn about them through a variety of experiences as we grow up.

Despite oppositions to Cosmides’ social exchange theory, her findings are inarguably interesting to test. Her experiment involved varies Wason card selection tasks. Her participants were all college students and when presenting these tasks, she altered the content and factors of the task. These factors include intentionality, benefit, and possibility to cheat. She hypothesized that if all three of these factors were included in the task then performance would be much better. She first varied the range of benefit among her permission rules which as stated earlier, should increase detection of cheating. She had three scenarios, one with high benefit, one neutral, and one not beneficial. Since benefit does not have a place in the pragmatic reasoning schema theory, under this theory, performance should be equivalent across all variations. However, Cosmides found that performance was best on the most beneficial permission rules. Following this study, Cosmides conducted another experiment that altered which included all three factors and found greater ability to detect cheating when intentionality, benefit, and ability to cheat were present.
Our study will be similar to that of Cosmides’, however, we are looking at a more diverse and larger group of participants.

Our stimuli will try to prove or disprove Cosmides’ social exchange theory. We are still creating our stimuli but we have decided to make the scenarios at least familiar as possible. The scenarios will fall into two categories, they will either portray a setting of social exchange or hazard management. Within the social exchange setting, there are six elements that will be changed: the presence or absence of a benefit, intentionality or lack of intentionality, and the ability or the inability to cheat. By keeping the scenarios identical other than these factors, we are eliminating familiarity playing a factor. Here is an example of a stimuli that either has presence of a benefit and absence of one:

Benefit:

In another part of the universe, there is a world populated by Ents. An Ent is a giant talking tree which has a life span of a thousand years. Ents are social creatures which interact with each other to form communities. However, only one Ent can lead each tribe and a new leader is only elected when the current one passes away. The leader is selected by a council of elders who will give the potential candidates a special golden flower. A candidate would have total control of the Ent community for a period of one year to test their leadership skills. However, there are some Ents who attempt to claim the candidate position by displaying a similar flower that they found in the forest and painted yellow. The candidate rule in the Ent community is:

“If you are a candidate, you must have an authentic golden flower.”

Imagine that you are a space traveler visiting this planet and the current leader has just passed away, and you are interested in whether the Ents obey this rule. Each card below represents an
Ent. On one side it sells what flower they have, and on the other side it says what claims they are making about being a candidate.

No benefit:

In another part of the universe, there is a world populated by Ents. An Ent is a giant talking tree which has a life span of a thousand years. Ents are social creatures which interact with each other to form communities. However, only one Ent can lead each tribe and a new leader is only elected when the current one passes away. The leader is selected by a council of elders who will give the potential candidates a special golden flower. A candidate would receive no special privileges and will be treated like any other Ent. After one year, the leader of the tribe will be chosen at random from another group, and no candidate will become the leader. However, there are some Ents who attempt to claim the candidate position by displaying a similar flower that they found in the forest and painted yellow. The candidate rule in the Ent community is:

“If you are a candidate, you must have an authentic golden flower.”

Imagine that you are a space traveler visiting this planet and the current leader has just passed away, and you are interested in whether the Ents obey this rule. Each card below represents an Ent. On one side it sells what flower they have, and on the other side it says what claims they are making about being a candidate.

The underlined portions are altered while the rest of the scenario stays identical. Both stimuli are followed by:

Indicate which card(s), if any, must be turned over to find out if anyone has broken the candidate rule.

<table>
<thead>
<tr>
<th>Claims to be a candidate</th>
<th>Has an authentic golden flower</th>
<th>Does not claim to be a candidate</th>
<th>Does not have an authentic golden flower</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>


By counterbalancing across all six conditions, we will be able to see which result in better performance and which do not. In the hazard management scenarios, the setting will be identical to the social exchange but the underlined portions will be varied upon efficacy or non-efficacy, hazard or no hazard, and precaution or no precaution.

Since we are still in the process of creating more stimuli, we will need to think of different scenarios and altering the factors without affecting the results of the study. In addition to creating stimuli, we are working on creating pre-screening tests through Qualtrics. Participants will be found from across the world using Amazon Mechanical Turk which will allow us to share our study with individuals and we will compensate them financially in return. The reason for looking for a diverse group of participants is that the Wason selection task and deductive reasoning has been widely studied among college students, which is only a small percentage of our population. By looking at all different people from different backgrounds, we will be able to prove or disprove Cosmides’ theory on a greater and more applicable level. We are also still in the process of deciding what pre-screening tests and surveys we want participants to fill out before participating in the main study. So far we have decided on the Advanced Raven’s Progressive Matrices and will be putting that into Qualtrics within the week.
Works Cited


Biological Motion Perception and Correlations with Autistic Traits

Ashley Vu

The processing of biological motion -- how others walk, move, or interact in a space -- is a fairly top-down mechanism. In comparison to inverted biological motion, upright biological motion is often much easier to extrapolate, and, thus, attend to. Nevertheless, past research has suggested that those clinically-diagnosed with ASD may not process these same social movements or interactive cues in the same manner. In fact, this kind of global processing of facial or biological motion may be greatly diminished (Cusack et al., 2015).

The present study induced a binocular rivalry stimulus by presenting to participants superimposed salsa dancers, one of which was upright and the other which was inverted. If a participant was better able to process biological motion, then the upright dancer was proposed to be, as a whole, more visible throughout each trial. The study aimed to investigate correlations between these visual biases toward biological motion and AQ score.

Methods

Materials

A chin rest and adjustable stereoscope was placed in front of a computer screen monitor, which was used to display stimuli created from the Psychophysics Toolbox. The presented stimuli displayed a black screen with two rival point-light male salsa dancers, one of which was inverted and presented to one eye, and the other of which was upright and presented to the other. The upright and inverted dancers each
consisted of nine randomly-sampled point lights, the topmost representing the head, and the other eight representing the localized joint and limb locations of the arms and legs. The dancers comprised of several fast-moving frames so as to simulate a salsa-dancing motion that lasted 26 seconds per trial. There were 32 trials. Each dancer alternated between displaying red or blue dots, and overlay one another when viewed through the stereoscope.

**Procedure**

Participants were instructed to adjust the chinrest and stereoscope to their best comfort before beginning the task. Once started, participants were presented with the dichoptic stimuli (both the red or blue inverted and upright male salsa dancers) and asked to signal whether the red or blue dancer was more dominant by pressing and holding down the keyboard’s right arrow key for red and left arrow key for blue; if both dancers appeared equally dominant, the participant was to press the down arrow key. Participants were to hold the keystroke for as long as their perception remained unchanged in one trial.

After all trials were completed, participants were given an Autism-Spectrum Quotient (AQ) questionnaire.
Rule Based Thought: Logic and Language

Brenda Gil

Introduction

The study at hand seeks to examine how interactions between the language processing regions of the brain, primarily Broca’s area, and the prefrontal cortex may play a role in deductive reasoning in humans. According to Hauser, Chomsky, and Fitch (2002) language involves ruled based thought and may have developed through evolution. Similarly, deductive reasoning also involves rule based thought and could possibly stem from the ability to comprehend linguistic rules. Therefore, in this experiment we seek evidence of either a possible relationship between the ability to comprehend rules of language and comprehending rules of logic or a lack of relationship indicating that these two cognitive systems may be independent of each other. Monti and Osherson’s (2012) review of neuroimaging studies led to the conclusion that the activation of language regions such as BA44/45 (Broca’s area) is a result of the lack of controls for the subject reading the stimuli presented. Therefore, they believe BA44/45 is not critical for the processing of rules of deductive logic. Overall, the goal of the following study is to demonstrate that deductive reasoning is independent of language processing areas. At the current time, the study is focusing on training and certification using the equipment and will enter the experimental phase later this summer. The following discussion will emphasize TMS and the study design.
The present study utilizes transcranial magnetic stimulation (TMS) in order to circumvent the limitations of fMRI technology that do not allow for causal conclusions. By utilizing TMS we are able to manipulate brain region activation as an independent variable to better understand it’s processes. Therefore, with TMS a temporary artificial lesion may be afflicted on the structure being manipulated (Miniussi, Ruzzoli, Walsh 2009). In this way, a double dissociation may be interpreted and causal information can be drawn from the results. Therefore, based on the aim of this study, we will alternate the disinhibition of Broca’s area and the prefrontal cortex to measure their involvement in deductive reasoning compared to baseline tasks with both areas activated.

Transcranial magnetic stimulation exerts a magnetic field that penetrates the scalp and skull in order to modify neural excitability in the subject’s brain (Rossi 2009). Depending on the parameters used, neurons can be stimulated or inhibited. In the current study, we will be using a form of repetitive TMS called theta burst stimulation in which short 50Hz bursts of repetitive TMS are repeated at a high 5Hz stimulation frequency that results in long term potentiation as proven through studies with rodents and the human brain (Oberman 2011). This method is also known as continuous TBS (cTBS) which results in inhibition of neural firing. Figure 1 provides a visualization of the type of rTMS used in this study. After application of TBS is complete, inhibition effects will last for approximately 30 minutes.

Safety concerns to keep in consideration when using TMS and recruiting participants include: seizures, syncope, burning of the scalp as a result of the heating of the TMS coil, displacement of head implants due to the magnetic field pulses, hearing alterations such as
tinnitus or aural fullness, neck pain due to forced posture or head immobilization, and other unforeseen events (Rossi 2009). In regards to the present study, hearing protection will be offered and mandatory when stimulating Broca’s area due to its proximity to the ear canal; the equipment utilized possesses a vacuum cooling system to prevent the coil from overheating; and neck pain may be prevented by monitoring the patient’s comfort levels from time to time. Criteria for exclusion will include alcoholism, sleep deprivation, metallic hardware and implants that may come into contact with the coil, history of epilepsy, and drugs that lower the seizure threshold.

Method

Participants

Participants will consist of 20 individuals 18-50 years of age recruited through previous TMS and MRI studies, SONA, and flyers posted in Franz Hall. Participants must perform the tasks with 60% accuracy or better to be included in the experiment. Compensation for completion of the study will be $25 or $10 per hour if the participant decides to withdraw. A telephone screening interview will be conducted to ensure that the potential participant does not meet any criteria for exclusion. The participants may not have past seizures, first degree relatives with seizure disorders, and must not be taken medication that may lower seizure thresholds. In addition, participants must be native English speakers, right handed, and have not had any formal training in deductive reasoning. If the potential participant has any further questions they will be directed to Allan Wu, MD.
**Design and Materials**

This study will consist of a 3-way within-participants design. The three independent variables consist of the brain region receiving TMS, the type of inference required by the task, and the degree of embeddedness of each type of inference. Each of these three variables will have two levels: the brain regions consist of BA8 (medial frontal cortex) or left BA 44/45 (Broca’s area); the two types of inference are deductive or linguistic; and the two degrees of embeddedness are one disjunction/conjunction or two disjunctions/conjunctions. All participants will experience cTBS to inhibit the brain regions. The three dependent variables to be measured are accuracy, response time, and change in accuracy and response time from the set of baseline trials.

The stimuli will consist of 96 pairs of sentences utilizing abstract elements W, X, Y, and Z. Half of these pairs will be logical premises with conditional deductive arguments using the connectives “if”, “then”, and “not”. The remaining half will be linguistic statements that avoid using connectives such as “saw”, “followed”, and “gave” instead. In regards to the degree of embeddedness variable, half of the stimuli pairs will have one disjunction/conjunction and the other half will have two disjunctions/conjunctions. These conjunctions and disjunctions will be expressed with the terms “or”, “and”, “neither”, and “both”. In addition, half of the pairs of stimuli will be valid and the remaining half will be invalid meaning they will have the same meaning or a different meaning. The baseline subset will consist of 24 grammar trials and conducted before administering TBS to the participant. Figure 2 depicts examples of the various
types of stimuli presented. Stimuli will be presented in black text on a grey background using Psychopy software. Each pair of arguments will be presented simultaneously with one statement above and the other below a horizontal black line.

Hypotheses

It is hypothesized that inhibition of BA44/45 (Broca’s area) will result in lower accuracy and longer response times in linguistic inference trials while accuracy and response time for trials with deductive inferences will not be affected compared to the baseline task. Furthermore, inhibition of BA8 (medial frontal cortex) will result in lower accuracy and response time for deductive inferences while accuracy and response time for linguistic inference trials will not be differ from the baseline task. In addition, it is predicted that when BA44/45 is inhibited accuracy will be lower and response time longer in linguistic inference trials than when BA8 is inhibited. When BA8 is inhibited, deductive inference trials will result in lower accuracy and longer response times than when BA44/45 is inhibited. In regards to levels of embeddedness, higher levels should result in lower accuracy and longer response times. There should be a greater difference in performance between high and low levels of embeddedness in deductive inferences when BA8 is inhibited than when BA 44/45 is inhibited. Finally, there should be a greater difference in performance between high and low levels of embeddedness in linguistic inferences when BA 44/45 is inhibited than when BA8 is inhibited.

Procedure
Subjects will attend a screening session in which they will experience the stimuli without any interventions to measure their accuracy in each subcomponent. Participants that receive a 60% accuracy or higher will be scheduled for their first session in a week at minimum to avoid practice effects. If the participant does not have an anatomical MRI on record, they will be directed to receive one at the Staglin Center of Cognitive Neuroscience. The TMS sessions will be held at the UCLA Brain Mapping Center. At the first session, the participants will read and sign an informed consent form and receive a copy of the Human Subject’s Bill of Rights. Baseline trials will occur at the beginning of each session.

Before commencing TBS, the active motor threshold (aMT) must be determined for each participant. After swabbing with a conductive gel, electromyography (EMG) electrodes are placed on the right first dorsal interosseous muscle (FDI) for subjects undergoing TBS on BA44/45 or on the right anterior tibialis (AT) muscle for subjects undergoing TBS on BA8. A ground electrode will be placed on the wrist or on the knee depending on which muscle the EMG electrode is placed. Using EMG software, evoked motor potentials (EMP) from the FDI or AT muscle are recorded. First, the individual’s resting motor threshold (rMT) must be recorded using single TMS pulses at the region of the left motor cortex where the FDI or AT muscle representation is located. Brainsight software and the participant’s anatomical MRI will be utilized to locate the region on the motor cortex that will be stimulated. Single pulses will be delivered at the intensity of 60% maximum stimulator output (MSO) at every 5 seconds. MSO will be reduced in 1% increments until MEP produced with amplitudes greater than 100µV occur
for 5 out of 10 trials. The rMT will be the lowest MSO needed to produce an MEP in 5 out of 10 trials. To find the participant’s aMT, they must gently squeeze a small item to generate a consistent EMG signal of 100µV. When the muscle is contracted, the Signal software will depict noisy waves of MEP. When single pulse TMS is applied to the motor cortex where the hand or leg representation is located, a silent period between MEP signals will appear. Again, intensity is reduced in 1% increments until the lowest MSO is found for MEPs when a silent period appears in 5 out of 10 trials. The aMT will be the lowest MSO needed to produce a silent period in 5 out of 10 trials. This process should only take 10 to 20 minutes.

Continuous TBS will be administered at 80% of aMT. During cTBS, short burst of 50Hz rTMS will be delivered at a rate of 5Hz for 40 seconds. Following this treatment, each stimulus pair will be presented for 15 seconds. Participants will be indicated to press “V” if the pair is valid and “N” if the pair is invalid. The total task will take approximately 30 minutes allowing enough time before the effects wear off. Participants with aMT equaling 55% MSO or higher will be excluded due to safety regulations requiring TBS to be administered at less than 45% MSO.

We are currently in the final stages of developing the stimuli. Within a few weeks participant recruitment will begin and the experiment will continue into Fall quarter.
Figure 1. In the right column, under patterned TMS, the patterns of cTBS and iTBS are illustrated.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Logic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If X then not Y.</td>
<td>If X or Y then not Z.</td>
<td>If W or X then neither Y nor Z.</td>
</tr>
<tr>
<td>If not X then Y.</td>
<td>If Y or X then not Z.</td>
<td>----------------------------------</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X saw Y.</td>
<td>X saw Y watching Z.</td>
<td>W saw X watching Y follow Z.</td>
</tr>
<tr>
<td>Y was seen by X.</td>
<td>X saw Z being watched by Y.</td>
<td>X was seen by W watching Y follow Z.</td>
</tr>
</tbody>
</table>

Figure 2. Examples of pairs of stimuli presented.
References


This report details the progress of the Quit Smoking App so far and comments on changes that still need to be made. I am considered a primary tester of UCLA BRITE Center’s “Quit Smoking App,” designed to help Korean teens quit smoking. We utilize cognitive science mechanisms proven by research to create a phone app that aims to trump California quitline in a randomized clinical trial. The app is designed to be a six weeks’ programs for the target participants. The participant engages in a particular session each week at the time of their choosing. Each session should not last more than an hour. Outside the session, participants will keep track of the amount of cigarette smoked, and can play challenges embedded in the app for reward. Keep in mind that this report is based on the iOS version. There may be variations on the Android version. In this report, I will be going over the main problems by milestones, organized as follows: milestone stone two deals with architecture, design and basic framework; milestone three deals with introduction, password, home page, and challenges; milestone four deals with my health, my stats, distract yourself, and settings milestone five deals with the first and second session; while milestone six deals with sessions three, four, five and six.

The list of menus on the screen are home, sessions, my calendar, my smoking album, my health, challenges, resources, settings, and instead of smoking. My smoking album is meant to help the
participant keep track of their smoking patterns. They are to take a picture each time they overcome a craving. My health page gives a list of trivia about the benefits of not smoking. Instead of smoking page gives a list of links the participants can click on to distract themselves from smoking, such as Facebook or Youtube.

Here are three things that I would suggest to change for the app as a whole. First, I highly recommend implementing an interactive walkthrough tutorial of the full app the first time the user opens the app. Given the ambiguous wording of the app and design, it is hard even for me to figure out what each page does. Second, I suggest that we add a small help button next to each function to guide what each button does. Third, a lot of the labels in the app is too ambiguous, a better wording for the labels is needed. For example, it is hard to understand what “My Health” page does. A better wording for that would be “Smoking trivia”.

Milestone three has several problems. First, the games and challenges are completely absent and not functional. Second, when entering the app, the password keypad does not have an “enter” button to hide the keypad after the user finishes entering the password. As of now, the user needs to click outside of the keypad to hide it. This system is very inconvenient. Please implement these two functions regarding the
keypad: when the user finishes typing his password, the keypad would hide automatically and unlock the app; second, implement a confirm button on the keypad for the user to touch when he finishes entering the password. A similar problem can be seen on the home screen. When the menu list is up, the user needs to touch the “show menu” button, symbolized by the three lines on the top left of the screen to hide it. This mechanism is a highly unusual design. Please make it so that the menu list can be hidden when the user press anywhere that is not the menu button, just like the way any other similar app implements this. Next, the “I smoked” button on the home screen is very confusing. It was meant for the user to click on when he smoked, but it is implemented under the “time since last cigarette:” label. The button may cause the time since last cigarette count to change, but we need to 1. Change the label of this button, and 2. Move the button somewhere else to avoid confusion. In the challenges screen, three of the buttons are shaded blue, and the two other buttons are shaded white like the background. These buttons look like labels. Please make them look like buttons by shading them.
Below are comments on milestone four. Firstly, the “my stats” button and end of the day push notification is non-functional. There is also no “Congrats” popup as detailed in the protocol when a user earns points. In the “instead of smoking page” the, capitalized buttons seem to have bigger fonts than the others, please make them consistent. The links in distract yourself do not work, and neither is the games in distract yourself. Next, there should be a logo in each box for “other things you can do besides smoking”. In “my smoking album”, after clicking submit, it says “submitted successfully” but nothing changes. In the calendar screen, every time we press a date the system interprets the action as wanting to change the session time, but the user should be able to press the dates just to see events planned for that date. There should be a separate button for changing the session date. In the “My Health” page there lack pictures, and the slide bars is very confusing. There are non-slidable slide bars for no reason. I would suggest the following: implement a slidable slide bar to represent time and when the user slide through the bar different texts appear. For example, at the twenty minutes mark, the screen changes and say “blood pressure decreases, temperature increases to normal”, and at the eight hours mark, the screen changes to say “carbon monoxide levels go back to normal, oxygen levels increase.”
Milestone five concerns the first two sessions. Firstly, each page of a session should have a button to go back to the first page, the content page, whenever the user wishes. The user should also be able to access the menu and exit when they wish. Then, the content page really should read “Contents.” Moreover, session 1 section 6 triggers 14 through 1, session 2 section 4 to 6, and the congrats screen are not functional. Section 1 trigger 4 “share” and “hear other responses” buttons are not working. In session 2, setting quit date only the “selecting a date” button works, but sending texts and contract is not implemented. Also, the push notification is not consistent. The push notification should come on every day morning and evening, with the time set by the user. Lastly, the sessions should be locked until the date that was set.

Milestone six concerns the last four sessions. Firstly, there should be push notification after 30 minutes of inactivity. Then, there is a list of sections that simply are not functional: session 3:2, 3:4-7, 4:4-6, 6:5. More specifically, team formation at the end of session three and the reveal page at the start of session four is not functional. Then, in session 6-4, the next button loops back on itself.

Most importantly, the design requires major revision, please use a more vibrant color and background as shown here. Please use the background that we provide with the designs of the
buttons and labels as provided. The menu should look like the one provided here as well, with logos on each screen. Every other page should have the same style.

Overall, as a frequent app user and developer myself, I want to say that the app is still very crude. It will not entice teenagers at all. They will probably only use the app if forced to. Major revisions are still needed. I think the most important thing we need to worry about is the design. Because we are trying to attract teenagers, a simple white background with ambiguous font and wording can really devalue an app that we spent so much time on. Even if the cognitive mechanisms were right, the first impression will scare away the user.
Defining Characteristics of Social Facilitation of Deductive Reasoning

Elliott Kim

Deductive reasoning can be defined as a logical process by which we come to a conclusion based on a set of assumedly true premises (e.g., if p, then q). Two common arguments of deductive reasoning are modus ponens (e.g., p, therefore q) and modus tollens (e.g., not q, therefore not p). It has been shown that modus ponens is a much easier deduction than modus tollens. One psychology experiment that exemplifies this observation is the Wason selection task, wherein one must choose two cards to examine in order to test a rule about the cards (examining one card involves modus ponens, and the other involves modus tollens): participants consistently examine the correct modus tollens card much less frequently than they examine the modus ponens card.

This selection task involves a rather unfamiliar scenario: one would not expect a direct application of this situation to everyday life. However, it has been shown that under certain circumstances, particularly social circumstances, the frequency of correctly applying modus tollens increases dramatically compared to the relatively abstract Wason selection task. The explanation for this has been hotly debated for some years now, with two main conflicting hypotheses being tested and argued for and against. One such hypothesis, put forward by Leda Cosmides et al., is termed “social contract theory”. This theory takes an evolutionary perspective, and states that the reason humans reason better in a social context is due to “social contract algorithms: a set of programs built by natural selection for reasoning about social exchange” (Cosmides et al., 2010). In other words, our capacity for social reasoning is highly specialized (supporting a modular hypothesis of the brain), selected for by application to everyday social exchanges and the need to detect “cheaters” in social exchange. Initially, Cosmides
presents her theory as a purely cost-benefit analysis: “A social contract related *perceived benefits to perceived costs*, expressing an exchange in which an individual is required to pay a cost (or meet a requirement) to an individual (or group) in order to receive a rationed benefit from that individual (or group)” (Cosmides, 1985).

This definition came under scrutiny by Cheng and Holyoak, who found fault with the requirement of an exchange between individuals among other aspects of the theory, and use a study of their own involving precaution but not social exchange to demonstrate their perceived flaw. They instead proposed a different set of schemas by which we reason: “Rather, they reason using abstract knowledge structures induced from ordinary life experiences, such as ‘permissions’, ‘obligations’, and ‘causations’ (…) termed *pragmatic reasoning schemas*” (Cheng & Holyoak, 1989). They also claim Cosmides’s theory to be a small aspect of their broader idea: “Her social contracts are therefore a subset of permissions (which are a subset of regulation schemas, which are a subset of pragmatic reasoning schemas)” (Cheng and Holyoak, 1989).

Cosmides’s response in 2010 refuted some of these criticisms. As an explanation to Cheng and Holyoak’s experiment involving precaution, Cosmides’s proposes an entirely separate schema evolved to deal with hazards: “these rules are processed not by social contract algorithms but by a different specialization with a distinct function: to monitor for cases in which people are in danger by virtue of not having taken the appropriate precaution” (Cosmides et al. 2010). She also employs experimental questions that do not differ in respect to the characteristics of pragmatic reasoning schemas (i.e., permission, obligation, and causation), but contain differences that apply or omit the requisites of social contract theory, in order to show that it is indeed the conditions of her social contract theory that
determine deductive reasoning performance, and not those of pragmatic reasoning schemas (which she refers to as “deontic logic”). Perhaps most importantly, Cosmides refines and furthers her definition of social contract theory to more specific conditions, and supports these inferences with experimental evidence: “First, intentional violations activate cheater detection, but innocent mistakes do not. Second, violation detection is up-regulated when potential violators would get the benefit regulated by the rule, and down-regulated when they would not. Third, cheater detection is down-regulated when the situation makes cheating difficult—when violations are unlikely, the search for them is unlikely to reveal those with a disposition to cheat” (Cosmides et al., 2010).

We wished to further investigate these three conditions (i.e., intention of violating the rule, presence of a benefit for violating a rule, and ability to violate the rule) and their effects on deductive reasoning. To this end, we are creating a study with questions to test the effects of each of these conditions in combination on deductive reasoning, specifically the ability to correctly apply modus tollens. In accordance with the results from Cosmides’s latest study, we hypothesize that the presence of each condition will increase the likelihood of correctly applying modus tollens, and that each condition have an additive effect (e.g., the intention of violation of a rule with the potential benefit of violating a rule will result in a higher correct percentage of applying modus tollens than the potential benefit of violating a rule alone). The null hypothesis is that these conditions have no effect on correctly applying modus tollens. We are also interested in the effects of other personal factors on deductive ability, particularly intelligence. To test and correct for this variable, we are including a problem set of Raven’s advanced progressive matrices, which has an ability to assess non-verbal intelligence, thereby testing equally among ethnicities. We hypothesize that intelligence will be positively correlated with deductive
reasoning ability, while the null hypothesis is that intelligence will not be correlated with deductive reasoning ability.

**Method**

**Participants**

We will recruit participants from the University of California, Los Angeles. Students will receive credit for certain psychology classes for their participation.

**Design**

This experiment uses a within subjects design. A combination of three variables will be manipulated: the intention to cheat, the ability to cheat, and the presence of a benefit of cheating. The number of questions with each combination has yet to be determined. Separately, there will be a test for intelligence and other possible confounding factors.

**Materials**

An online survey application, Qualtrics, will be employed for the purposes of this study. Participants will be monitored during the survey, as they will still have to take the survey in our lab for consistent environmental conditions.

The questions used to manipulate the independent variables will include a short paragraph describing an unfamiliar situation/community. The situation must be unfamiliar in order to minimize the possibility that participants will draw from their own personal experience during these questions. Certain rules, traditions and beliefs in this community will establish the independent variables. These sentences that establish the independent variables will be minimally altered between combinations of independent
variables. At the end of the paragraph, there will be a rule in the form of an if-then statement, which serves as the premise for the following deductive reasoning task (e.g., if there is an even number on one side of a card, then there must be a vowel on the other). The deductive reasoning task will be in a similar form to the Wason selection task, in that it will have four cards with only one side visible, with the participant being required to pick two to turn over (one modus ponens and one modus tollens) to verify that the rule stands.

There will also be a Raven’s advanced progressive matrices problem set (its placement has yet to be determined). Participants will have 15 minutes to complete 12 questions selected from the second set of the Raven Advanced Progressive Matrices (Arthur & Day, 1995). These problems consist of a 3 x 3 grid of visual patterns, with one of the nine squares missing. Participants will be asked to select from a number of possible squares that correctly complete the pattern, and their scores will be used to measure intelligence.

**Procedure**

Participants choose a time they are available to come to the lab to take the test from a schedule of possible times. They are seated at a computer with Qualtrics prepared to administer the test. When they are ready, participants start the test. After the test is complete, the study is over.

**Results**

As we have yet to run the study, no results can be reported. However, Figure 1 taken from Cosmides’s 2010 paper shows the general trend of what we expect from our combinations of variables testing for social contract theory. The results from her study display an additive effect of the number of conditions on accuracy on modus tollens application.
Discussion

As we have yet to run the study, there are no results to discuss. However, an additive effect of the number of conditions on modus tollens application accuracy would support Cosmides’s definition of social contract theory and its constituent factors. That being said, in order to support the evolutionary perspective of social contract theory, it would be more prudent to consider a biological approach. Cosmides et al. references a neurological study to demonstrate the modularity and independence of social exchange from precautionary reasoning: “Neuroimaging results and evidence that brain damage can selectively impair social exchange reasoning (while sparing precautionary reasoning) support the evolutionary hypothesis that these are two distinct specializations, not one superordinate deontic” (Cosmides et al., 2010). Refining a neuroimaging technique to distinguish this social contract reasoning as unique and activating independently of other types of reasoning would provide physical evidence of evolution’s effects that a psychological survey cannot. Distinguishing each of Cosmides’s social contract theory conditions as independent but additive neurological activation networks would further support Cosmides’s specific criteria for this specialized deductive reasoning.

References


Figure 1: Percentage of correct choice of cards depending on number and type of conditions present (B is benefit present, I is intention present, and A is presence of an ability to cheat) (Cosmides et al., 2010)
The Effects of Feedback and Length of Retention Interval on the Testing Effect

Hamoon Azizi (SSC)

The primary purpose of education is to transfer of knowledge to the students, in a way that this knowledge can be retained in the long term memory. Research has indicated that testing is a more effective way to increase long term retention, as relative to restudying material. Numerous studies have been conducted on how tests affect learning, memory, and recall; traditionally tests were considered as measures of memory, however research shows that they can be treated as learning events, and memory modifiers (Bjork, 1975). This phenomenon, known as the testing effect, which has received an outpouring of interest recently, states the conclusion that intermediate retrieval practice can increase the eventual recall of previously studied materials, more so than the intermediate restudy of the originally encoded material (Roediger, Putman, & Smith, 2011).

In a sequence of experiments, Storm, Friedman, Murayama, & Bjork (2014), conducted studies to test out the implications of the theory of disuse by Bjork and Bjork (1992). This theory suggested that sometimes, prior study may be assist the restudy to be more effective than retrieval practice. In their study, participants were presented with pairs of English-Swahili words, after the initial study phase, some of the word pairs were revisited and restudies, while others had only the Swahili word of the pairs presented and subjects were tested without feedback and asked to recall the corresponding English word. A delay of 1-week was used as a retention.
interval, and then the participants were tasked with completing several final tests, all of which entailed the recall of the English word once the Swahili associate was presented. Results showed that recall in the retrieval-tested condition was better in the first final test as opposed to the restudied items; as it was expected by the testing effect. An important factor of the design of this experiment is the provided feedback and additional study opportunities after each of the final tests, which in turn impacted recall on the second final test and reversed the expected testing effect, while maintaining a better recall performance in the restudy as compared to the retrieval practice condition. The same results persisted for the remaining four final tests.

Theses mentioned results from the Storm et al. (2014) signify that in order to measure the consequences of prior retrieval practice or studying, a single final test will not be enough, and more will be needed. As outlined, the addition of provided feedback through the restudy phases after the initial final test can possibly reverse the testing effect. Some apparent disadvantages of restudying in comparison with prior retrieval practice in the initial final test were due to a greater boost in memory strength given to the items by the retrieval practice as opposed to restudy; thus many of the items were still accessible in the recall threshold after the 1-week delayed interval. In conclusion, the Storm et al. (2014) results suggest that the initial restudy of already encoded material may be more effective for later recall than retrieval, given that the final test had feedback.

In another related study, Halamish and Bjork (2011) sought to answer similar questions about the limits of the testing effect. It was argued that the benefits of testing, even without
feedback, tend to be more visible at longer retention intervals, or when recall is tested instead of recognition. Final-test delay and final-test formant and the effects of testing versus restudying was a focus point, and it was further suggested that successful retrievals are a stronger way of learning than restudying, and also that they interact with the distribution of the memory strength. Difficulty of the initial test as dictated by its delayed retrieval or format, was claimed to increase for initial testing rather than restudying. Similarly, the difficulty of the final test was mentioned to determine the overall benefits of the testing effect.

Result of the Halamish and Bjork (2011) experiment supported their hypothesis that the interaction between final test’s difficulty and format benefited recall ability by initial testing rather than restudying. In sum, these data showed that testing and restudying skew and shift the items’ distributions as a consequence of their initial testing and restudying. Generally, the findings of this experiment and the Storm et al. (2014) are in agreement over the testing effect, and the role that factors such as feedback, retrieval practice delay, and final test formant play for such interaction.

This present study aimed to replicate the findings of the Storm et al. (2014) study, which claimed that provided feedback may impact the testing effect. Additionally, this research intended to examine if providing feedback after an initial test will reverse the expected testing effects, and if feedback can be more advantageous for already encoded material which do poorly in recall tests, or the ones that have higher recall levels. The design of this study allows to examine such hypothesis over a short and long retention interval between the retrieval-practice
test and the final test, and it is expected that the short retention interval produce better recall performance as compared to the long retention interval. Furthermore, it is also expected that initial retrieval practices which will produce better recall functioning on the first final test, given that the initial retrieval practice was difficult.

**Method**

**Participants**
Undergraduate students enrolled at University of California, Los Angeles participated in the study. There were 16 students in the sample: 11 were female and 5 were male (age range= 18-25 years, mean age = 20.5 years). Participation in the study received course credits.

**Design**
The study employed a two way, within-subjects design, with two levels of each independent variable, therefore resulting in a 2 x 2 within-subjects factorial design. The first independent variable was the timing delay between retrieval-practice and final test and it had two levels: short interval (30-minute delay: immediate), and long interval (1-week delay: delayed). The second independent variable was the final test: initial test, and the subsequent tests. The dependent variable was the number of correctly recalled associated English words on a ratio scale of 0 (none recalled correctly) to 1 (all recalled correctly).

**Materials and Apparatus**
Thirty-six Swahili-English pairs (e.g., Pombe-Beer, etc.) were selected. The words were basic nouns ranging from three to eight characters in length in both Swahili and English. All of
the Swahili words were pronounceable. The words were adopted from the Storm et al. (2014) study. The words were adopted from the Storm et al. (2014) study, thus they were already counterbalanced in subsets based on the appropriate strength of association levels. Each participant’s set of words was randomly divided into two sets, which were used separately for each delayed condition. The particular sets assigned to each condition was counterbalanced across participants. Standard laboratory computer workstations, all of which shared identical resources, such as hardware, software, user interface, and the same wired internet network were used to display all pairs in the center of a white background in Arial black font with a fixed font size.

**Procedure**

A unique identification code was assigned to each participant, this was done so the program would be able to use the remaining set of randomized pairs of words for the second part of the study (long delay interval), for each participant.

In the initial study phase of the experiment, participants were told that they will be tested on their ability to recall and remember the English words once the associated Swahili words is given as a cue. Participants studied the list of 36 Swahili-English pairs, which were presented one at a time on the computer screen for a duration of 10 s each.

In the initial testing/restudy participants were immediately given 3 repeated testing sessions after the initial study phase. Only half (18) of the original pairs were used, the cue words and a fragment of the target word was shown and the participants were asked to complete
the target word in the allotted 5 seconds. No feedback was given. In the short interval session, a 5-minute distractor task separated the initial testing from the delayed final tests.

In the delayed final tests participants while in short delay condition were again tested on the first set of 18 words, however a 2 second correct response feedback was added after the second final test. Upon returning a week later for the long interval delayed final tests, the participants were tested on the second set of 18 words (which was not restudied before), and like before, the 2 second correct response feedback was provided after the initial round of questions.

**Results**

Prior to discussing the obtained results and observed effects, it must be noted that the estimated required number of participants for this experiment (36) was again not achieved, and therefore the results discussed below are using the first 16 participants. This experiment will continue in the following available quarters and the final reported results and effect may be different than this overview paper.

Figure 1 presents the average number of correctly recalled associated English words on a ratio scale of 0 (none recalled correctly) to 1 (all recalled correctly) when the subjects were presented settings in which the 2 different timing schedules are compared as a function of their final tests. Looking at the pattern of results displayed in Figure 1, it appears that there were significant variations in the number of correct recalls for the initial finals test versus the subsequent tests in both immediate and delayed conditions. It is also apparent that there is a
significant difference of number of correct recalls between the delayed and immediate timing schedules.

To test these apparent effects, the data were analyzed using a two way within-subjects ANOVA, which revealed that the main effect final test was found to be significant, such that average participant’s number of correct recalls was significantly different $F(2, 32) = 101.91, MSE = .008, p < .05$. The main effect for the timing was also found to be significant $F(2, 32) = 31.55, MSE = .008, p < .05$. Additionally, Figure 1 indicated that there was an apparent significant interaction between timing and testing, $F(2, 32) = 4.56, MSE = .008, p < .05$.

To compare individual condition means, multiple dependent (paired-samples) t-tests, with a Bonferroni correction to maintain an alpha level of .05, were conducted. When initially tested, the average number of correct recalls was significantly smaller when the test was delayed ($M = 0.12, SD = 0.12$) than when the test was conducted immediately ($M = 0.37, SD = 0.25$), $t(16) = 6.28, p < 0.0125$. Similarly, the second successive delayed test’s measures were significantly smaller ($M = 0.45, SD = 0.27$) than the immediate one ($M = 0.61, SD = 0.30$), $t(16) = 3.70, p < 0.0125$. Likewise, the third final delayed test appears to have significantly smaller ($M = 0.58, SD = 0.27$) than the immediate one ($M = 0.70, SD = 0.25$), $t(16) = 3.31, p < 0.0125$.

**Discussion**

This primary purpose of the current study was to replicate the findings of the Storm et al. (2014) study: to show if provided feedback impacts the testing effect. Additionally, it was
intended to examine if providing feedback after an initial test will reverse the expected testing effects, and if feedback can be more advantageous materials which are already and do poorly in recall tests, or the ones that have higher recall levels. To test these hypotheses this study was devised to allow examinations over a short and long retention interval between the retrieval-practice test and the final test, and it was expected that the short retention interval produce better recall performance as compared to the long retention interval. Furthermore, it was also expected that initial retrieval practices produce better recall functioning on the first final test, given that the initial retrieval practice was difficult. The initial finding of this limited research did not support the expected reversal of the testing effect, however we did find that feedback positively and significantly impacted the recall after a delayed testing. However, once the required number of subjects have participated we may see that the final results do support the hypothesis.
References


Figure 1. Average correct response when the testing is done immediately versus when it’s done with a delay, as a function of the testing schedule and initial final test versus the subsequent ones. Significant main effects, and interaction were found.
The Effects of Feedback and Length of Retention Interval on the Testing Effect

Hamoon Azizi (SSA)

The primary purpose of education is to transfer of knowledge to the students, in a way that this knowledge can be retained in the long term memory. Research has indicated that testing is a more effective way to increase long term retention, as relative to restudying material. Numerous studies have been conducted on how tests affect learning, memory, and recall; traditionally tests were considered as measures of memory, however research shows that they can be treated as learning events, and memory modifiers (Bjork, 1975). This phenomenon, known as the testing effect, which has received an outpouring of interest recently, states the conclusion that intermediate retrieval practice can increase the eventual recall of previously studied materials, more so than the intermediate restudy of the originally encoded material (Roediger, Putman, & Smith, 2011).

In a sequence of experiments, Storm, Friedman, Murayama, & Bjork (2014), conducted studies to test out the implications of the theory of disuse by Bjork and Bjork (1992). This theory suggested that sometimes, prior study may be assist the restudy to be more effective than retrieval practice. In their study, participants were presented with pairs of English-Swahili words, after the initial study phase, some of the word pairs were revisited and restudies, while others had only the Swahili word of the pairs presented and subjects were tested without feedback and asked to recall the corresponding English word. A delay of 1-week was used as a retention interval, and then the participants were tasked with completing several final tests, all of which
entailed the recall of the English word once the Swahili associate was presented. Results showed that recall in the retrieval-tested condition was better in the first final test as opposed to the restudied items; as it was expected by the testing effect. An important factor of the design of this experiment is the provided feedback and additional study opportunities after each of the final tests, which in turn impacted recall on the second final test and reversed the expected testing effect, while maintaining a better recall performance in the restudy as compared to the retrieval practice condition. The same results persisted for the remaining four final tests.

Theses mentioned results from the Storm et al. (2014) signify that in order to measure the consequences of prior retrieval practice or studying, a single final test will not be enough, and more will be needed. As outlined, the addition of provided feedback through the restudy phases after the initial final test can possibly reverse the testing effect. Some apparent disadvantages of restudying in comparison with prior retrieval practice in the initial final test were due to a greater boost in memory strength given to the items by the retrieval practice as opposed to restudy; thus many of the items were still accessible in the recall threshold after the 1-week delayed interval.

In conclusion, the Storm et al. (2014) results suggest that the initial restudy of already encoded material may be more effective for later recall than retrieval, given that the final test had feedback.

In another related study, Halamish and Bjork (2011) sought to answer similar questions about the limits of the testing effect. It was argued that the benefits of testing, even without feedback, tend to be more visible at longer retention intervals, or when recall is tested instead of
recognition. Final-test delay and final-test formant and the effects of testing versus restudying was a focus point, and it was further suggested that successful retrievals are a stronger way of learning than restudying, and also that they interact with the distribution of the memory strength. Difficulty of the initial test as dictated by its delayed retrieval or format, was claimed to increase for initial testing rather than restudying. Similarly, the difficulty of the final test was mentioned to determine the overall benefits of the testing effect.

Result of the Halamish and Bjork (2011) experiment supported their hypothesis that the interaction between final test’s difficulty and format benefited recall ability by initial testing rather than restudying. In sum, these data showed that testing and restudying skew and shift the items’ distributions as a consequence of their initial testing and restudying. Generally, the findings of this experiment and the Storm et al. (2014) are in agreement over the testing effect, and the role that factors such as feedback, retrieval practice delay, and final test formant play for such interaction.

This present study aimed to replicate the findings of the Storm et al. (2014) study, which claimed that provided feedback may impact the testing effect. Additionally, this research intended to examine if providing feedback after an initial test will reverse the expected testing effects, and if feedback can more advantageous for already encoded material which do poorly in recall tests, or the ones that have higher recall levels. The design of this study allows to examine such hypothesis over a short and long retention interval between the retrieval-practice test and the final test, and it is expected that the short retention interval produce better recall performance
as compared to the long retention interval. Furthermore, it is also expected that initial retrieval practices which will produce better recall functioning on the first final test, given that the initial retrieval practice was difficult.

Method

Participants

Undergraduate students enrolled at University of California, Los Angeles participated in the study. There were 10 students in the sample: 8 were female and 2 were male (age range= 18-21 years, mean age = 20.0 years). Participation in the study received course credits.

Design

This study employed a one-way, within-subjects design, with two levels of the independent variables. The independent variable was the delay between retrieval-practice and final test and it had two level: short interval (30-minute delay), and long interval (1-week delay). The dependent variable was the number of correctly recalled associate English words.

Materials and Apparatus

Thirty-six Swahili-English pairs (e.g., Pombe-Beer, etc.) were selected. The words were basic nouns ranging from three to eight characters in length in both Swahili and English. All of the Swahili words were pronounceable. The words were adopted from the Storm et al. (2014) study. The words were adopted from the Storm et al. (2014) study, thus they were already counterbalanced in subsets based on the appropriate strength of association levels. Each participant’s set of words was randomly divided into two sets, which were used separately for
each delayed condition. The particular sets assigned to each condition was counterbalanced across participants. Standard laboratory computer workstations, all of which shared identical resources, such as hardware, software, user interface, and the same wired internet network were used to display all pairs in the center of a white background in Arial black font with a fixed font size.

**Procedure**

A unique identification code was assigned to each participant, this was done so the program would be able to use the remaining set of randomized pairs of words for the second part of the study (long delay interval), for each participant.

In the initial study phase of the experiment, participants were told that they will be tested on their ability to recall and remember the English words once the associated Swahili words is given as a cue. Participants studied the list of 36 Swahili-English pairs, which were presented one at a time on the computer screen for a duration of 10 s each.

In the initial testing/restudy participants were immediately given 3 repeated testing sessions after the initial study phase. Only half (18) of the original pairs were used, the cue words and a fragment of the target word was shown and the participants were asked to complete the target word in the allotted 5 seconds. No feedback was given. In the short interval session, a 5-minute distractor task separated the initial testing from the delayed final tests.

In the delayed final tests participants while in short delay condition were again tested on the first set of 18 words, however a 2 second correct response feedback was added after the
second final test. Upon returning a week later for the long interval delayed final tests, the participants were tested on the second set of 18 words (which was not restudied before), and like before, the 2 second correct response feedback was provided after the initial round of questions.

**Results / Discussion**

The estimated required number of participants for this experiment (36) was not achieved, thus the recorded data was not analyzed. Due to the insufficient number of participants, this present experiment will be continued throughout the next available quarter, and the results will be analyzed and discussed in a separate paper in detail.
References


Virtual Reality Language Learning - A Procedural Summary of AA2 with Scanned Participants

Hugo Shiboski

UPDATED FROM ORIGINAL TO INCLUDE fMRI PORTION

Virtual Reality Language Learning - A Procedural Summary of AA2

People have always been very fascinated with the idea of being remotely immersed in alternate environments than the one of the present moment. Something about being able to experience an exotic foreign world with the reassurance of being safely on a couch is very appealing. Recent advances in technology have made this possibility more immersive than ever. According to Joseph Psotka in a 1995 article of Instructional Science, VR technology has reached the level where user can easily be convinced that a virtual environment feels like real life. However, there are still drawbacks to even the most cutting-edge technology. For example, many users experience motion-sickness while experiencing virtual reality, a problem that engineers still have not been able to overcome (Psotka, 1995). Despite these challenges, virtual reality has been recognized as a potentially valuable educational resource. Psotka states that high schools have used VR as a platform for learning in maths, sciences and the humanities.

The present experiment is specifically interested in virtually reality as a tool for language learning. A 1998 study by Plass, Chun, Mayer and Leutner on visually-aided language learning found that students remember German words better when presented with both a visual and auditory cue simultaneously (Plass, Chun, Mayer & Leutner, 1998). The following is a summary
of the experimental procedures of an experiment that currently being conducted at UCLA under the direction of Jesse Rissman. The experiment’s goal is to gain more insight on the potential of virtual reality as a medium for language learning.

Methods

Participants

Participants are required to be 18-26 years of age to take part in the study. All participants are fluent English speakers and readers and are not bilingual (although some participants may have participated in high school language courses). All participants have normal or corrected vision and are in good mental health. Participants are made sure to not be prone to motion sickness. All participants are right handed. Each participant receives monetary compensation and/or UCLA SONA credit as compensation for their participation. Participants are required to fill out an fMRI scan form, ensuring that they do not have any sort of metal fragments or devices inside or on their body.
Design

The study uses a 2x2 mixed factorial design with two manipulated variables: learning environment and language. Subjects are randomly assigned to conditions.

Procedures

Participants are each randomly assigned a counterbalanced condition (language order, mission order, and language learning task environment). On the first day, before beginning the experiment, participants are asked to leave all possessions in the main lab area (separate from the lab room) and are instructed that cell phone use is prohibited for the entire duration of the experiment. Subjects are then instructed to read and sign a consent form. Each participant begins with a practice session that acclimates them to the virtual environment. They are given time to adjust to the controls on the computer keyboard, then are shown an instructional video that informs them on the language-learning procedure. The procedure involves approaching pedestals
that are placed in fixed locations in the learning environments. Each pedestal has an animated object floating and spinning above it with the English word for the object displayed alongside it. Participants are to recite the English title once before clicking on the pedestal, which triggers the word to be translated to the designated language (Chinyaja or Swahili) along with an audio recording of the word. This audio recording repeats three times, after which the participant is to click the pedestal again and proceed along a newly revealed path to the next pedestal. After the video, participants are given an opportunity to practice this with a couple of pedestals in the practice environment with the designated language being set as pig-latin (for practice purposes). Once comfortable with the VR and the language procedures, the participant proceeds to the first mission (corresponding to their assignment on the counterbalance sheet), consists of exploring one of the two language-learning environments (Fairlyland Gardens or Moon Base). Participants are told that the missions are purely for purposes of internalization, and the participant is asked to absorb as much detail as possible while exploring the environment. The experimenter, sitting in the same room as the participant on the other end of a screen with a monster to observe the participant’s actions, verbally guides the participant as they make their way through the virtual world. For each sub-area of the map, the experimenter calls out the name of the area three times: once as they are entering the area, once as they have arrived at the center, and a final time as they are leaving the area. The missions do not have pedestals, and are intended purely for the participant to internalize the environment. Once the mission is completed, the participant is given a two-minute break to use the restroom or have a drink of water. Following the break, the
participant begins round one of the first language task. The location of this task is always the same as that of the first mission. Language tasks utilize the instructions of both the initial practice task and the missions. Participants are instructed to explore the world following the same path as in the mission, but each sub-area of the world contains multiple pedestals that the participant must interact with as per the instructions in the practice session. There are a total of 40 pedestals (i.e., 40 words) in each world. In round one of the first language task, the participant simply learns the foreign-language words. No recall is involved in this task. Following the first round, the participant immediately begins the second. In this round, the participant is required to recall the foreign-language translation of the word after calling out the english translation, but before clicking on the pedestal for the first time. If they cannot recall the word or any part of the word, they are instructed to say “pass” before clicking the pedestal. All of the recall tasks are recorded (each round but the first, in which there are no pedestals). Before the third round, the participant is given another two-minute break. The third round is identical in procedure to the second. After all three rounds of the first language task, the participant begins Mission 2, which is identical to Mission 1 in procedure, but takes place in the second (whichever one the participant hasn’t experienced yet). Following Mission 2 is another five-minute break, and then
round one of the second language task which takes place in whatever environment the participant is assigned to on the counterbalance sheet (sometime, this is a repeat of the first language task). The participant proceeds through three rounds of this task, just like the the first. Round three of the second language concludes Day 1 of the experiment. Before the participant leaves, the experimenter asks them if they are available a week from the first day of the experimenter for a follow-up phone call which will serve as a recall test. The participant is told nothing about the purpose of the phone call other than that they will be asked some questions and that the phone call should take no longer than 30 minutes. Day 2 of the experiment takes the following day. The participant is again asked to leave all of their possessions in the main lab area before beginning the experiment. Day 2 consists of two separate parts. The first is a continuation of the Day 1 language tasks. The participant first completes one more round of language one recall (whichever language was learned first the previous day in whatever environment in which that task took place). Following this, the participant does one more round of language two recall (whichever language was learned second in it’s respective environment). The second section of Day 2 consists of a recall task in the fMRI which does not involve the virtual environment. The participants are shown an instructional video of the task. The video explains to them that they
will be put through imagination and language tasks. There is one imagination and one language
task for each word that the subject recalls, and the subject is required to go through all 80 words
from the language tasks (40 for each language). For the imagination task, the participant is given
an auditory cue with the name of an area (corresponding to a sub-area of one of the worlds), the
name of a language, and a word. They are then told to imagine themselves in there as vividly as
possible, and then mentally perform a 360° rotation in the given location, pressing number 1 of
the computer keypad at the start of the rotation, number 2 halfway through, and number 3 when
they have completed the rotation. They are given approximately 10 seconds to do this. They are
then prompted with a beep, after which they are to open their eyes. Following this is the
language task, in which they hear another beep and are given 3 seconds to recall the word to the
best of their ability or say “pass”. After these two tasks, there are two more tasks. The first
of these is a rating task in which the participant is asked to rate how vivid the mental rotation
felt by pressing a number from 1 to 4 on the keypad (1 being the least vivid, 4 being the most
vivid).

The second is a number task in which two numbers are displayed onscreen and the participant is
required to press the 1 key if the product of the two numbers is odd and the 2 key if the product is even. These 4 tasks are repeated for each word in an order that corresponds to their assignment on the counterbalance sheet. Before doing all 80 words, the participant is given a test run to make sure they have the process internalized. The participant is walked to the fMRI after this practice run and the participant is quizzed on their comprehension of the task on the walk over.

Upon arrival at the scanner, the participant’s weight and height measurements are taken and they are screened using a metal detector. Once clear to proceed into the fMRI, the subject is set up inside the machine by a research supervisor and a research assistant. The task is recorded using PsycoPy, a statistical data collection software. When the PsychoPy tasks have been completed, the participant exits the fMRI and fills out a post experimental survey and is done with the Day 2 of the experiment. A week later, on day 8, the experimenter calls the participant at the previously agree-upon time and the subject is tested on all 80 words.
References


Can the Retrieval Difficulty Effect be Reversed?

Rufei Fan

Learning and memory behaviors have always been intriguing topics. Prior research has focused on how testing affects learning and memory. As a very important background research, the “Testing Effect” is a widely studied finding that indicates tests improve retention (Roediger & Butler, 2011). For example, the study by Roediger and Butler examined the Testing Effect with the following procedures: At the study phase, all the participants were exposed with the same stimuli and study the same word pairs. At the practice phase, participants were divided into two groups. One group went through a restudy session and the other went through retrieval practice (no feedback). After a delay period (usually one week), all the participants took a Final Test and the practice group outperformed the restudy group.

The Testing Effect can be explained by the Bifurcation Model (Kornell, Bjork, & Garcia, 2011). As shown in Figure 1, the two columns represent the two conditions of Roediger and Butler’s study. A distribution of strength of memory of all studied items is shown on the graphs. After the study phase, a distribution of all studied items is formed. The B row shows the effect of the practice phase that both distributions shift towards the right for certain amount due to the strengthening of the memory. In the restudy condition, the whole distribution shifts. However, in the retrieval practice condition, only the successfully retrieved items get strengthened, and shifted for a large amount, larger than the restudy condition because to successfully retrieve an item requires more effort. The C row shows the distribution after the delay period. Forgetting the
studied items is inevitable so both curves shifted towards left. Note here, the solid line in the middle is the threshold of recall, which means that only the items that fall beyond this line can be recalled at this point. The initial shift in this portion was so large that most of this part still remained beyond the threshold after the regression. This is why the test group outperform the restudy group. Essentially, the Testing Effect revealed the phenomenon that intermediate retrieval practices between the study phase and the final memory test can significantly enhance the final test performance compared to the restudy trials do.

Another important background research focused on the Retrieval Difficulty Effect, which shows that the harder the practice test is, the better participants performed in the final test (Carpenter & DeLosh, 2006). In this study, participants were first asked to study related English word pairs in the study phase. In the following practice phase, participants were exposed to a problem on a computer screen asking them to generate one word from each of the previously studied word pairs with one cue letter provided. If they couldn’t solve it, they could choose to press the spacebar for one more cue letter, and they could do so three times. At last, participants received a final test phase to test their memory on studied word pairs without any memory cue. The result showed that words which were successfully retrieved during the hardest practice test (with one cue letter) were recalled most, and ones retrieved during the easiest practice test (with four cue letters) were recalled least.

The final background research focused on the Feedback Effect (Storm, Murayama, & Bjork, 2014). Similar to other research, in this study, there was a study phase, a practice phase,
and a final test phase. Participants studied Swahili and English word pairs. In the practice phase, half of the words were presented in restudy condition (simply showing the word pair again) and the other half were presented in retrieval practice condition. The final test phase consisted of 6 final tests in total, and feedback was provided after every final test question. The result shows that in the first final test, there was a Testing Effect that words which were tested in the practice phase were better recalled than ones not tested previously. However, after the first final test, the Testing Effect was reversed in that words which were simply restudied had a higher retrieval rate than tested words, which is known as the Feedback Effect. We can also explain the Feedback Effect by the bifurcation model: similar to the paradigm of the Testing Effect, Figure 2 shows that after Row C (where the Testing Effect occurs), the remaining part above the recall threshold is larger under the retrieval practice condition. However, after having been given the feedback of the first final test, both curves shift towards the right again because their memory is once again strengthened. The restudy condition curve was continuous this entire time, so that more items would now be shifted above the threshold, which led to the reversal of the Testing Effect (the Feedback Effect).

Now that we have gone through multiple previous studies focused on the learning and memory behaviors, and we know that because of the Retrieval Difficulty Effect, increased difficulty in practice tests enlarges the size of Testing Effect. However, under the conditions of practice by testing or by restudying, the larger Testing Effect that resulted from testing can be reversed by providing feedback in final retention tests. Therefore, one question can be raised
here: will the larger Testing Effect resulted from a harder practice test be reversed because of the Feedback Effect? In our study, we intended to examine whether the Retrieval Difficulty Effect will also be reversed when feedback is provided. On the basis of prior work, we hypothesized that a difficult initial retrieval practice would produce better recall performance on a first final test than an easy initial retrieval practice (Carpenter & DeLosh, 2006). Such a finding would be consistent with the bifurcation model which assumes that the magnitude of the increase in strength is a function of retrieval effort and, based on the view that difficult retrieval practice should demand more retrieval effort than easier retrieval practice, difficult retrieval practice should result in a larger boost in memory strength for successfully retrieved items than easy retrieval practice does. We also hypothesized that the Testing Effect would be reversed under the difficult practice condition. That said, the Feedback Effect was hypothesized to also be revealed.

**Method**

**Participants**

The participants were 34 undergraduate students from University of California, Los Angeles. All students participated in this study in order to earn course credit. All the participants had normal or corrected vision and understood instructions in English language in order to finish the experiment tasks.

**Design**

This study employed a within-subjects experimental design with one independent variable: the difficulty of retrieval practice phase. It consisted of two levels: difficult retrieval
practice test (prompt participants to generate the target words with no cue provided) and easy retrieval practice test (prompt participants to generate the target words with two cue letters provided). The dependent variable was participants’ memory of the words studied. It was operationalized as measuring the number of words each participant correctly recalled during the final test phase, out of a total of 30 word pairs which were studied in the preceding study phase.

**Materials and Apparatus**

The experiment was created using *Collector*, a PHP-based open source program which generated experiments on a web page, automatically showing stimuli on a computer screen and recording responses inputted by the participants. This experiment was administered in an experiment lab. All participants needed computers with mice, keyboards, and internet connection in order to finish the experiment tasks.

The materials consisted of 60 related English word pairs with association strengths of between .15 and .20 according to Nelson, McEvoy, and Schreiber (1998). This strength index indicates that when people were shown the first word in the pair, 15–20% of those individuals produced the second word in the pair as the first associate that came to mind.

**Procedure**

The experiment session consisted of four segments: a study phase; a practice phase; a one-week delay; and a final test phase.

At the first screen, participants were informed that they would study word pairs which would appear on a computer screen one at a time. Then they were shown the study list in a
randomized order, which was randomized for each individual when the program started, so the order for each participant was effectively unique. The words were presented on the center of the screen, at a rate of five seconds per word pair. After the study phase, there was a distractor task of playing the Tetris game for 45 seconds, which was designed to be just long enough to prevent rehearsal of the word pairs. Before the task began, there was an instruction intended to teach participants how to play the game.

After the distractor task, there was the practice phase when all the manipulation took place. Participants received a practice phase during which a random half of studied items (30 word pairs) practiced in a difficult retrieval test mode, in which participants were asked to produce the target word (the second word of the pair) with no letter cue presented; the other half of word pairs were practiced in an easy retrieval test mode, that participants were asked to produce the target word with one letter cue (the first letter of the word was shown on the screen) presented.

After the practice phase, there was a delay of one week till the second part of the experiment. After one week’s delay, participants came back to the lab and received the final test phase, which included three back to back hard retrieval test questions (produce the target word without any cue provided) on all of the 30 words. After each question, the feedback of the correct answer was provided.

**Results and Discussion**
Figure 3 contains the average number of words correctly retrieved during the three final tests as the function of the type of practice test (either easy or difficult). As shown in Figure 3, in the first final test, words were retrieved at a similar amount no matter how they were practiced ($M_{\text{Easy practice-Final test 1}} = 17.76, SD = .59$; $M_{\text{Hard practice-Final test 1}} = 17.69, SD = .59$). In the second and the third final test, words were retrieved more when they were practiced with easy practice tests than with difficult practice tests ($M_{\text{Easy practice-Final test 2}} = 27.67, SD = .92$; $M_{\text{Hard practice-Final test 2}} = 26.65, SD = .89$; $M_{\text{Easy practice-Final test 3}} = 29.3, SD = .97$; $M_{\text{Hard practice-Final test 3}} = 28.95, SD = .97$). A t-test was conducted to analyze the data, and the effect between the easy practice condition and the difficult practice condition was not found to be significant.

Looking at our hypothesis and results together, they didn’t show a significantly consistent pattern. Our experiment still had great room for improvement: first of all, we recruited a relatively small sample size. The small sample size might cause the lack of power of the data. Secondly, the task design seemed to be easy and the results showed a tendency of ceiling effect in the third final test, that almost all the participants reached 94% correct rate no matter under which condition the word was practiced. Therefore, the future direction of our study can be increasing the sample size and the difficulty of the learning task, such as using English and foreign language word pairs instead of English only word pairs.
References


Storm, B. C., Friedman, M. C., Murayama, K., & Bjork, R. A. (2014). On the transfer of prior tests or study events to subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(1), 115.
Figure 1. Testing Effect in Bifurcation Model.
Figure 2. Feedback Effect in Bifurcation Model.
Figure 3. Average amount of words correctly retrieved as a function of the type of practice test performed.
Experience

I used to think research work is far away from me, an undergraduate student who has no enough knowledge, skill, and resource to conduct a holistic research. However, the experience of working in the Bjork Lab helped me overcome this sense of distance. As a research assistant, I made friends, learned important skills, and participated in a serious academic research. I got to see how a research was being designed. I helped with organizing and monitoring each participant during the experiment. I learned how to collect, sort, and analyze data efficiently. I was also trained to present the study in a semi-formal conference, which was a fantastic opportunity to prepare me for the future academic presentations. Oliver, my research mentor, is a kind and outgoing scholar who is passionate about his field of study. He is supportive and helpful on all the questions I had. His enthusiasm is so contagious that I enjoyed the process of research very much.

Besides the research work I directly participated in, Bjork Lab is an awesome space that provide me with all kinds of support and resources. The weekly minifog meeting conducted by Saskia, our lab coordinator, effectively improved my skill of reading and analyzing academic publications. During these events, I made friends with other research assistants and scholars working in the lab. Everyone is friendly to each other, supporting each other, and open to advices. I appreciate the constant positive vibe of learning and encouraging hovering around in the lab. Moreover, Mr and Mrs Bjork, our lab mentor professors, are also very warmhearted and easygoing. They are very supportive to every member in the lab, and gives advices as our sincere
friends. So far, I have already been working here for two quarters. It would be my honor if I have
the chance to stay here to contribute and learn more.
Perceptual Processing and Computational Lab

Tania Jarjoura

Our brains are able to perceive and distinguish between two-dimension and three-dimension based on small visual cues. The threshold for this ability is still being explored and the brain’s ability to garner its own perception from a displayed image is truly a baffling quality. The brain is able to take the physical image and expand on it, distort, and amplify to make an individual’s distinct perception. While working in the Zililab, I have been surrounded by these perceptual ideas when experimenting with ellipses and cones.

In the lab, participants are looking at a rotating ellipse as well as a moving dot. They are looking at this moving figure and attempting to perceive a rotating cone. Once this image is perceived, the rotating and elongated ellipse serves as a smooth round base for a cone. How does this happen? How is the brain able to move between these two dimensions? With this idea, I decided to delve into the topic of perception and representation in the mind – how a physical object in the world is relayed and stored into the brain, how it is changed, and perhaps the new meaning that accompanies it.

All objects that we perceive in our mind are representations. When looking at a physical object in the world, that image does not get projected into the mind, rather it is our perception of its representation. This representation is grounded in the image witnessed, and is symbolic of the item in the real world. The physical object itself is referred to as a referent. The manner in which
the referent and the symbol are related is through its intentionality, or the mapping between the two objects that makes them meaningful (Friedenberg). Something interesting to note is that one will never know the symbol used in another’s brain to represent the same referent. It is variable and any form of introspection is an unreliable way to detect what is meaningful and symbolic to the person. This idea permeates into the Zililab study and raises an interesting question as to how some people are able to perceive a rotating cone as opposed to the spinning ellipse and dot. Perhaps some of the individuals have been engulfed in the study of geometry, their ability to perceive the rotating cone may have a lower threshold compared to an individual with little geometric or artistic experience. Just as with optical illusions that show both a young lady and an old woman, depending on prior experience and consequently, priming, some individuals may be more prone to see a certain shape.

Perhaps this ability of priming can be one exercised before performing the demonstration and experiment. Researchers can show a progression from a well-defined cone to slowly removing some of the lines and edges to identify the specific point where the shape is no longer perceived. At this point, we should be able to pin point at which point the figures moves from three dimension to two dimension. Will there be a consistent trend in the stopping point among all participants? Will each participant’s past play such a large role that makes such extreme variability in perception? Normally, individuals are able to retain images based on geons, which are the underlying geometric figures within an object. At first, it was supposed that images were held through the theory of feature detection, yet changing the lines and edges that defined each
feature ultimately had no effect on response time in a study of pigeons. However, there was a 
change in response time when changing the basic, single geon, showing evidence for memory of 
visual images to be held in these three-dimensional structures (Importance). Perhaps, the ability 
to recognize this geon in the priming before the experiment will help bolster the results of cone 
identification.

At this point there meets the crossroads between perception and the physical world where we 
attempt to find a mathematical baseline for our perception and a consistent formula in which the 
manner we perceive and store information. One must question the ability of the brain to derail 
from the norm -- the math, the geons -- and to take its perception and representation beyond the 
bare bones of the image provided. In the Zililab the dot on the rotating ellipse is positioned on 
either the major or minor axis which helps the participant perceive the image as a cone. Why 
does one axis more favorbly encourage this perception? Is there a mathematical formula that 
helps describe this phenomenon? This portrays the idea as the physical world as a generalized 
foundation for thee mind to build upon with creativity and imagination. The results of this cone 
identification are not always consistent, depending on the brain’s ability to imagine and look 
beyond the bare bones given. Perhaps this general image is present to give us humans the ability 
to create and imagine more complex things.

Expanding on this idea of imagination, the most interesting part of the lab for me is the ability to 
help the participant during the demonstration period and inquire on what they see. Some are 
completely engulfed in the spinning ellipse and dot that they cannot see beyond this direct image
to the cone. However, during the demonstration period, I am able to hint at and describe the cone form and see if the participant is able to see it. More often than not, they are, but only with my help. Once told, some participants can now only perceive the image as a cone and no longer a spinning ellipse and dot. This is the most fascinating part of the experiment when participants go from being completely blind to the cone to now completely blind to the ellipse. Why does this happen? How is their perception able to go through such a drastic change in image representation?

All studies targeted at perception have led to little conclusive results given the limited scope in studying the mind. Besides examining its structure and function, the attempts at objective analysis from language, the way in which humans communicate, will ultimately always become skewed. One will never know how information is altered once it is vocalized. With complex machines such as PET scans and fMRIs we are given more scope into the depths of the brain but there remains to be so many unanswered questions.

Perception is characterized by the ability to organize, interpret and create a meaningful picture of the world. This ideal image of perception is one that is constantly changing and influenced by many factors. Feelings such as attitudes and interests influence perception (Sree). Depending on the mood perhaps one can see an angry face in an inkblot as opposed to a happy one. It can also be influenced by time. Haste in looking at an object can prevent an individual from looking into the detail and prevent them from seeing another image. Another variable feature of perception is the object’s similarity to something already experienced, therefore making it easier to perceive
for one individual as opposed to someone without priming or prior experience. Although two individual can be presented with identical information, to them, the meaning and representation may be completely different.

An interesting idea regarding the topic of perception and representation could be a study that analyzes these differences in mental representation. In this speculative study, two individuals can be presented with the same physical information and tested if the same areas of the brain are stimulated by the image. They can then recount their perception of the image and the feelings that accompany it. These variations can then be mapped to see if there are certain differences, such as in memories of events or feelings that are consistently brought up by the majority of individuals, or if there is complete variability. Having complete variability would show the expansive knowledge of the brain and how each individual takes the stimuli and makes it their own. If there were to be trends in the objective stimuli, then perhaps we could assume that the brain functions and looks for certain things. For example, perhaps a loved one will be a recurring image, or food, because from an evolutionary perspective these are some advantageous qualities for survival.

Perception is a concept that really entices my interest and makes me think about the grand scheme of things and the variability among the human race. It makes me want to continue pursuing this field and attempt to identify the differences among individuals and possible explanations for variability. The Zililab has taught me that even with the simplest of tasks, there
will always be differences and that is something unique to the human race. We each have a creative and expansive mind of our own and that’s why we are superior organisms.
References


Effects of Aging in Memory and Learning Strategies

Uyen-Ly Nguyen (SSC)

Natural and healthy aging can affect memory in several ways. Research has shown that, in addition to decreases in memory capacity, aging also affects overall memory quality. Studies have shown that aging is associated with impaired recollective experience in older adults compared to younger adults in recognition tasks (Bugaiska et al., 2007). Furthermore, retrieval of autobiographic memory was shown to decline with age, resulting in fewer recalled episodic details when generating narratives of past events (Piolino et al., 2010). However, through “strategic control of attention and memory”, older adults can compensate for their reduced memory capacity by selectively focusing on higher valued information over less important material (Castel et al., 2011). Further support for this idea is provided by the findings of Spaniol, Schain, and Bowen (2013), which suggest that aging preserves the positive impact of reward anticipation on memory encoding.

Bugaiska et al. (2007) investigated the effects of aging on conscious recollection in recognition by utilizing the Remember/Know paradigm. Following a study phase for a series of words, subjects were instructed to indicate whether their recognition for words in a testing phase were due to remembering (R), knowing (K), or guessing (G). R responses indicate conscious recollection of an event, and thus correspond to episodic memory. On the other hand, K responses correspond to semantic memory, accompanied by an absence of episodic details. It was found that age had a significant effect on “remember” responses but not on “know”
responses. Older participants showed impaired conscious recollection compared to younger participants, suggesting a decline in episodic memory with increasing age (Bugaiska et al., 2007).

A proposed explanation for this decline is the speed mediation hypothesis, which suggests that decline in processing speed is responsible for cognitive declines in aging. Throughout the lifespan, processing speed of the central nervous system and consequently performance of mental operations slow down, requiring more time to perform. This effect could negatively impact both quantity and quality of memory, especially when encoding complex and sensory-rich experiences (Bugaiska et al., 2007). Another possible explanation for diminished episodic memory in aging is the executive-aging hypothesis, which suggests that executive function decline is the mediator for age-related differences in episodic memory. Aging is associated with a decline in executive functioning, which may play a role in interference control, conscious awareness, and memory of temporal order (Bugaiska et al., 2007). The impact of these mechanisms on episodic memory was investigated by additionally testing participants’ performances on the Wisconsin Card Sorting Test to evaluate executive functioning, and the Letter-Comparison Test to evaluate processing speed. They found that, though performance on both measures correlated positively with R responses, executive functioning appeared to be the more fundamental mediator of age-related variance in R responses.

The more prominent role of executive functioning in mediating conscious recollection is supported by the findings of Piolino et al. (2010), which investigate the effects of aging and the
influence of executive functioning and feature binding in working memory on generating autobiographical memory. Piolino et al. (2010) employed a novel verbal autobiographical fluency (VAF) task that had participants retrieve as many items as possible for four different levels of autobiographic memory with increasing specificity from “lifetime periods” to details of specific events. The two more general levels were more semantic in nature and the two more specific levels were considered to involve more episodic memory. They also utilized tests evaluating feature binding and the executive functions: updating, shifting, and inhibition.

For the VAF task, it was found that older adults recalled fewer items, especially for the more specific episodic levels of autobiographic memory, than younger adults. In addition, older participants indicated that their recall of these items was associated with a weaker sense of conscious recollection and field point of view (whether they recall the experience from a first person point of view versus that of an outside observer) than for younger adults, suggesting a relationship between old age and more limited episodic memory traces. Piolino et al.’s (2010) findings paralleled those of Bugaiska et al. (2007) in that age had a much more negative effect on episodic autobiographical memory than on memory for semantic information.

Similarly to Bugaiska et al.’s (2007) findings, Piolino et al. (2010) found that decline in executive processes (updating and inhibiting) played a large role in mediating age-related deficiencies in episodic autobiographical knowledge. These executive functions are likely involved in the complex retrieval process of more specific episodic autobiographical memories. In order to retrieve specific autobiographical memories, the updating function of the central
executive likely contributes by adjusting search criteria for retrieval and by updating current information in working memory with new external or internal information. Inhibition possibly plays a role in autobiographical memory by blocking irrelevant information from working memory (Piolino et al., 2010). Deficits in the associative binding function of the episodic buffer of working memory may also contribute to diminished recall of specific autobiographical memory for older adults by failing to effectively bind different aspects of an experience together to form a cohesive internal episode during retrieval (Piolino et al., 2010).

Executive function could also play an important role in facilitating the initiation and controlled execution of effective encoding of a lasting episodic memory. The frontal lobes, which are needed for executive control processes, are also involved in integrative encoding of multiple aspects of an experience. Frontal executive processes are needed to ensure that peripheral episodic details, such as spatial and temporal context, are adequately attended to integrated properly to form a full internal representation of an episode (Piolino et al., 2010). Thus, for older adults, aging-related executive decline is associated with a deficit in both encoding and retrieval of episodic information, in varying degrees. The extent to which executive function decline affects either process is unknown.

Another topic that is particularly salient in daily life concerns the effect of age on memory capacity and ability to selectively focus on valuable information. Castel et al. (2011) investigated these matters with different age groups across the average life span. They used a selectivity task in which a series of words were shown to participants with an accompanying
point value, and participants were later tasked with recalling words to earn as many points as possible. Often times, this requires strategic remembering of high value information at the expense of less valuable information. Cognitive factors that may influence selective ability are metacognitive monitoring of one’s own memory ability and capacity, inhibitory control of less relevant information, memory capacity, and attentional control (Castel et al., 2011). It was found that memory capacity and ability to strategically attend to high value information follow different patterns across the lifespan. Memory capacity peaks at young adulthood and declines throughout later aging. In contrast, the ability to be selective in value-based learning seems to surface during young adulthood and persist through healthy aging until older adulthood (Castel et al., 2011). This allows older adults to maximize their memory effectiveness by strategically focusing on fewer higher value items over low value items.

A possible reason for the dissociation between aging trends in memory capacity and strategic encoding is that the frontal lobe plays a large role in facilitating selective encoding through inhibition of lower valued information and goal maintenance. The frontal lobe is still developing throughout adolescence, which may explain why strategic value-based learning is not as efficient until young adulthood. In contrast, memory capacity may be driven by different neural systems such as the hippocampus (Castel et al., 2011). Another factor contributing to the different developmental pattern of memory capacity and ability to selectively encode higher valued information is that there are metacognitive strategic processes in play concerning selectivity in learning. These processes involve monitoring and evaluating past performance in
order to modify encoding strategies to improve memory effectiveness for subsequent memory tasks. Metacognitive skills seem to be more immature in childhood and adolescence, likely due to the fact that the frontal lobe is still developing at this time. Thus, during these early stages of life, individuals may be compromising memory effectiveness by utilizing more inefficient memory strategies.

Spaniol et al. (2013) confirmed the finding that long-term encoding based on reward anticipation is less affected by healthy aging in adulthood. It was found that encoding for individual words is enhanced by the accompaniment of cues indicating the possibility of higher future reward. This positive effect of reward anticipation was shown for both younger and older adults, and was thus preserved throughout aging. Interestingly, this effect of reward was not significant on an immediate memory test, but was significantly enhanced after a 24 hour delay, suggesting that the memory enhancement from the anticipation of reward extends to long term memory consolidation for both young adults and older adults (Spaniol et al., 2013).

In conclusion, research has shown that aging in adulthood has a negative effect on autobiographical episodic memory recall, whereas semantic or familiar memory is relatively preserved. This is possibly due the fact that aging is also accompanied by a decline in executive functioning, which plays a role in integrative encoding and in the complex retrieval of episodic information regarding past experiences. Aging also has an effect on memory capacity, which peaks in early adulthood and declines throughout the rest of the life span. However, a different trend is shown for “strategic control of attention and memory” (Castel et al., 2011). Ability to
selectively encode higher valued information matures during young adulthood and is relatively preserved into old age, where it experiences a slight decline. This ability allows older adults to compensate somewhat for their diminished memory capacity by strategically focusing on more important information at the expense of lower valued information, thereby “maximizing” their memory’s effectiveness.
References.


Memory Measurements

Uyen-Ly Nguyen

Current research by doctoral student Joseph Hennessee in the Knowlton Lab at UCLA examines the effects of value on learning and memory. During a study phase, participants are shown words on a computer screen, each with a corresponding numerical point value. Then, during a testing phase, participants must decide whether the subsequent words shown to them are “old” words that they saw previously, or “new” words that were not shown to them during the study phase. In order to examine the memory quality for each word, the experiment assesses three different measures of retention: recognition, remember-know, and source memory.

This experiment measures recognition simply by asking participants to indicate whether or not they think they have seen a word before in the study phase, and to rank their confidence in their answer. A positive answer expresses that there is some level of retention for the word, but holds no indication for what exactly is retained, or for the quality of retention. The sole yes/no measure of recognition treats memory as one-dimensional, even if a confidence rating is included. It ignores possible underlying mechanisms in memory, essentially treating memory as a single process. In an experiment studying the impacts of value on memory, it is important to further examine specific ways in which memory is impacted. For example, one would want to know what exactly can be recalled about the experience, or if recognition for a learned item invites only a feeling of familiarity. These questions could be especially pertinent in the real world, and in order to examine them, a simple indication of recognition is insufficient.
Recognition can be divided into two forms: remembering and knowing (Dudukovic & Knowlton, 2006). Thus, in order to add another level to the memory measurement in our experiment, participants are additionally prompted to indicate whether they recognize a word due to remembering (R), knowing (K), or guessing (G). “Remembering” corresponds to conscious recollection of an actual experience. Recognition would count as remembering if participants are able to recall specific details about when they first learned an item, such as contextual details or conscious thoughts that they had. “Knowing”, on the other hand, is recognition in the absence of such conscious recollections, and is more associated with feelings of familiarity without accompanying episodic details.

There are two main models explaining remember-know in recognition. One of these is the dual process model, which advocates the separate and distinctive nature of remembering and knowing as two forms of recognition. This model associates R responses with episodic memory, a type of explicit (declarative) memory for autobiographical events tied to a certain time and place in one’s life (Dudukovic & Knowlton, 2006). “Knowing” simply draws on semantic memory – memory for facts and knowledge about the world – with an absence of episodic recollection. Some evidence for remember-know as separate aspects of recognition come from dissociations of circumstances that affect the two processes. R responses are affected by different levels of processing in learning, whereas K responses are not. Divided attention leads to a decrease in R responses, but not in K responses. Shallow maintenance rehearsal increases K responses but not R responses (Knowlton, 1998; Dudukovic & Knowlton, 2006).
The dual nature of remember-know in recognition is further explained by the redundancy view: that the underlying processes in remembering are the same as those active in knowing, but with additional elements that supply recollection of episodic detail. In this view, “knowing” processes must be active in order for “remembering” to occur. This parallels the idea that episodic memory for an experience must also involve semantic memory for aspects within that experience (Knowlton, 1998). There is evidence from neuroimaging studies that suggest a redundant neurological dissociation between the two processes of remembering and knowing. Retrieval for items with K responses elicits activity in the hippocampal formation and the medial temporal lobe (Knowlton, 1998). For retrieval in items that are “remembered”, additional activity in the frontal lobes has been reported in numerous studies. It may be that the frontal lobes are active parts of the brain in retrieving additional episodic information from memory, which provide more vivid recollection of previous experiences, leading to R responses rather than K (Knowlton, 1998). It was also speculated that activity in the frontal lobes could be related to the “volitional retrieval” of information leading to conscious recollection associated with R responses. Lack of frontal lobe activity could contribute to the unconscious nature of information retrieval leading to K responses (Knowlton, 1998). R responses were also found to elicit activity in regions of the brain specific to the type of content being retrieved during the recognition task, whereas no such activity was found for items meriting K responses (Dudukovic & Knowlton, 2006).
Further evidence for the redundancy view comes from experiments examining the behavior of R-K responses over time. It was predicted that under the redundancy view, many responses that were “remembered” at the time of a first test would either convert to K responses one week later due to the loss of accompanying R processes, or would become misses that were neither remembered nor known, due to the loss of both R and K processes associated with the item to be retrieved (Knowlton, 1998). Indeed, it was found by Dudukovic and Knowlton (2006) that R responses decreased after one week, whereas the number of K responses remained fairly stable. The conversion of R responses to K responses somewhat compensated for the loss of K responses in the conversion of previous K responses to misses over one week, contributing to the stability of K responses over time. Additionally, the conversion of R responses to K responses over time complements the loss of contextual details for the study episode. These findings, in addition to those of the neuroimaging studies mentioned above, strongly indicate that remembering and knowing represent two different forms of explicit memory: likely episodic and semantic, respectively.

Another remember-know model is the single (or “decision”) process model, which favors the idea that remember and know are simply two thresholds within a “recognition strength dimension” (Donaldson et al., 1996). Essentially, this model claims that there is a continuous spectrum in strength of recognition. In this way, the single process model parallels that of signal detection theory. Items are marked as “recognized” when they exceed some criterion set for recognition. Items are “remembered” when they exceed an even stricter criterion, and are
“known” if they fall in between the lower “recognition” and higher “remember” criteria (Donaldson et al., 1996). This model suggests that recognition, remembering, and knowing are part of a single process and simply represent varying strengths of memory traces.

Since this model claims that the only difference between R and K responses is a quantitative difference in recognition strength, it suggests that there are still some memory traces of episodic details in K responses, contrary to what the dual process model suggests (Dudukovic & Knowlton, 2006). Indeed, some past studies have found that K responses were accompanied by smaller amounts of contextual details rather than none at all. Dudukovic and Knowlton (2006) speculate that this finding could have been due to the fact that those studies only tested recognition once, right after a study phase. In Dudukovic and Knowlton’s (2006) experiment however, they found that after one week, K responses reflected familiarity with no contextual details from the study phase, as expected by the dual process model. They explained that this discrepancy could be because at a short study-test interval, the overall recollection levels for items were high, leading to very few true K responses with no contextual details. The demand characteristics of the test may have led to participants simply using K responses for episodic memories that were not as strong as those with R responses, even though those items actually should have merited R responses. In such cases, it seems that the R-K responses agree with the single process model indicating memory strength.

This debate concerning what R-K assessments actually measure is one of the weaknesses of remember-know as a measure of memory. Overall, it should be noted that different subjects
can apply the R-K distinction differently, depending on the demand characteristics of the experiment and the way that the distinction is explained to them (Dudukovic & Knowlton, 2006). Past studies have found that participants’ understanding of the R-K distinction and how they apply it is easily affected by the instructions given to them in the study (Dudukovic & Knowlton, 2006). This is another weakness of the R-K measurement, due to the fact that this can lead to inconsistency in findings across different studies using different explanations for the R-K distinction; these R-K measurements might end up meaning different things in terms of underlying processes of recognition. The subjectivity of R-K measurements and the fact that this method relies on introversion is a substantial flaw. For this reason, it is necessary to also include a more objective measure of memory.

Source memory – memory for specific contextual details of an episode – is examined in our experiment by additionally prompting the participants during the testing phase to report the color and point value of the words they recognized from the study phase. Source memory adds another level to memory measurement by objectively quantifying the extent to which episodic details contribute to R-K judgments and to overall quality of memory (Dudukovic & Knowlton, 2006). Source memory assessments have enabled researchers to determine whether “know” responses are accompanied by contextual details, leading to the findings described above. It has been reported that “remember” responses in recognition correlate highly with correct recall of episodic details, whereas “know” responses were accompanied by a mean number of correct source judgments that was no greater than chance. This finding supports the theory that R
responses are an adequate measure of episodic memory, and that K responses measure non-episodic semantic memory (Knowlton, 1998; Dudukovic & Knowlton, 2006).

It makes sense that remember judgments and source memory are related to one another; if a participant is able to correctly recall a detail about the study experience, then their memory of the corresponding episode warrants a “remember” response, based on the definition of an R response (Donaldson et al., 1996). In fact, an electrophysiological study found that R responses have the same neural correlates (the minimum brain activity necessary for a specific experience) as correct source recall (Dudukovic & Knowlton, 2006). Some parts of the brain active in source memory include the hippocampus, frontal lobes (especially the medial prefrontal cortex), medial temporal lobe, and the striatum, according to Elward et al. (2015). Many of these areas are also part of the “core recollection network” active in remembering.

Despite the fact that source memory measurements contribute an objective extra level to recognition and remember-know assessments, this measure of memory quality has its weaknesses. In Dudukovic and Knowlton’s experiment (2006), 24% of R responses were unaccompanied by any successful source memory judgments. A possible explanation for this is that recollection of a specific experience can take the form of episodic memories for internally generated information as well as episodic memory for external context. Source memory assessments like those used in our study can only examine external details related to the way items are presented, such as color and location. They do not test whether participants can remember what their thoughts or feelings were at the time of item presentation, which is an
important factor for episodic memory in “remember” responses (Dudukovic & Knowlton, 2006). Additionally, source memory tests cannot possibly be fully comprehensive examinations of what the participants remember from the study episode. It is entirely plausible that participants can remember items that are not accounted for in the test phase. This may result in slightly under-representative scores for source memory.

Remember-know distinctions and source memory evaluations substantially build upon the standard recognition test and play a significant role in our study examining learning and value. Such findings can hold a particular salience in the real world. For example, in reviewing a victim’s testimony recounting their experience of a crime, it is critical to know how the victim’s high personal investment in the situation could possibly influence aspects of what they can and cannot remember about the experience. There are still some flaws in utilizing these measurements to determine the nature of memory for certain events, but hopefully further research will be able to lead to some way of resolving discrepancies in remember-know findings as well as procure some comprehensive way to assess memory for both internal and external episodic details.
References


The Scientific Study of Human Consciousness Using Psychophysics and Metacognition

Yasha Mouradi

The scientific study of human consciousness has historically been controversial due to the objective nature of scientific inquiry and the inherently subjective nature of conscious experience. At first glance it may seem impossible to gain insight into some aspects of subjective experience because there is no direct way to objectively measure them, however, we have gained much insight into various physical phenomena (e.g., planets or particles) by indirect observation and measurement. Assuming that consciousness arises in the physical brains of some humans, it should be possible at least in principle, to gain valuable insight into the nature of consciousness through careful objective measurements and scientific experimentation. Beyond having intrinsic value, knowledge about consciousness can prove very useful from a clinical point of view as there exist debilitating disorders of consciousness that could potentially be prevented, cured, or alleviated, once properly understood.

Which discipline is best suited for studying the human mind? No single discipline can provide a complete account of mental phenomena. The mind can be analyzed and studied from many perspectives and at various levels, and therefore, the best approach is interdisciplinary research. Cognitive science brings together the disciplines of psychology, linguistics, neuroscience, computer science, anthropology and philosophy, with the ultimate goal of establishing “a unified science that would discover the representational and computational capacities of the human mind and their structural and functional realization in the human brain” (Miller, 2003). Consciousness
is one aspect of the human mind and the aim of this work is to review some of the contemporary methods and tools that are employed to facilitate the scientific study of consciousness. Specifically, I will present an overview of modern psychophysics and explain how metacognition can shed some light on consciousness, which is viewed as an informational process that takes place within a larger informational system.

Psychophysics

Psychophysics is the study of the relationship between objectively measurable stimuli and subjective experience or behavior. Psychophysical experiments often involve weak stimuli, about which, participants make perceptual judgements after very brief presentation intervals. Stimulus energy is measured and adjusted during calibration of equipment to ensure precise systematic manipulation by the researcher.

Entire psychophysical experiments are performed using regular computers and calibrated display units. The display units are gamma-corrected so that the relationship between input voltage and output luminance is linear because it is the input voltage that is directly manipulated. The researcher is often proficient enough in a computer programming language (e.g., MATLAB and Psychtoolbox) and creates a program that presents stimuli at desired energy levels and time intervals, captures responses and response times from participants, and performs data analysis on the results.
Ideally, perceptual responses are the outcome of perceptual judgements that are made by an observer on the basis of available information, through an optimal process. Participants either detect the existence of a particular stimulus during various trials, or discriminate between similar stimuli. Using Detection Theory (DT), we can examine the effects of varying the amount of available information on the performance of participants and consequently infer their decision making strategies. The analysis takes into account the information (signal to noise) available during stimulus presentation, and based on the participant hit and false alarm rates, estimates the sensitivity of the participant in detection or discrimination of stimuli separate from the internal decision criterion that is used by the participant to make the perceptual judgement (Macmillan & Creelman, 2004). The researcher is able to determine the sensitivity of the perceptual system (affected only by the signal and noise distributions representing the stimuli) independently of subjective response bias, which affects the decision only. One can use the hit and false alarm rates to construct a curve referred to as Receiver Operating Characteristic (ROC), which signifies the detection or discrimination sensitivity of the participant. Each point on the ROC reflects a particular criterion and yields the corresponding hit and false alarm rates. Figure 1 shows an example of the ROC curve with multiple criteria points. Even though each point on the curve corresponds to different hit and false alarm rates, they all correspond to a constant discrimination sensitivity. Figure 2 shows ROC curves representing different sensitivity values (the diagonal represents chance performance). Each curve has its own set of criteria points.
Figure 1. Receiver Operating Characteristic curve with various criteria points.

Figure 2. Receiver Operating Characteristic curves for different sensitivity values.

The researcher often deploys an adaptive algorithm within the program, that manipulates the amount of information available to the perceptual system (participant) until performance stabilizes to some level (e.g., chance), which corresponds to a particular stimulus energy. This is referred to as the threshold stimulus energy that the participant requires as perceptual input in order to succeed in detection or discrimination some particular proportion of the time.

What is the role of consciousness in perceptual judgements? Consciousness is not necessarily required for above-chance performance in perceptual decisions involving certain stimuli. This is particularly interesting in cases of blindsight, where the participant achieves well-above-chance performance in a forced-choice perceptual judgment task and yet reports not having seen the stimuli.

Retrospective Metacognition

The relationship between metacognition and consciousness has been the subject of much debate and controversy in recent years (Lau & Rosenthal, 2011). blindsight patients demonstrate the link between metacognition and visual awareness most vividly. These patients achieve performance levels in perceptual tasks similar to healthy individuals, yet they are unable to reflect on their decisions to effectively rate their own confidence. Metacognitive judgements are
ideally based on the information that is predicted to be available (prospective metacognition) or based on information that is available in memory (retrospective metacognition) (Grimaldi, Lau, & Basso, 2015). Researchers often ask participants to rate their confidence in their perceptual judgements. In this approach the participant is making the perceptual judgement, about which some information is made available to the metacognitive system for making the confidence judgement. The first type of judgement is referred to as type 1 task and the metacognitive decision is referred to as a type 2 task. DT has been used to calculate a measure of metacognitive sensitivity based on data from confidence ratings of perceptual judgements (Maniscalco & Lau, 2012). Metacognition has been used to argue that blindsight cannot be induced in healthy subjects, since above-chance performance always accompanies metacognitive sensitivity that is also above chance, even though performance is higher than metacognitive sensitivity in many cases (Peters & Lau, 2015). It has also been argued that visual awareness and metacognition are independent processes and using metacognitive sensitivity to study consciousness will prove to be misleading (Jachs, Blanco, Grantham-Hill, & Soto, 2015).

Conclusion

Psychophysics offers methods and tools for the objective manipulation of stimuli and the observation of resulting changes in subjective experience. DT and computational modeling in psychophysical experiments can help researchers identify the mechanisms of perceptual decision making. Combining perceptual judgement tasks with confidence ratings allows for the investigation of metacognition and calculation of metacognitive sensitivity, which is perhaps
intimately linked with visual awareness and consciousness. Other tools and techniques, such as brain imaging and non-invasive brain stimulation, are available to researchers today, which only continue to be improved upon along with advances in technology. The scientific study of consciousness is a rapidly growing field and arguably, one of the most fascinating areas of contemporary scientific research.

References

Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. Neuroscience & Biobehavioral Reviews, 55, 88-97.


