<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
<th>Faculty Mentor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubber Hand Illusion: Ownership and Proprioceptive Drifts</td>
<td>3-15</td>
<td>Ladan Shams, Ph.D</td>
</tr>
<tr>
<td>Second Order Adaptation</td>
<td>16-23</td>
<td>Phil Kellman, Ph.D</td>
</tr>
<tr>
<td>FMRI and Reward Memory</td>
<td>24-29</td>
<td>Barbara Knowlton, Ph.D</td>
</tr>
<tr>
<td>Learning Patterns and Facts Study – An Analysis of Spacing Delays</td>
<td>30-42</td>
<td>Phil Kellman, Ph.D</td>
</tr>
<tr>
<td>Neocortical Reinstatement</td>
<td>43-49</td>
<td>Jesse A. Rissman, Ph.D</td>
</tr>
<tr>
<td>Object Recognition</td>
<td>50-54</td>
<td>Hongjing Lu, Ph.D</td>
</tr>
<tr>
<td>Interaction of Perceptual Learning Modules and Explicit Instructions in Statistics</td>
<td>55-82</td>
<td>Phil Kellman, Ph.D</td>
</tr>
<tr>
<td>Improving Computer-Based Learning: How the Introduction of Game-Like Elements May Improve Learning</td>
<td>83-90</td>
<td>Elizabeth Ligon Bjork, Ph.D</td>
</tr>
<tr>
<td>Thyroid Hormones Dynamics Stimulation in MATLAB</td>
<td>91-98</td>
<td>Joseph Distefano, Ph.D</td>
</tr>
<tr>
<td>Representing Brain Activities During Processes of Reasoning</td>
<td>99-106</td>
<td>Keith Holyoak, Ph.D</td>
</tr>
</tbody>
</table>
The Effect of Social Exclusion on Social and Cognitive Working Memory Tasks........107-113
Caitlin M. Black, Faculty Mentor: Mathew Lieberman, Ph.D

Predicting Point Light Action during Binocular Rivalry.............................................114-121
Ye Eun Chun, Faculty Mentor: Hongjing Lu, Ph.D
Rubber Hand Illusion: Ownership and Proprioceptive Drifts

Albert Chung

Introduction

The Rubber Hand Illusion was first formally demonstrated in the work by Botvinick and Cohen in 1998. It was defined as a “the feeling that a rubber hand belongs to one’s body brought by stroking a visible rubber hand synchronously to the participant’s own occluded hand.” (Rohde, Luca, and Ernst 2011) In this study, it was found that the Rubber Hand Illusion induces significant displacement in a ‘reach judgment task’ which is a task analogous to measuring proprioception drift. Also, a questionnaire task which indirectly collected subject responses for the effects of ownership provided further support that there were significant visual and tactile sensory integration induced by the Rubber Hand Illusion. Such result was reproduced in many other studies such as Armel (2003), Tsakiris (2005), and Constantini (2007). In a recent study by Rhode, Luca, and Ernst (2011), it was found that the mechanism for feeling of ownership and proprioceptive drift may be separate mechanisms.

In the current study, with the aims to investigate the difference in mechanisms of ownership and proprioception, we explored different possibilities of visual-tactile integration and visual-proprioception integration. The experimental methods used to
collect data were similar to the methods that were introduced in the study by Rohde, Luca, and Ernst (2011) with addition of slight modifications. These modifications were intended to produce a better control over the sensory stimulation, and reproduce the cleanest possible Rubber Hand Illusion. Furthermore, in an attempt to model the sensory integration, a Bayesian causal inference model presented by Kording et al. (2007) was adopted. The causal inference model is a widely accepted model for multisensory integration. The model mathematically describes tendency to establish sensory integration decision on some probabilistic function (p-common), and when sensory information from multiple modalities are proximate enough to be integrated, the outcome estimate is calculated as some combination of the two sensory likelihood. The current study probed for this change in the integrated estimate, based on an assumption that our proprioceptive localization likelihood is driven by our tendency to expect that both our arms exist at mirror location at rest.

*Experiment 1*

In experiment 1, the conditions of sensory stimulation were manipulated. By varying visual and tactile stimuli, we attempted to induce different level of sensory integrations.
Method

Participants

A total of 92 subjects were recruited in this experiment. These subjects were students volunteers from University of California, Los Angeles and were recruited by SONA-system. These participants were given 1 credit on SONA for each participation session as compensation. 12 of these were excluded as they deviated from the standard population in their pre-test proprioception or did not comply with given instructions.

Design

In this experiment we were interested in two main dependent variables: ownership and proprioception drift. These data were collected across four different conditions, each controlling for different combinations of stimuli. “Synchronous” condition received synchronous visual and tactile stimulation, “Asynchronous” condition received a delayed tactile stimulation after the visual stimulation, “Hand” condition received no tactile stimulation, and “No Hand” condition did not receive visual or tactile stimulation. This independent variable was measured as between subjects. Both ownership and proprioception was measured before and after the stimulus.

Material
A spray painted goggles, a black cloth, two brushes, a rubber hand, medical tape, cardboard and a handcrafted wooden box were used as prop for inducing the Rubber Hand Illusion. The handcrafted wooden box was composed of three wood boards, two of them parallel to each other and one of them connecting the two at the bottom. Matlab installed Macintosh desktop was used for running the Matlab coded program which was designed to collect and save data.

**Procedure**

Each participant was asked to read and sign the consent form before any procedure was done. They received a copy of the signed consent form and participant bill of rights. Also, each participant was asked to check if they satisfied any exclusion criteria depending on their health state.

Participants were seated in front of a Mac desktop monitor in a specific position marked with a tape and were asked to be blinded by putting on the spray painted goggles. The Matlab based program was executed at this point and was used during the remainder of the entire procedure. The goal of this phase was to blind participants from their own hand for 240 seconds until pre-stimulus proprioception could be measured. The timer counting down 240 seconds started and was displayed on the screen from the moment the
participants were blindfolded. This displayed number was designed to jiggle at a random moment to be used as a distractor task throughout the experiment. While the timer counted down, the experimenter placed the handcrafted box in front of the participant and placed his or her left arm in a predetermined position. The left index finger was aligned in a particular spot for all subjects who were recorded in the Matlab based program. The index finger was fixed to stay there with use of a medical tape. The participant’s right arm was positioned outside the box. Then a cardboard was placed on top of the handcrafted box, and was covered with the black cloth. The spray painted goggle was removed, and the participants were asked to look at the counting down seconds and mentally keep track of how many times the number jiggled for the remainder of 240 seconds.

After the 240 seconds, the subjects were asked to perform 4 practice trials and 40 proprioception measurement task trials. In this task, the participants were asked to move the mouse cursor which was fixed across Y axis to the location that best fit the description “wherever the left index finger feel like it is” and clicked. Matlab based program measured and recorded this location and automatically calculated the distance between the actual index finger location and the response location. At the end of 40 trials, the participants were blinded again by the spray painted goggles. Then, the experimenter placed a cardboard wall between the participant’s real left arm, and where the visual stimulus
(rubber hand) would be placed. Depending on the experimental condition that the participants were assigned, either the rubber hand was placed if they were in “Synchronous”, “Asynchronous”, or “Hand” conditions, and nothing was placed if they were in “No hand” condition. Participant’s right arm was placed in the box symmetric to the rubber hand as it was reported that it amplifies the Rubber Hand Illusion.

The spray painted goggle was removed after this point, and the participants were asked to rate on the scale of -3 to 3 on how much the rubber hand felt like it was their own hand based only on visual stimulation if they were not assigned in the “No hand” condition. This was recorded as “pre-test ownership” score. Then, the participants received visual and tactile stimulation for 240 seconds. If they were assigned in the “Synchronous” condition, subjects observed brushes on the rubber hand while they received the same tactile stimulus on the same location of their real hand. In the “Asynchronous” condition, the tactile stimulus was spatially congruent, but was temporally delayed, creating an asynchronous stimulus. In the “Hand” condition, the subjects only received visual stimulation, and in the “No Hand” condition, they were given no task to do, but were asked to stare at a fixed point where the rubber hand would have been placed.
At the end of the illusion phase, the participants were blindfolded with the goggles, and the same setup as previous proprioception measurement was restored. The participants were asked to do the proprioception measurement task, this time without the practice trials. Then, a short questionnaire asking for how much the rubber hand felt like it was their hand after the illusion phase was asked this answer was recorded as “post-test ownership” score.

Results

The ownership scores were only collected in “Synchronous”, “Asynchronous”, and “Hand” conditions as “No hand” condition did not receive any visual stimulation by the rubber hand. The ownership score was calculated by measure the difference between the subject’s response before and after the Rubber Hand Illusion induction. We conducted the analysis of variance (ANOVA) for this change. This test reported a statistically significant interaction as predicted in our hypothesis \[F(2; 60) = 5.06; p < 0.01\]. (Figure 1)

The change in proprioceptive measure, delta bias, was calculated by measuring the difference of proprioceptive drift between pretest and posttest. We also tested for ANOVA and found statistically significant result \[F(3; 80) = 3.52; p < 0.05\]. (Figure 2)
Further post-hoc t-tests revealed that proprioception drifts were only significant across “Sync” group and “Async” group ($t_{40} = 2.50; p < 0.05$). The t-test across “Sync” group and “No stroking” group failed to reach significance, however reached a trending ($t_{40} = 1.81; p = 0.078$). Lastly an additional t-test between “no hand” and “sync” group to verify the effectiveness of the “no hand” control group which resulted ($t_{40} = 2.94; p < 0.01$).

**Experiment 2**

An additional condition similar to the “sync” group in experiment 1 was made. The only difference in this group was that the subject’s real arm was placed outside the wooden box during stimulation phase, altering the mirrored hand location assumption. This manipulation expected a weakened proprioceptive drift.

**Participants**

20 additional students in UCLA psychology department were recruited by SONA systems. These participants were also given credit for their participation.

**Materials and Procedure**

The exact same materials and procedure used in experiment 1 was adopted. However, during the stimulation phase, the experimenter removed the participant’s arm
out of the box, occluding their arm from the visual field. All the data collection phase was exactly same as the “sync” condition in experiment 1.

Results

Unfortunately, the collected proprioceptive drift failed to reach significance when compared to the “sync” group from experiment 1. However, the means were trending towards weakened proprioceptive drift. This suggests that there could be an effect due to the manipulation in this experiment that just happened to fail to reach significance.

General Discussion and Conclusion

While the results of the current study replicate most of the findings in the previous studies, it also showed a novel finding that the feeling of ownership is achieved mostly by visual stimulus, though it is either retained or lost after tactile stimulation depending on how synchronous the visual and tactile integrations are.

The drifts in proprioception were only significantly different between “Sync” group and “Async” and between “Sync” and “No hand” while not significantly different between “Hand” group and “Async” group and between “Hand” group and “No Hand” group. However, because the “Hand” group was also not significantly different from “Sync” group,
we can presume that the “Hand” group behaves as some midpoint between the experimental (“Sync”) and two control groups (“Async” and “No Hand”). In conclusion, this suggests that even though the feeling of ownership could be solely achieved by visual stimulus, the proprioceptive drift requires visual-tactile integration. Although the experiment 2 was failed to reach significance, the trending means were suggesting that the proprioceptive drift is caused by integration of our expected hand location (presented by vision) and tactile stimulation.

**Future Research**

To justify the proposed Bayesian causal inference model, a good approach would be constructing more simulation results. When a number of conditions that generate unusual integration behavior are found by this constructed model, this specific condition can be adopted as an experimental condition. Once the simulation results are replicated in a laboratory setting, our proposed model will gain strength.
Figure 1: As predicted in our hypothesis, the change in ownership is most robust in the synchronous condition. The large negative change in asynchronous group reflects that subjects originally had a large ownership score without inducing any tactile stimulus.
Figure 2: Proprioceptive drift in four different conditions were measured with n=20. As predicted, synchronous group showed most proprioceptive drift due to visual-tactile integration while the other conditions do not show as much of a drift. It does show that there is an effect of just presenting visual stimuli as significant difference in “async” and “no stroke” from “no-hand group” show.
References


Previous versions of adaptive learning modules are generally based on expanding presentation intervals determined by estimated parameters, like the Atkinson model (1972). Mettler, Massey and Kellman (2011) developed an adaptive learning system called Adaptive Response Time Based Sequencing (ARTS), which uses accuracy and response time as inputs into a sequencing algorithm for a learning program. Accuracy and response time are used as measures of fluency, as detailed in Benjamin & Bjork (1996) and Pyc and Rawson (2009), who show that response time correlates with learning strength. Bjork & Bjork (1992) describes how difficult retrieval improves memory through the idea of “desirable difficulties.” One of the goals of adaptive learning is to predict the longest interval between item presentations before the learner forgets. In a sense, the idea is to stretch the retention interval, but retain sufficient information to further solidify the learning experience.

Mettler used a priority score system to optimally space the items according to reaction time. All items were assigned scores indicating relative importance of future presentation.
These scores are a function of response time ad learning strength, which is calculated from performance data. The next trial presents the highest priority item at that time. The Priority Score for item i ($P_i$) is given by:

$$P_i = a(N_i - D) [b(1 - \alpha_i) \log (RT_i/r) + \alpha_i W]$$

where:

- $N_i$ = number of trials since item i was presented
- $D$ = enforced delay constant (trials)
- $a, b, r$ = weighting constants
- $\alpha_i$ = 0, if I was last answered correctly
- $= 1$, if learning item was last answered incorrectly
- $W$ = priority increment for an error
- $RT_i$ = response time on most recent presentation of item i

Initial priority scores are given to all items and then the scores are updated after each trial.

This system delivers rapid reappearance of missed items due to high priority scores and retention intervals expand as an inverse function of response time, in the cases of accurate responses. Faster responses produce longer intervals between appearances.

Methods
Mettler compared the adaptive reaction time-based sequencing model (ARTS) with the Atkinson model (1972) for adaptive learning. The subjects were 50 undergraduate psychology students at the University of California, Los Angeles. The subjects were tested on African geography. They were shown a map of Africa with 24 countries on it and were asked to select the name of a country as it was outlined from a list on the side of the map. The subjects were split into three conditions. The first condition received a fixed equal interval schedule for item presentation (5-5-5), the second received a fixed expanding interval schedule (1-5-9) and the third group received an adaptive sequencing schedule based on response time. Each item was presented 4 times in all conditions. There were 24 countries being tested and 12 filler items. After each response during the learning phase, the student would receive feedback telling them whether they got the answer right or not. If the subject gave an incorrect answer, the correct answer would be highlighted in the list. After the learning phase, they immediately completed a post-test and also came back one week later to complete a delayed post-test. Figure 1 and Figure 2 show screen shots of what the students saw during the experiment. Figure 1 shows the map of Africa and the list of countries on the side. Figure 2 shows the screen shot of what happens when a subject selects the wrong answer.
Results

Figure 3 shows the results of the subjects in each condition on the pre-test, during the training/learning phase, on the immediate post-test and on the delayed first test. As is evidenced by the figure, while the adaptive learning scores are marginally higher, there is not much difference between the scores during the learning phase and on the immediate post-test. However, adaptive learning scores were significantly higher than the expanding and fixed groups on the delayed post-test. The scores for the fixed equal condition are the lowest for learning, immediate post-test and delayed post-test. There is little variation in the pre-test, which was expected due to the randomization of the subjects.
Figure 3 – Average scores across all participants in each condition

Also, there appear to be significant differences at presentation, across items and for participants, but there does not appear to be an interaction of user and item. Figure 4 shows the average delay by item and figure 5 shows the average delay by participant.
Discussion

These results show that adaptively generated delays lead to subjects maintaining retention of material for longer periods of time relative to subjects in the fixed expanding and fixed equal presentation conditions. Also, in the adaptive condition, the proportion of incorrect answers after the 3rd presentation decreases significantly in the training period. Karpicke & Roediger (2007) claim that the first delay, between presentation 1 and 2, was critical, yet these results indicate that the delays between presentation 2 and presentation 3 are even more critical. It also appears that limiting the number of presentations of the items may decrease the advantage of adaptive presentations at a delay.

Future Research

The most efficient algorithmic model for determining presentation intervals for one subject may not be the same model that is the most efficient for another subject. The idea of 2nd order adaptive learning is that the priority algorithm should be adjusted from learner to learner to create the most efficient presentation of items. For some learners, the delays become quite high after correct answers. Remaining questions include how and when this adjustment should be made within the learning process. One possible scenario is to measure the number of incorrect responses after the 2nd delay, and if the number of incorrect answers drops below a certain predetermined number, it can be reasoned that
the delays are too long and weighting constants should be adjusted. Measuring the percentage of incorrect answers to determine when the delays are too long is also an option. It might also be useful to try implementing delay intervals that were adapted for one subject onto a new subject and then comparing this to both fixed equal and fixed expanding conditions. This type of experiment would give more insight onto whether a certain adapted schedule can be applied to multiple learners.
References


This quarter, we have been running an fMRI experiment that examines the differences in brain activity between younger and older adults during value-based memory tasks. One task requires subjects to learn words that are assigned point values, and will be later asked to recall as many words as possible in addition to achieving as high a score as possible. Another task asks subjects to quickly press a button when they see a white square on the screen. Preceding this white square, a monetary value appears on the screen that tells subjects an amount they can win if they press the button on time (either no money, ten cents, or a dollar and ten cents.) This experimental design has drawn from a previous study by Adcock, Thangavel, Whitfield-Gabrieli, Knutson, and Gabrieli (2006), that used a similar procedure and discovered that areas implicated in reward learning became active while subjects completed these tasks.

In the Adcock et al (2006) paradigm, the task wherein subjects can win money by pressing a button for a white square is known as a monetary incentive delay task, or MID task. In our own study, we have used this same task, but with different reward amounts. The other task in the Adcock study, called the monetary incentive coding task, or MIE task, required subjects to view pictures that were also preceded by a monetary cue, of either
$0.10 or $5.00. One day later, subjects would perform a recognition task, which would reveal if subjects were better able to remember the pictures associated with the high values or the low values. The authors explained that the two different tasks were meant to tease apart motivated learning and learning by itself.

Results showed that subject had better memory for images associated with high values than those associated with low values. The fMRI data showed that the high value images that were later remembered activated the ventral tegmental area, the nucleus accumbens, and the hippocampus. Higher hippocampal and ventral tegmental activation was also correlated with better memory. Furthermore, activity in these areas did not correlate with words that were later forgotten. Activity in the ventral tegmental and the nucleus accumbens occurred before encoding the stimuli, while medial temporal lobe activation was occurred during encoding of the stimuli. Subjects who showed more brain activity in these regions before actual encoding also had better memory overall.

The authors interpreted these results to imply that activity in these regions prior to encoding was associated with later memory for events. Previous studies in rats found that ventral tegmental dopamine projections to the medial temporal lobe are important in order to form memories. Also, projections from the medial temporal lobe to the nucleus accumbens and ventral tegmental can have influences over their activity. In conjunction
with previous findings, the current study seems to suggest that the ventral tegmental, medial temporal lobe, and nucleus accumbens work together during motivated memory formation, with the help of the neurotransmitter, dopamine.

The monetary incentive delay task used in this study, or MID task, that we have also used in our own study, was first described in a study by Knutson, Westdorp, Kaiser, and Hommer (1999). They too collected fMRI data to find out which areas were related to reward memory.

In their study, Knutson et al (1999) asked their subjects to complete three tasks. One task was a control task, where subjects would not win any rewards. Another, a reward task, where subjects could potentially earn money. The last task was a punishment task, where the subjects could lose money. As described previously, subjects would see a white square, and afterwards would have to respond as quickly as possible. Before the white square, however, they would see a colored square. Orange squares would indicate that they could win money, red and blue squares would indicate that they could neither win nor lose money, while yellow squares indicated they could lose money. Only after they had responded to the white square, which followed the colored square, would the subjects know if they won or lost any money, as well as a total sum of how much they had won so
far. This MID task was adapted from similar tasks that had been used previously on monkeys, and had found that ventral tegmental neurons would fire while doing so.

The authors hypothesized that the reward task would lead to activity in the striatum as well as the mesial forebrain, and both of these structures make use of dopamine. Their results showed that many structures were active during both rewards and punishment, such as the mesial prefrontal cortex, the insula, caudate and putamen. During the punishment task, activation was also visible in the anterior cingulate and the thalamus.

They concluded that their findings were in line with previous animal research that had found that structures in the midbrain containing dopamine neurons contribute to reward-related processes. While it is difficult to conclude whether or not dopamine played a role during these tasks using only fMRI data, the authors believed that dopamine is important for the monetary incentive task as well.

Overall, the two studies discussed both used the MID task, where subjects can win or lose money depending on how they respond to a cue. In the study we are running through the Knowlton lab, we have also used a version of this MID task, and will see whether there are differences in activation between younger and older adults.
This quarter, I learned a great deal while working as a research assistant for the Knowlton lab. I learned about how to set up an MRI scan, safety protocols for conducting scans, and I also learned more about running behavioral studies. I enjoyed conducting the MRI scans the most out of all the work we did. I feel that it prepared me to run MRI studies of my own someday, and I value this in particular as I aspire to be a cognitive neuroscience researcher. Running scans in conjunction with my class in functional neuroimaging, as I was able to apply the concepts and ideas that I learned about to the scanning process. I would have also liked to learn about pre-processing and analyzing the data we collected, but fortunately I have some previous experience with this and have also had the chance to do this in my neuroimaging class.

After working with Michael this past school year, overall, I feel that this experience was very valuable and eye-opening. Working in the Bjork lab introduced me to running behavioral studies and working hands-on with the participants. Attending the Bjork lab meetings and the journal clubs were also enjoyable. The Bjork lab meetings were a great way to learn about designing experiments, and how time consuming and effortful it truly is, which I was surprisingly unaware of previously. The journal clubs were valuable in that they taught me to think critically, and how to read research articles in the most effective way possible.
Additionally, working in the Knowlton lab was especially enjoyable, as the research was most pertinent to my interest in cognitive neuroscience and brain imaging, although I discovered from the Bjork lab that memory is also a fascinating subject. While the Knowlton lab meetings were smaller and had fewer undergraduates present, they were also very informative and interesting, and I enjoyed them very much.

Altogether, I am very grateful that I was able to be a part of these labs this year. My only regret is that I am unable to stay here longer and keep learning about conducting research!
Learning Patterns and Facts Study – An Analysis of Spacing Delays

Stephanie Dunn

Introduction

The present study in the Kellman Lab is testing learning algorithms within a data set of African countries to assess learning effects on retention. Supportive research derives from the Mettler, Massey, and Kellman (2011) study that employed a learning algorithm called Adaptive Reaction Time-Based Sequencing (ARTS) in teaching geography and mathematics in an educational environment. The algorithm is adaptive because it is weighted by the user’s reaction time and learning strength (indicated by performance data) to predict optimal spacing between item presentation. The adaptive item presentation, via delays, allows for retention intervals to expand as an inverse function of response time for accurate responses, where faster responses produce longer intervals. In Mettler et al., this learning algorithm was compared to the old pattern of Atkinson (1972) that used a Markov model. The comparative learning patterns showed higher accuracy recall in the ARTS learning pattern condition than in the Atkinson condition.

The Mettler et. al (2001) study used insights found in a Storm, Bjork, and Storm (2010) study that provided support for expanding intervals produced by ARTS. Researchers presented stimuli to participants to-be-learned on equal and expanding
intervals. The results showed much higher accuracy of recall for the expanding condition.

The implication of this study was that the expanding schedule only caused enhanced recall when the task between successive retrievals was interfered with memory for the passage. The results suggest that the extent to which learners will benefit from expanding retrieval practice depends on the degree to which the to-be-learned information is likely to be forgotten. Storm et. al reasoned that when tests are given immediately following presentation, learners are able to access information from memory in a way that affords little benefit. Rather, when tests are delayed, the to-be-tested information has become less accessible and learners are inclined to engage in the type of processing that promotes learning and long-term retention. Storm et. al combatted the tendency of participants to unsuccessfully retrieve the item from memory with increased delays with the expanding schedule.

The present study observes two different training schedules: Fixed and adaptive. The adaptive condition is based on an algorithm similar to Everett et. al (2011) and changes as a function of a user’s response time. The other condition has two fixed schedules: Equal and expanding at a 1-5-9 pace. The data analyzed has revealed significant retention for the adaptively generated spacing schedule, compared to the fixed schedules.
These encouraging results have guided the analysis into a fine-grained approach to spacing delays, response time, and accuracy.

Figure 1 shows the average accuracy for each segment of the experiment. The first (zero) segment is before the learning phase and the second is the post-learning phase test. The third is the delayed post-test. Unsurprisingly, accuracy initially boosts after the material is learned, but drops off considerably at the delayed post-test. This is due to the participants not being exposed to the information for some time, ensuring the results demonstrate long-term retention. The details to pay attention to in Figure 1 include the condition that has the highest drop off (fixed equal), and the condition has the highest overall accuracy (adaptive). These results mean that the fixed equal condition has the highest forgetting rate and the adaptive has the lowest forgetting rate and highest overall accuracy.

The data analysis for the present study is guided by the retention benefits observed in taking a test is better than restudying. This finding is can be found in the Roediger and Karpicke (2006) study that observed expanding versus equally spaced testing after the practice of vocabulary word pairs. The researchers found retention benefits for gaps between practice and testing. More specifically, the delay in the first test improved long-
term retention regardless of how the repeated tests were spaced. This implication is relevant for our discussion later.

The analysis is also supported by the fact that temporal space or delay between presentations benefits memory. A study by Cepeda, Vul, Rohrer, Wixted, and Pashler (2008) examined the joint effects of gap and retention interval. The researchers found, for each retention interval, final performance always rose initially with increasing gap and then fell as gap was further increased. Conversely, Bjork and Simon (2001) studied retention effects in massed versus spaced practice with keyboard pattern memorization. The massed practice was more popular amongst the participants, but lead to poorer results after retention. Thus, massing schedules lead to declined retentions, where delays between presentations lead to non-monotonic increases in retention.

Data Analysis

The data show that the adaptive condition generates spacing delays of varying lengths according to the user’s response time and accuracy. The distribution of spacing delays, shown in Figure 2, illustrates the delays after success, on the bottom, and spacing delays (enforced delays of 2) after failure, on the top. The accuracy after presentation is indicated by the highlighted portion of the bars, blue for success, red for failure. The figure
displays the increase in success after delay after the second and third presentations, and the total failure within the responses decreasing as presentations increase. This figure also shows that the adaptive algorithm produces much longer delays than the fixed scheduling, with the average delay being upwards of 10. This is relevant because it appears that the longer the delay, the better the accuracy.

The results also revealed that the highest overall accuracy is obtained after the fourth presentation. Accuracies for each presentation as a function of average delay are shown in Figure 3. A couple things are worth pointing out in this graph: the first being that accuracy drops off during after the third presentation (as compared to accuracy after the second presentation), and second, there appears to be a positive correlation between delay and accuracy. Accuracy is probably higher after the second presentation than after the third presentation because most participants will have an incorrect response after the first presentation. After this incorrect response, the participant receives an enforced, short delay, and thereby sees the item shortly after. After this success, they receive a longer delay and accuracy decreases again. Though this result is not ideal, the important point here is that accuracy is highest after the fourth presentation.
Further data analysis has lead to the observation of features in correlation with post-test accuracy. Table 1 displays the correlation that several variables (response time, accuracy, delay, etc.) have with accuracy, and whether this correlation is significant or not. This table is for the immediate post-test scores and shows that presentation 2 RT, length of delay 3, presentation 2 accuracy, and presentation 3 accuracy are significantly correlated with immediate post-test accuracy. Similarly, Table 2 displays the same information but for the *delayed* post-test after a week of not studying. The linear model illustrated in Table 2 reveals that the length of delay 2, presentation 3 RT, presentation 1 accuracy, and presentation 3 accuracy are significantly correlated to delayed post-test accuracy.

*Conclusions*

From the analysis, it can be concluded that adaptively generated delays lead to longer retention of material relative to fixed expanding and fixed equal presentations. Additionally, as revealed from the linear model, delays between presentation 2 and 3 appear critical. This is notable because it is contrary to Karpicke and Roediger (2007) who found the first spacing delay to be the most important gap. Also important, presentation 3 generates longer delays and has lower and more variable accuracies than presentation 4.
The results reveal that, in general, larger adaptively generated delays lead to better post-test performance.

**Future Research**

An important thing to keep in mind with this study is the effectiveness of the adaptive condition. Is it producing the right delays that are specific to that participant? The results appear to confirm this notion. However, a way to ensure the specificity of the adaptive delays, we could observe the accuracy results from participants who receive other participants’ adaptive delays. The hypothesis is that this would produce significantly poorer retention than those with their own adaptive delays. If the results revealed the converse of this hypothesis, this could have implications for adaptively produced delays; are only a few schedules needed that work for many learners, or should every participant receive their own individualized sequence?
Figure 1. The average accuracy for each condition across the different segments of the experiment.
Figure 2. The distribution of spacing delays after success (bottom), and spacing delays (enforced delays) after failure (top).
Figure 3. Average delay and following accuracy at each presentation by item.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Slope</th>
<th>Error</th>
<th>t</th>
<th>p</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>4.78E-02</td>
<td>4.37E-02</td>
<td>1.093</td>
<td>0.274521</td>
<td></td>
</tr>
<tr>
<td>delay1</td>
<td>2.95E-03</td>
<td>4.21E-03</td>
<td>0.701</td>
<td>0.483422</td>
<td></td>
</tr>
<tr>
<td>delay2</td>
<td>3.27E-05</td>
<td>1.64E-03</td>
<td>0.02</td>
<td>0.984032</td>
<td></td>
</tr>
<tr>
<td>delay3</td>
<td>7.60E-03</td>
<td>2.56E-03</td>
<td>2.972</td>
<td>0.003 **</td>
<td></td>
</tr>
<tr>
<td><code>pres1-acc</code></td>
<td>6.73E-02</td>
<td>4.46E-02</td>
<td>1.51</td>
<td>0.131278</td>
<td></td>
</tr>
<tr>
<td><code>pres2-acc</code></td>
<td>1.03E-01</td>
<td>2.93E-02</td>
<td>3.511</td>
<td>0.000459 ***</td>
<td></td>
</tr>
<tr>
<td><code>pres3-acc</code></td>
<td>2.07E-01</td>
<td>2.78E-02</td>
<td>7.456</td>
<td>1.41E-13 ***</td>
<td></td>
</tr>
<tr>
<td><code>pres4-acc</code></td>
<td>2.27E-01</td>
<td>3.05E-02</td>
<td>7.455</td>
<td>1.42E-13 ***</td>
<td></td>
</tr>
<tr>
<td>preAc</td>
<td>-3.35E-02</td>
<td>4.46E-02</td>
<td>-0.752</td>
<td>0.452341</td>
<td></td>
</tr>
<tr>
<td>delayedAcc</td>
<td>2.88E-01</td>
<td>2.27E-02</td>
<td>12.689</td>
<td>&lt; 2e-16 ***</td>
<td></td>
</tr>
<tr>
<td><code>0.RT.Mean</code></td>
<td>2.44E-07</td>
<td>1.44E-06</td>
<td>0.169</td>
<td>0.865438</td>
<td></td>
</tr>
<tr>
<td><code>2.RT.Mean</code></td>
<td>-1.36E-05</td>
<td>1.30E-06</td>
<td>-10.454</td>
<td>&lt; 2e-16 ***</td>
<td></td>
</tr>
<tr>
<td><code>3.RT.Mean</code></td>
<td>1.52E-06</td>
<td>1.42E-06</td>
<td>1.071</td>
<td>0.284398</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. The linear model of correlations for immediate post-test accuracy.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Slope</th>
<th>error</th>
<th>t</th>
<th>p</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-8.86E-02</td>
<td>4.44E-02</td>
<td>-1.993</td>
<td>0.04637</td>
<td>*</td>
</tr>
<tr>
<td>delay1</td>
<td>-5.41E-03</td>
<td>4.28E-03</td>
<td>-1.263</td>
<td>0.20669</td>
<td></td>
</tr>
<tr>
<td>delay2</td>
<td>7.13E-03</td>
<td>1.65E-03</td>
<td>4.312</td>
<td>1.71E-05</td>
<td>***</td>
</tr>
<tr>
<td>delay3</td>
<td>4.75E-03</td>
<td>2.60E-03</td>
<td>1.825</td>
<td>0.06821</td>
<td>.</td>
</tr>
<tr>
<td><code>pres1-acc</code></td>
<td>1.88E-01</td>
<td>4.51E-02</td>
<td>4.164</td>
<td>3.28E-05</td>
<td>***</td>
</tr>
<tr>
<td><code>pres2-acc</code></td>
<td>4.60E-02</td>
<td>2.99E-02</td>
<td>1.539</td>
<td>0.12405</td>
<td></td>
</tr>
<tr>
<td><code>pres3-acc</code></td>
<td>1.16E-01</td>
<td>2.86E-02</td>
<td>4.063</td>
<td>5.07E-05</td>
<td>***</td>
</tr>
<tr>
<td><code>pres4-acc</code></td>
<td>8.76E-02</td>
<td>3.15E-02</td>
<td>2.785</td>
<td>0.00541</td>
<td>**</td>
</tr>
<tr>
<td>preAc</td>
<td>5.17E-02</td>
<td>4.53E-02</td>
<td>1.14</td>
<td>0.25459</td>
<td></td>
</tr>
<tr>
<td>postAcc</td>
<td>2.98E-01</td>
<td>2.35E-02</td>
<td>12.689</td>
<td>&lt; 2e-16</td>
<td>***</td>
</tr>
<tr>
<td><code>0.RT.Mean</code></td>
<td>4.74E-06</td>
<td>1.46E-06</td>
<td>3.255</td>
<td>0.00116</td>
<td>**</td>
</tr>
<tr>
<td><code>2.RT.Mean</code></td>
<td>4.93E-07</td>
<td>1.36E-06</td>
<td>0.362</td>
<td>0.71776</td>
<td></td>
</tr>
<tr>
<td><code>3.RT.Mean</code></td>
<td>-8.84E-06</td>
<td>1.43E-06</td>
<td>-6.199</td>
<td>7.12E-10</td>
<td>***</td>
</tr>
</tbody>
</table>

*Table 2. The linear model of correlations for delayed post-test accuracy.*
References


Mettler, E., Massey, C.M., Kellman, P.J. (2011) Optimizing learning technology through the use of response times.

Neocortical Reinstatement

Elizabeth Shek

The brain is a very powerful organ that typically weighs less than 3 pounds. The cerebral cortex alone has approximately 15-33 billion neurons; each neuron has from one to ten thousand synapses. This results in over one hundred trillion connections in our brain that are always changing with incoming information and experience. The brain is constantly activated; it is absorbing information, retrieving memories and processing information that enters through our senses (Kumaran, Maguire, 2009). No matter how much researchers try, the complexity of our brain cannot yet be fully replicated by any man made machine. This paper focuses on one small but highly integrative area and its fascinating capabilities: the hippocampus.

The hippocampus is a component of the limbic system located in the medial temporal lobe of the brain. It is thought to play a critical role in the formation and retrieval of memories and spatial navigation. This brain region encodes short-term memory into long-term memory, retrieves stored memories and controls spatial navigation through visual and movement related cues. The focus of the research study I have been assisting with this quarter is the intricacy of the hippocampus's retrieval mechanisms.
The hippocampus is thought to be at the end of the line of multiple sensory processes and as such can consolidate information from various brain regions and combines different types of details into one coherent memory. This allows us to incorporate smell, taste and feelings associated to the memory and store it as one whole. Furthermore, when we retrieve a specific memory, research shows that we are reactivating the same regions in our brain that were activated during encoding. Even when a single stimulus is activated in retrieval, the relevant and associated sensory and emotional processing regions are reactivated by default, (Danker & Anderson). This is a phenomenon called pattern completion. The more general idea of reactivating same brain regions during encoding and retrieval is called cortical reinstatement and will be the focus of this paper.

In the first phase of our study, which looks to evaluate mismatch detection and neocortical reinstatement, we present subjects with names and associated pictures of objects, animals or locations. In the second part of the first phase, we ask them to rate their memory of the associated picture when presented with a neutral retrieval cue (the name) as “Strong memory”, “Moderate Memory”, “Weak Memory” and “No Memory” and then determine if a picture presented after the name is an “Exact Same Picture” a “Similar Picture”, or a “Very Different Picture” or if they have “No Recollection”. Between the presentation of the name and the picture, there is a delay period when subjects should be
picturing the original picture paired with the name. After we perfect this study, in the second phase in the scanner, we will train a multivariate pattern classifier to decode patterns of activation for a given visual category. This will allow it to later predict which category has been reinstated during the delay between name and image and let us quantitatively assess how reliably distinct mental states are differentiated by their brain activity patterns.

In their article, “The Ghosts of Brain States Past”, Danker and Anderson review numerous studies that provide evidence that indicate when people retrieve an episode or association from memory, the regions of the brain that were originally involved in processing that episode or association are reactivated. The studies shared a similarity in their methods: subject were shown neutral stimuli accompanied with either a sound, odor or visual cue; during the retrieval/test phase subjects were asked to make some kind of decision or judgment about what they were viewing. In all of the studies, the parts of the brain that were associated with processing the sensory or emotional aspects were reactivated during retrieval, showing clear evidence for cortical reinstatement in multiple regions of the brain. Researchers also manipulated emotion and showed greater retrieval with cues that had a stronger emotional connection (Danker & Anderson, 2010).
Evidence for emotional valiance enhancing memory can be directly applied to our research. Some subjects can hold stronger emotional ties to specific random visual cues presented to them than others. For instance, if the subject was born in San Francisco then the visual cue of the Golden Gate Bridge will be more easily encoded, since they have more emotional attachment and more readily retrieved than for a subject born elsewhere. The same idea holds for the neutral names presented to the subjects. For some people, those names may not be neutral. Since subjects are not showing 100 percent familiarity with the names they are given, we can assume that some names hold more meaning, and hence stronger emotional ties, and are therefore easier to remember than brand new names that they are seeing for the first time.

Furthermore, “Fidelity of Neural Reactivation” article quantifies the degree of brain patterns that are reactivated during retrieval to understand competition between similar memories. This allowed them to quantify the strength of reactivation and showed that the strength or fidelity of a memory is diminished when there are similar competing memories stored in your mind (Kuhl, Rissman, Chun & Wagner, 2011). In our example, if a subject sees a name and immediately associates that name with a random object based on some underlying memory, then the visual association actually presented will be diminished and memory for that association will be reduced. Furthermore, if subjects have a memory that
is confused with another, and they are not sure if a name was paired with one picture or anther, this competition would also degrade the strength of the reinstatement and our classification accuracy.

Additionally, the article titled: “Recollection, Familiarity and Cortical Reinstatement” showed that cortical reinstatement is not solely restricted to episodic retrieval or recollection, but is also present during feelings of familiarity. During familiarity reinstatement, the same pattern of brain regions is activated that is associated with remember-related effects (Johnson, McDuff, Rugg & Norman, 2009). This study also used multivariate classification analysis to make a quantitative assessment to deduce this effect. In our study, subjects gave a rating of how similar the presented image was to the original on a scale from “Exact Same Image”, “Similar Image”, “Very Different” or “No Memory”. If the image on the screen seemed familiar to the original image, even without the exact memory of the image, we should still see brain activation. However, this idea is subjective because people make abstract prototypes and come up with unique and specific similarities, so we would not be able to pinpoint which exact visual cues will be most effected by this phenomenon.

The hippocampus is a very complex region that is integral in the formation and recovery of memory. In the current stage of our research, we are manipulating variables to
account for the biases discussed in this paper. In the actual experimental phase, we will look at what subjects are using to represent their confidence of memory, try to quantify the degree of neocortical reinstatement and relate the degree of reinstatement to hippocampal sub-field activation.
References


Object Recognition

Nicole M. McIntyre

Introduction

In 2001, a study by Fisner and Aslin demonstrated that humans can learn higher-order statistical regularities in visual scenes. They first trained subjects to learn a pair of shapes and then tested how well they were able to detect the true pair by displaying it with a foil (i.e. false) pair and asking them which was more familiar. The foil pairs were different in their shapes and their positioning. They found that the learning of shapes was both rapid and automatic, as although subjects were not instructed to attend to any part of the display, they still learned information about the objects. For example, subjects didn’t only learn single-shape frequencies, but also learned higher-order aspects of the structures (e.g. absolute shape-position relations, shape-pair arrangements independent of position, etc.). Subjects were very accurate at being able to distinguish the true pairs from the foil pairs, even without consciously encoding the information they needed to do so. The real world implications of this study are that humans learn information about objects and visual scenes even when not purposely doing so.

A follow-up study by Fisner and Aslin in 2005 looked deeper into this finding and sought to explain exactly how humans encode information about objects. Fisner and Aslin (2005) looked at how adults encoded and remembered parts of multi-element scenes that were comprised of recursively embedded shape combinations. They again trained individuals to recognize a pair and then tested them on their ability to recognize it when posted against a foil; this time, however, the pairs were embedded in quadruples. Fisner and Aslin (2005) found that embedded shape combinations (i.e. embedded pairs) were less
well-remembered than non-embedded ones (i.e. quadruples and non-embedded pairs). The results suggested humans are not able to detect embedded pairs, only standalones. This finding brought to question how humans encode information about objects. Why can humans encode pairs when they are by themselves but not when they are embedded in a larger figure?

Our study sought to further explain human object recognition. We found the findings of Fisner and Aslin (2005) very interesting based on the real world observation that if a human is shown a box with a lid and then later just show them the lid, chances are that they would be able to recognize the lid. So, we sought to answer the question of whether or not humans can recognize the embedded structure if they learn the counterpart of a particular structure. In other words, can humans learn the counterpart to the embedded pair (FH) if we trained them on both the quadruple (EGFH) and the embedded pair (EG)? We used the same method as did Fisner and Aslin (2001, 2005) and tested subjects on their ability to detect the counterpart to the embedded pair by displaying it next to a foil pair.

Method

Participants

Twenty four young adults (age range: 18-22 years) from the University of California, Los Angeles were recruited through classes offered by the Psychology Department. Participants were given extra credit as compensation for their participation in the study.

Design
The independent variable was positioning and there were two levels - embedded or standalone. The dependent variable was recognition and was measured in percent correct. We implemented a between-subjects factorial design.

Materials and Apparatus

We conducted the experiment on a computer; participants saw images on the screen and recorded their answers using a standard keyboard. A chin rest was used to ensure that all participants saw the screen from the same distance.

Procedure

Training consisted of two phases. The first phase ensured that the participant got a solid representation of the trained embedded pair before seeing it embedded in a quadruple. This phase consisted of a one minute viewing of the target embedded pair (EG). Next, there was a detection task in which participants were asked to detect the pair in 96 different scenes. Finally, there was a familiarity judgment test in which subjects were asked to say which was more familiar – the trained embedded pair or a randomly generated pair. Then, during phase 2, the observers viewed 104 distinct training scenes without doing any task. Each scene consisted of one quadruple and one pair. Frequencies were arranged such that all the untrained embedded pairs occurred at the same rate. Lastly, there was the final testing session which consisted of twelve trials. There were two trails for each of the following conditions: target quadruple, non-target quadruple, pair, counterpart to the embedded pair, untrained embedded pair, and trained embedded pair.

Results

First, we tested to see how effective we were at training the trained embedded pair EG. We found that the training session was very effective at enabling the participants to
detect EG (M = .94, SD = .13). In addition, participants were able to detect EG very well after phase 1 (M = .99, SD = .44).

We also wanted to see how well participants could recognize the counterpart to the trained embedded pair FH, as opposed to another pair such as AC or DB. We found the participants indeed found FH more familiar (M = .82, SD = .24) than AC or DB (M = .66, SD = .34).

Conclusions

The finding that FH, the counterpart to the trained embedded pair EG, was more familiar than AC or DB suggests that humans can detect embedded pairs when they are trained to recognize the embedded structure and one of its pairs. These findings give us important insight into the process of object recognition; they suggest that humans must see the entire objects (e.g. a box) and one of its part (e.g. its lid) to be able to recognize the other part (e.g. lidless box).
References


Interaction of Perceptual Learning Modules and Explicit Instructions in Statistics

Joseph Truong

Abstract

Public education is more focused on conceptual learning, beginning with a semantic introduction. Recent studies (Goldstone & Barsalou, 1998) show that perceptual learning—based on discovering patterns to identify and classify information, and with practice comes fluency with the subject being learned—is a more effective method of instruction. In this study, students are taught statistical main effects and interactions, depicted with graphs, using perceptual learning and conceptual learning to see whether there is an interaction depending on the order it is learned in. Two studies were conducted, both using a perceptual learning module, one using a video, another using a textbook excerpt. The results from the first study were not able to find a significant main effect, as does the second ongoing study, containing large error margins from a low number of participants. There may be an improvement to be made in designing the conceptual learning portion, which will be addressed in a third study.

Introduction

Traditionally, students are taught to learn a subject using a lecture as a cornerstone, and building a mastery of the subject with a hands-on experience or practice, and more lectures. Another view of education by Goldstone & Barsalou suggests that conceptual learning and perceptual learning are very similar, such as the way they both represent and process information. In addition, they believe that perception is the foundation for
conception, that “perceptual processes guide the construction of abstract concepts” (Goldstone & Barsalou, 1998). This means to say that students use perception to help set up a starting base for learning a subject, in which learning takes place in an abstract manner, then allowing conceptual learning to fill in the gaps with specific details more effectively.

Perceptual learning involves extracting information from structures and categories, allowing for the identification and classification of objects, such as identifying a given animal despite varying appearances. Kellman & Garrigan (2009) state that perceptual learning does not depend on the container metaphor of learning, in which information is simply stored in memory and retrieved when needed. Instead, information is obtained dynamically by using pre-made associations and relations. Thus, we can say that perceptual learning is essentially limitless; learning starts with a basis that is constantly built on and interwoven with new information.

In addition, perceptual learning involves discovery and fluency—discovery referring to pattern recognition by discovering features similar to one another, while fluency refers to being able to respond almost automatically because of practice when facing a problem (Kellman, Massey & Son, 2009). Experience correlates with perceptual learning, while formal instruction correlates with conceptual learning. In Kellman, Massey, and Son’s (2009) study, a perceptual learning module (PLM) was implemented to teach participants to recognize patterns and be able to answer questions quickly (fluently) as they continued with the module. The control group received a packet to work on, with an answer key provided to check on their work, but without providing feedback. Conceptually, this would
make participants process problems in terms of memorization as opposed to seeing the problems in terms of patterns. The problems given on both the PLM and the packet consisted of translation problems for linear equations, asking participants, given a word problem, graph, or equation, to choose an answer that depicts a new representation as either a graph or an equation, resulting in four types of translation problems (Kellman et. al, 2009). Results indicated that participants who underwent through the PLM were better at transferring what they learned onto different cases and presentation types.

Learning is also perceptual in arithmetic operations, as shown in a study conducted by Goldstone, Landy, & Son (2009). They found that participants “tended to look at the multiplication portion of the expression” initially in a problem such as “2 x 3 + 4.” This shows that participants were able to automatically apply the precedence rule just by searching for the multiplication sign first and calculating the math that occurs there. This result was seem as “people’s perceptual systems becom[ing] rigged up over practice ... to automatically gravitate” to where the equation should be handled first (in this case, multiplication) (Goldstone et. al, 2009). In addition, the multiplication sign attracted more attention than the addition sign, regardless of the problem. This shows that through perception, selective attention is given to operators of higher precedence for varying problems, non-specific to the type of calculations actually needed.

Using these ideas, this current study, conducted in the Human Perception Lab at the University of California, Los Angeles, was designed to test whether learning was more effective when participants went through perceptual learning followed with conceptual learning, or when participants went through conceptual learning followed with perceptual
learning. Perceptual learning is believed to provide a better basis to learning statistical main effects and interactions than conceptual learning first.

*Experiment 1: Using a video for conceptual learning*

**Methods**

**Participants**

For the first study, 35 students (12 men, 23 women; age approximately 18-25) from the University of California, Los Angeles were recruited in the first study. They participated for course credit and did not receive monetary compensation. Participant registration was restricted to students who did not take an introductory psychological statistics course or the psychology research methods course, which emphasizes learning statistical main effects and interactions, but has been changed to allow students to participate in the experiment if they are concurrently enrolled in an introductory psychological statistics course.

**Design**

The experiment used a 2 x 2 experimental design with the PLM and the video as one independent variable and the order in which the PLM and the video were used. The PLM was designed to be mostly perceptual, while the video was designed to be mostly (if not entirely) conceptual because perceptual learning and conceptual learning are closely tied to one another. The use of a PLM for perceptual learning was effective in that explanations of main effects and interactions were not explicitly stated verbatim, but instead, it was designed to have participants pickup patterns from each problem and transfer each pattern
regardless of the scenario issued (Kellman et. al, 2009). The dependent variable was the number of correct answers chosen on the paper tests issued after each instance of training from the PLM and from the video, measured on an interval scale. Other data was collected from classification exercises conducted before and after the PLM step, measured on an interval scale. Statistics is a subject known to be rigorous and to take a large amount of time and effort to master. Within statistics, main effects and interactions are able to be depicted with graphs and words, which is compatible with perceptual learning (picking up visual patterns) and conceptual learning (verbatim facts and abstract meanings) respectively. Thus, main effects and interactions were chosen as the subject to learn.

Materials

The experiment used a PLM and a video both about main effects and interactions, and did not require that participants had a background in the subject matter.

The PLM was designed to ask participants one of three different trials: participants were given a graph and told to choose the correct statement verbally describing the effect(s) shown; participants were given a statement and told to choose a graph depicting this; and participants were given a graph and asked to describe the main effect or the interaction in terms of ‘greater than,’ ‘less than,’ or ‘equal to.’ Before and after the PLM task, a classification task consisting of 16 questions was issued, which contained questions identical to those asked in the PLM but without feedback. The responses to these questions were recorded in a database to be analyzed. The PLM and video were run-able on either a Macintosh or Windows platform, with the PLM having a point-and-click interface similar to the MultiRep PLM (Kellman et. al, 2009). The situations used were produced by an
experimenter and were equally paired with each combination of main effects and/or an interaction in order to hint to the participant that the situation was an arbitrary factor in thinking about main effects and interactions. The experiment contained eight possible combinations of main effects (one of each) and/or an interaction, resulting in 24 different questions. Both bar graphs and line graphs were used. The module was designed to ask the participants a select number of the 24 different types of questions, and was adaptive in retiring certain questions in which participants were able to answer quickly and correctly while adding a new question for each one retired. The program would later reintroduce the question to check for whether the participants had mastered the question to the point of the question being answered almost automatically. Feedback was given for each question, where each incorrect answer was compared with the correct answer. After 15 trial questions, a progress report was shown to each participant detailing their accuracy and reaction time. Mastery was defined as answering the question correctly five out of six times, having a reaction time of 10 seconds or less, and having answered for both line graph and bar graph scenarios. The PLM was designed to be adaptive, in which the PLM would not end unless all mastery levels were achieved.

The video was around 10 minutes long, consisting of definitions for main effects and interactions, and drawing multiple examples of each using bar graphs and line graphs. The presence/absence of main effects and interactions for each example used were stated, along with the means compared to find the main effects and interactions. Explanations of the presence/absence of main effects and interactions were shown through a printed statement, using visual cues to hint at where to look in the graphs, and read aloud by a
speaker in the video. Each learning method was designed to consist of a majority of perceptual learning and conceptual learning, placing an emphasis on each one respectively.

Three similar paper tests were issued to the participant: one before the experiment; one after either the PLM or the video, and a final test after the follow-up learning mechanism (the PLM if the video was shown first and vice-versa). These tests contained 5 questions, asking the participant to identify and explain main effects and interactions, and draw a graph corresponding to the presence or absence of main effects and interactions, and translate problem solving to a 2x3 study situation. The paper tests were graded by two graders to check for consistency and for inter-rater reliability. The experiment was conducted in a three hour session. A log sheet created on Microsoft Excel was used to record demographic information regarding name, sex, date, statistics background, year in school and major, additional notes about any issues that came up in the experiment.

Procedure

Participants were asked for some demographic information and assigned to a condition, either PLM first or video first. Next, participants were given the pretest paper test to complete as a diagnostic test. They were told to answer all questions and write, "I don't know," or some variation or it if they had no idea what to do. Otherwise, if they wanted to guess or had some idea on how to answer the questions, they were encouraged to answer clearly as necessary. After completing the paper test, they turned in the test to proceed to either the PLM or the video, depending on their condition. For the PLM step, a powerpoint presentation was used to introduce the functionality of the PLM, in order for the user to be familiar with it, and to stress the PLM’s adaptive feature. After presenting the
presentation, the experimenter created a new account for the participant, and the participant logged in to complete the PLM. Before doing the actual PLM, participants completed a short, 16 question classification task. While the participant worked on the PLM, the experimenter would periodically monitor the participant’s progression in the program, and try to encourage breaks if the participant was not cooperating. Participants were told to continue the module until they mastered each of the 24 different questions, at which point, the module would come to a halt and present a final classification task. If participants could not complete the PLM within two hours time, the experimenter would manually stop the participant. The experimenter would then set up a new account, and tell the participant to only complete the classification task. For the video task, participants were given a 10 minute video to watch, listen, and study for a maximum 20 minutes, although they were allowed to move onto the next step before the 20 minutes as long as they watched the entire video at least once. After either condition presented first, a midtest paper test was issued after the condition, with the same instructions as the pretest paper test. After the midtest, the second condition was given, followed by the final paper test. After this, the experiment ended and participants were debriefed.

Results

Figure 1 shows the mean test scores for the three tests for each condition. Looking at this pattern or results, there is a main effect increase from pretest to midtest, and pretest to posttest, but not significant main effect between midtest to posttest. There was not a significant interaction. Figure 2 shows the mean test scores for the three tests for each condition, ignoring those who did not get at least one mastery level. Looking at these
results, the same pattern is seen, as with Figures 3 & 4. Figures 5, 7, and 8 do not show a significant main effect in yes/no-type questions, draw-type questions and 2x3 problem-type questions. Figure 6 does show a significant difference in the midtest between PLM first and video first. Figure 9 shows a main effect between prePLM and postPLM in the classification task, and a possible interaction.

Discussion

The current experiment’s results show no main effect other than that between the pretest and midtest, pretest and posttest, and a difference in the explain-type questions between the PLM first and video first condition. In the classification task, a main effect was present between prePLM and postPLM, and a potential interaction was spotted. Based on these results, we conclude that just after one type of instruction, a large increase in learning is seen. The video and the PLM both increase student learning to a ceiling effect. The video may have been designed to be too perceptual, because it contained graphical illustrations and comparisons of each type of main effect and interaction combination. A later study conducted replaced the video with a textbook description excerpt about main effects and interactions. The results contain a large error margin, due to the small number of participants in each condition. In addition, the three hours allotted for the experiment may have had a fatigue effect between the midtest and posttest, which would explain why there was little increase in test scores. A future study may try to break the experiment into smaller modules, or conduct the experiment along a set of days, to avoid the fatigue effect.

Experiment 2: Using a textbook description for conceptual learning

Methods
Participants

Twenty-eight students (10 men, 18 women; age range: approximately 18-25 years) from the University of California, Los Angeles were recruited for a second study. Participant restriction was the same as the first study.

Design

The experiment used a 2 x 2 experimental design with the PLM and the textbook description excerpt as one independent variable and the order in which the PLM and the description were used. The PLM was designed to be mostly perceptual, like the first study, while the description was designed to be highly conceptual and more realistic from an educational stance. The dependent variable was measured in the same way as the first study: the number of correct answers chosen on the paper tests issued after each instance of training from the PLM and from the video, measured on an interval scale. Other data was collected from classification exercises conducted before and after the PLM step, measured on an interval scale, as was in the first study.

Materials

The experiment used a PLM with the same functionality as the first study. Instead of a video, a textbook description excerpt about main effects and interactions was used.

The description was printed on three sheets of paper, (1.5 pages front and back). It consisted of definitions for main effects and interactions, and contained a few examples defining main effects and interactions. The presence/absence of main effects and
interactions for each example used were stated, along with the means compared to find the main effects and interactions.

The same three paper tests and log sheet was used, as was from the first experiment.

Procedure

The same set of procedures was used as was in the first study, except with the incorporation of the textbook description in place of the video.

Results

In Figure 10, there is a main effect between pretest and midtest, and pretest and posttest, but not between midtest and posttest, like in the first study. Figures 11 and 12 show the same effect between participants who completed at least five mastery levels and those who completed all 11 mastery levels respectively. There is a large error margin because of the low number of participants in this study. In Figure 13 & 14, showing the results of the classification task, there is a main effect between prePLM and postPLM results, showing that the PLM did help those who worked on it. Those with the description are expected to get better results because they received training in main effects and interactions from the textbook description excerpt, prior to completing the PLM. Figure 15 & 18 show no main effect between pretest, midtest, and posttest for Yes/No-type questions and 2x3 problem-type questions respectively. Figure 16 & 17 show a main effect between pretest and midtest, and pretest and posttest, but no main effect between midtest and posttest.
Discussion

This second study is currently being conducted, but based on the data gathered so far, the results seem to be consistent with the first study. We cannot conclude anything from this study until it is completed with enough participants. From what has been gathered so far, there may be a better way to approach the problem of perceptual learning first or conceptual learning first. We may decide to break up statistical main effects and interactions into two different studies. This would help decrease the experiment time from three hours, in order to avoid fatigue effect and ensure participants to complete the PLM. It is interesting to note that after one method of learning, students have learned approximately to the same level.

Implications for this study can apply to the educational setting, changing the way students are taught in order to improve fluency of a subject, if perceptual learning first is found to be more effective than conceptual learning first. By being more fluent, one is able to learn new things much easier using older subjects learned as the foundation.
References


Figure 1. Mean test scores of each test for each condition, using all valid participant data. A main effect was found between Pretest and Midtest, and Pretest and Posttest, but not between Midtest and Posttest, nor was an interaction found.
Figure 2. Mean test scores of each test for each condition, using valid participant data of those who completed at least 1 mastery level. A main effect was found between Pretest and Midtest, and Pretest and Posttest, but not between Midtest and Posttest, nor was an interaction found.
Figure 3. Mean test scores of each test for each condition, using valid participant data of those who completed at least 5 mastery levels. A main effect was found between Pretest and Midtest, and Pretest and Posttest, but not between Midtest and Posttest, nor was an interaction found.
Figure 4. Mean test scores of each test for each condition, using valid participant data of those who completed all 11 mastery levels. A main effect was found between Pretest and Midtest, and Pretest and Posttest, but not between Midtest and Posttest, nor was an interaction found.
Figure 5. Mean test scores for Yes/No type questions for all valid participants in Study 1.
Figure 6. Mean test scores for Explain-type questions for all valid participants in Study 1.
<table>
<thead>
<tr>
<th></th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Video</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>1.25</td>
<td>0.232737</td>
<td>1.357143</td>
<td>0.307545</td>
</tr>
<tr>
<td>Mid</td>
<td>2.5625</td>
<td>0.157288</td>
<td>2.357143</td>
<td>0.169263</td>
</tr>
<tr>
<td>Post</td>
<td>2.5625</td>
<td>0.18186</td>
<td>2.571429</td>
<td>0.137253</td>
</tr>
</tbody>
</table>

N=16    N=14

Figure 7. Mean test scores for Draw-type questions for all valid participants in Study 1.

<table>
<thead>
<tr>
<th></th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Video</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>2</td>
<td>0.312916</td>
<td>2.214286</td>
<td>0.321886</td>
</tr>
<tr>
<td>Mid</td>
<td>2.6875</td>
<td>0.315816</td>
<td>3.428571</td>
<td>0.507789</td>
</tr>
<tr>
<td>Post</td>
<td>3.3125</td>
<td>0.415519</td>
<td>3.571429</td>
<td>0.398683</td>
</tr>
</tbody>
</table>

N=16    N=14
Figure 8. Mean test scores for 2x3 Problem-type questions for all valid participants in Study 1.

<table>
<thead>
<tr>
<th></th>
<th>pre</th>
<th>std error</th>
<th>post</th>
<th>Std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLM first</td>
<td>6.076923</td>
<td>0.635101</td>
<td>12.15385</td>
<td>0.648926</td>
</tr>
<tr>
<td>Video first</td>
<td>8.333333</td>
<td>0.84686</td>
<td>11.33333</td>
<td>0.74196</td>
</tr>
</tbody>
</table>

N=13, N=12

Figure 9. Mean classification task scores. Shows a main effect between prePLM and postPLM, and a possible interaction.
Figure 10. Mean test scores of each test for each condition in Study 2, using all valid participant data.
Figure 11. Mean test scores of each test for each condition in Study 2, using all valid participant data for those who completed at least 5 mastery levels.

<table>
<thead>
<tr>
<th>≥ 5 ML</th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Desc</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>8.66667</td>
<td>0.607196</td>
<td>8.29167</td>
<td>1.270973</td>
</tr>
<tr>
<td>Mid</td>
<td>11.0833</td>
<td>1.404583</td>
<td>12.29167</td>
<td>1.391884</td>
</tr>
<tr>
<td>Post</td>
<td>14.75</td>
<td>1.294657</td>
<td>15.43333</td>
<td>1.077267</td>
</tr>
<tr>
<td>N=12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11 ML</th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Desc</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>9.2</td>
<td>0.982344</td>
<td>10.7</td>
<td>1.079352</td>
</tr>
<tr>
<td>Mid</td>
<td>12.8</td>
<td>2.390607</td>
<td>14.7</td>
<td>1.3</td>
</tr>
<tr>
<td>Post</td>
<td>15.4</td>
<td>2.486966</td>
<td>17.6</td>
<td>0.966954</td>
</tr>
<tr>
<td>N=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 12. Mean test scores of each test for each condition in Study 2, using all valid participant data for those who completed all 11 mastery levels.

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Desc</th>
<th>Std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td></td>
<td>6.111111</td>
<td>0.789593</td>
<td>7.916667</td>
<td>0.690392</td>
</tr>
<tr>
<td>Post</td>
<td></td>
<td>10.111111</td>
<td>0.857069</td>
<td>11.91667</td>
<td>0.792563</td>
</tr>
</tbody>
</table>

N=9       N=12

Figure 13. Mean classification task scores. Shows a main effect between prePLM and postPLM.
Figure 14. Mean classification task scores for participants who completed over 5 mastery levels.
Figure 15. Mean test scores for Yes/No-type questions for all valid participants in Study 2.

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Mid</th>
<th>Post</th>
<th>N=13</th>
<th>N=14</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLM</td>
<td>4.692308</td>
<td>3.769231</td>
<td>4.923077</td>
<td>4.571429</td>
<td>4.857143</td>
</tr>
<tr>
<td>std error</td>
<td>0.286101</td>
<td>0.342574</td>
<td>0.329365</td>
<td>0.388209</td>
<td>0.32912</td>
</tr>
</tbody>
</table>

Figure 16. Mean test scores for Explain-type questions for all valid participants in Study 2.

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Mid</th>
<th>Post</th>
<th>N=13</th>
<th>N=14</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLM</td>
<td>1.230769</td>
<td>2.307692</td>
<td>3.807692</td>
<td>1.178571</td>
<td>3.107143</td>
</tr>
<tr>
<td>std error</td>
<td>0.351806</td>
<td>0.511215</td>
<td>0.52969</td>
<td>0.275398</td>
<td>0.426391</td>
</tr>
<tr>
<td>Desc</td>
<td>1.178571</td>
<td>3.107143</td>
<td>3.428571</td>
<td>0.275398</td>
<td>0.426391</td>
</tr>
<tr>
<td>std error</td>
<td>0.178571</td>
<td>0.275398</td>
<td>0.426391</td>
<td>0.275398</td>
<td>0.426391</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=13</th>
<th>N=14</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>
Figure 17. Mean test scores for Draw-type questions for all valid participants in Study 2.

<table>
<thead>
<tr>
<th></th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Desc</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>1.076923</td>
<td>0.264612</td>
<td>1.071429</td>
<td>0.266526</td>
</tr>
<tr>
<td>Mid</td>
<td>2.461538</td>
<td>0.243252</td>
<td>2.142857</td>
<td>0.231031</td>
</tr>
<tr>
<td>Post</td>
<td>2.538462</td>
<td>0.243252</td>
<td>2.5</td>
<td>0.228468</td>
</tr>
</tbody>
</table>

N=13      N=14

2x3 Problem

<table>
<thead>
<tr>
<th></th>
<th>1) PLM</th>
<th>std error</th>
<th>2) Desc</th>
<th>std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>2.076923</td>
<td>0.361825</td>
<td>2.142857</td>
<td>0.364754</td>
</tr>
</tbody>
</table>
Mid | 2.615385 | 0.449907 | 3.107143 | 0.489389  
Post | 3.769231 | 0.422202 | 3.728571 | 0.392992  
N=13 | N=14  

Figure 18. Mean test scores for 2x3 problem-type questions for all valid participants in Study 2.
Improving Computer-Based Learning: How the Introduction of Game-Like Elements May Improve Learning

Brooke Hollyfield

In recent years, the use of technology has increased dramatically. In current education, technology is used to create multimedia presentations and simulations that enhance the teaching methods of scientific concepts. In addition, the entertainment industry has capitalized on technology, designing exciting video games that captivate children's attention. Growing up, children are told that they must finish their homework before playing video games; this reinforces the idea that video games and education are separate entities of technology. However, researchers have started to investigate how the highly motivational environment created by video games could actually be used to improve learning technology. We live in a society in which a student could be an expert at HALO, but unable to pass a basic math test. If we have perfected the art of captivating attention and motivation through video games, why can we not successfully apply this to computer based learning as well?

In my experience at the UCLA Learning and Forgetting Lab, many experiments have been discussed that involve computer-based learning that uses multimedia presentations about scientific topics. Most of the research presented this quarter discussed ways to improve computer-based learning by applying memory principles (i.e. massed presentation vs. spaced presentation), or by measuring how different groups are affected by computer-based learning (i.e. ADHD individuals vs. non-ADHD individuals). The goal of this paper is to explore how computer-based learning could potentially be improved by
applying game-like principles, specifically competition. I will present two articles that examine the consequences of adding competition to computer-based learning. In the first article, Van Eek et. al (2002) examine situations in which competition may not be beneficial to enhancing transfer in computer-based learning. The second article by DeLeeuw et. al (2011) looks at how competition elements were found to positively effect students’ levels of attention but not motivation in computer-based learning.

Van Eek et. al (2002) wanted to demonstrate how competition and contextualized advertisement would effect transfer in computer-based learning. Competition has previously been shown to increase intrinsic motivation to learn by motivating the player to beat their own score or the scores of others. Competition is perhaps one of the most captivating elements of video games. Hence, by adding this game-like element to computer-based learning, effectiveness should be increased. However, Van Eek et. al point out that this may not be the case. Competition has been shown to be ineffective in enhancing learning when students are in the acquisition phase of learning (the process of acquiring a set of skills), and only effective when students have reached the practice stage of learning (where they are starting to mostly give right answers.) The effectiveness of competition also depends on the students being able to do better in the competitive environment than in a non-competitive environment, or else they will not be motivated to improve. Lastly, competition is not effective in motivating learning if the student does not have the ability to complete the problems; for example, competition won’t help in a situation where the student continually gets stuck on answering problems. Due to these controversial effects of competition, it is unclear competition will effect computer-based learning.
Advertisement is another element that could be potentially added to computer-based learning to make it more game-like. Many video games include characters that provide useful information and guidance for solving problems within the game. Most computer-based learning programs also have some form of help available, but it is usually a “help” icon. Van Eek et. al wanted to test how an integrated form of help that was consistent with the storyline (like the kind used in video games) would effect computer-based learning. Van Eek et. al hypothesized that conceptualized advertisement that was consistent with the storyline of the game would be most effective because it would not disrupt the fluency of the game.

Van Eek et. al (2011) designed a study that manipulated competition and contextualized advertisement to see how they effected performance on transfer questions. Transfer questions are questions that do not explicitly replicate types of problems seen in the learning phase, but that require the same set of skills to be applied. They selected transfer learning because it is a deeper form of learning that requires making connections to prior knowledge and understanding how that knowledge can be applied to diverse situations. Participants in the study would play a simulation computer game. The premise of the game was that the player was asked to help fix things in his/her aunt and uncle’s home remodeling business. The different things needing fixing in the house would involve math concepts including number sense, measurement, and geometry. For example, one task would be to paint the living room, which required finding the areas of all the walls and then calculating how many gallons of paint would be needed for the task. There were four different conditions: competition and contextualized advertisement, each with two levels: competition and non-competition and contextualized advertisement and non-
contextualized advertisement, respectively. In the contextualized advertisement conditions, participants had immediate, constantly available access to problem solving assistance. They had the option of pressing a walkie-talkie, which would call their aunt and uncle into the room and allow for them to ask them questions about the problems. In the non-contextualized condition, they had no access to this help. In the competition condition, participants were told that they were playing against a computer character in both time and accuracy. In the non-competitive condition, there was no computer opponent.

Eek et al (2011) hypothesized that participants in the contextualized advertisement condition would have higher transfer scores than those in the non-contextualized advertisement conditions. They also hypothesized that non-competitive conditions would have higher transfer scores than competition conditions. Their results were somewhat consistent with their predictions; a significant interaction was found between competition and contextualized advertisement. They found that in competitive conditions, participants that had contextualized advertisement performed worse than those with non-contextualized advertisement. In the non-competition conditions, participants with conceptualized advertisement performed better than those with non-conceptualized advertisement. These results suggest that competition makes it difficult for contextualized advertisement to be processed effectively. Eek et al propose that this could be due to higher levels of anxiety and arousal due to the competition, which decreases the resources that can be used for processing deeper learning. Competition may also increase cognitive load and make it harder to make connections to prior learning, hence negatively affecting transfer. Non-competitive conditions seem to benefit from contextualized advertisement. Without the anxiety of competition, it’s possible that participants in the non-competition
group were able to devote more resources to the contextualized advertisement and use them to process deeper learning. These results reveal that it is important to take context into account when deciding whether or not to add competition and contextualized advertisement to computer-based learning. Competition may only be a beneficial element when there is no contextualized advertisement involved. The conclusions drawn from Eek et. al (2002) do not provide enough evidence for whether or not competition should be withheld completely from computer-based learning. A second article, by DeLeeuw et. al (2011) reveals that there is a small benefit to introducing competition in computer-based learning.

DeLeeuw et. al (2011) also wanted to analyze the introduction of competition to computer-based learning. They used a computer-based learning game called the Circuit Game, which uses ten different progressively difficult levels to teach circuitry. The last level is an embedded transfer test that measures how well students can apply knowledge learned in circuits to problems with light bulbs. DeLeeuw et. al added elements to make the Circuit Game more competitive. These included inserting a graphic at the side that showed the participants’ current score, and introducing a ticket system based on success at each level. If the participant completed a level correctly, they would receive a ticket. Their current ticket count would be displayed before each level. Each ticket would be entered into a raffle with the chance of winning $50. The experiment tested two groups: one with the competition elements and one without. Participants would complete a pre-study questionnaire, followed by the Circuit Game, then a questionnaire about mental load and a retention test. The transfer test was embedded within the last level of the Circuit Game.
DeLeeuw et. al found that the addition of competitive elements did not significantly improve scores on the transfer question. This result is consistent with Van Eek et. al (2011)’s finding that transfer test was unaffected by competition. However, DeLeeuw et. al found that participants in the competition condition and non-competition had no difference in mental effort ratings. This suggests that the difference in performance on transfer tests may not be due to competition increasing working memory load and hindering processing at a deeper level. More research on how competition effects mental effort and working needed to further explore this discrepancy between the experimental findings of DeLeeuw et. al and Van Eek et. al. On the retention test, DeLeeuw et. al did found that the participants in the competition condition did better than those in the non-competition condition. This finding supports the idea that competition enhances memory straightforward information. This could be problematic for the long-run effects of computer-based learning, because retention measures are not generally related to increasing long-term knowledge of information. Nevertheless, it does suggest that there is some inherent benefit to being intrinsically motivated by competition in a learning task.

The two studies discussed both found evidence supporting the idea that competition somehow hinder the acquisition of deep knowledge required for transfer in computer-based learning programs. This is disheartening for the prospects of incorporating game-like design into computer-based learning. One of the central features that makes video games appealing is the intrinsic motivation to continue the task, which is presumably caused by competition. Since competition appears to hinder rather than help the understanding of deep concepts, it is unlikely that it will be an effective addition to computer-based learning programs that have the goal of improving transfer. However,
competition is only one example of a game-like element that could be useful in improving transfer in computer-based learning. Other game-like elements including color coding, virtual reality, narrative theme, and user interaction could potentially be more beneficial to improving computer-based learning. More research is needed to fully understand how the captivating quality of video games can be adapted for use in improving the modern education computer-based learning system.

Thoughts about the Lab Experience:

I thoroughly enjoyed my experience as a research assistant in the Learning and Forgetting Lab under my graduate student advisor Michael Cohen. The only thing I would have liked to change about my experience is to have had the time to conceptualize and design one of my own experiments, which unfortunately I did not have the opportunity to do to my brief time in the lab. I felt that I learned the most by attending lab meetings and journal clubs. I become more successful at analyzing research articles due to the weekly discussions and lab presentations. I had never felt comfortable tackling papers as complex as the ones we read every week for journal club, but after the quarter I have more confidence in my ability to successfully understand experimental design, point out flaws of the studies, and understand how conclusions are drawn from complex statistical methods. Most importantly, I enjoyed working with graduate students, professors, and lab managers who were clearly incredibly passionate about their areas of study.
References


Thyroid Hormones Dynamics Stimulation in MATLAB

Benjamin Steeper

Introduction

Eisenberg et al (2006) developed a feedback control system model to simulate TH regulation in the hypothalamo-pituitary-thyroid axis of adult humans and used it to assess bioequivalence of levothyroxine (L-T4) in treatment of CH\(^2\). That model was improved (Eisenberg et al, 2008) to include submodels for the brain and was validated with pharmacokinetic data. They found optimal doses of L-T4 for various treatment administration methods in adult humans\(^3\). The model encompasses the dynamics of the pituitary, thyroid, and hypothalamus as their role in hormone production and the degradation and elimination of T3, T4, and TSH as the sink.

In the 2006 model, a six compartment feedback control systems model was implemented. This six compartment model can be seen as the central six compartments in Figure 1. The six compartments represented the slow, fast, and free pools for T3 and T4 in the adult human. The conversion of T4 in to T3 was approximated with three Michaelis-Menten hill functions which represent the extravascular conversion enzymes. One equation was used for the fast pool but the slow pool was split into two equations to represent the 80:20 conversion ratio of deiodinase D2 and D1. Binding equations were incorporated for T3 and T4 to represent the free hormone in the plasma as a function of the total hormone, which included that bound three binding proteins thyroid binding: globulin, albumin, and transthyretin\(^2\).
The model was updated in 2008 to include a brain submodel in order to include TSH as a controllable variable and to be able to predict TSH values given specific cases. Eisenberg et al accomplished this by incorporating an oscillating sinusoid into a secretion rate equation for TSH which was dampened by a function of T³ in the brain. That secretion rate equation simulated the circadian rhythm of TSH being released into the plasma. Plasma TSH therefore became another compartment, which in turn, affected the original six compartment model. This made the model more accurate and allowed for the measurement of TSH values.

**Figure 1.** Full implementation of the model in SAAM II. The brain submodel is in the lower right and its output enters the rest of the system through SR3 and SR4, both of which are delayed by the delay blocks. The other compartments, q14 through q17 are for exogenous input, i.e. pills. The original model is seen in the inner six compartments, q1 through q6.
The goal of this project is to represent the model developed by Eisenberg et al\textsuperscript{3} in MATLAB and Octave, thereby making it more distributable. The goal for the MATLAB model is to convert the finished model into Systems Biology Markup Language so that it can be made available to schools for instruction via the internet. The goal for the Octave implementation is an online interface for public use.

\textit{Method}

The equations from Eisenberg et al, 2008\textsuperscript{3} were converted to MATLAB code by hand. Many errors were found in the equations along the way and those were corrected both in the MATLAB code and in the equations list. Parameters from the model were converted to consistent units and label p1 through p48 according to the order in which they appeared in the equations.

The state variables were labeled q(1) through q(13) and their rate equations were denoted q1dot through q13dot. Six extra compartments were added to implement the time delay in SR3 and SR4 (equations 1 and 2).

\begin{equation*}
SR_3(t) = S_3 TSH_p(t - \tau) \quad SR_4(t) = S_4 TSH_p(t - \tau) \quad \text{(Equations 1 and 2)}
\end{equation*}

The six delay compartments were labeled q(14) through q(19). Those six compartments take TSH as input and act to delay it by eight hours. The delay effects are shown in figure 2. The resultant delayed value, q(19), replaces TSH in the equations above and thus feeds the appropriate values into the T3 and T4 pools. Furthermore, an extra parameter, k\text{delay}, was added to the list of parameters and used in each of the six delay
equations. Kdelay was needed to make the delay block, as a whole, delay its input for 8 hours.

\[
\text{Figure 2} \quad \text{The plot of amount of TSH in } q(14) \text{ (top) through } q(19) \text{ bottom versus time when a bolus input of 1 is input at time equals zero.}
\]

We had exceptional difficulty with q(8), T3 in the brain, which exploded at time equals thirty minutes. The equation was described in error but careful observation led to the simple correction and all of the values steadied after that.

Initial conditions were pulled from Eisenberg et al (2008)\textsuperscript{3} and from Achour’s previous Python model\textsuperscript{1}. The equations were run with both sets of initial conditions. MATLAB’s library ODE solvers were used, primarily ODE23s and ODE45. No difference was shown between the two.

\textit{Results}
When run with the initial conditions from the Eisenberg et al paper, stable but incorrect output was returned. This output is generated in Figure 3. With these initial conditions T4 peaks to 28.06 μg/dL, T3 peaks at 14.19 μg/dL, and TSH peaks to 473.3 μg/dL.

**Figure 3** Concentration versus time using the Eisenberg et al initial conditions.

When the model was run with initial conditions from the Python code, much more realistic results were returned. These results are detailed in Figure 4.
With both set of initial conditions, the values reached at the end of 700 hours have leveled off to similar values. In Figure 3, T4 has leveled off to 8.94 μg/dL and in Figure 4 it is 8.67 μg/dL. T3 has reached 1.26 μg/dL in Figure 3 and 1.32 μg/dL in Figure 4. TSH is 3.49 μg/dL in Figure 3 and averages to 3.72 in Figure 4.

**Discussion**

All of the peak values corresponding to Figure 3, which used the Eisenberg et al initial conditions\(^3\), are much too high, because the model is supposed to be in steady state at time equals zero. That means that the graph should remain steady or in oscillation.
around the same mean for the duration of the simulation. This is because there are no exogenous inputs so the system is only affected by the initial conditions, secretion rate, equations, and sinks, which should be in steady state if the initial conditions are correct. The graphs do level out to the approximate steady state values that are expected, after 700 hours.

When run with the initial conditions from Achour’s model\(^1\), the MATLAB model performs very well. The initial peaks are very low, which is desirable because if the system is already in steady state then there should be no peak at all. One would expect just a smooth oscillation.

The fact that Achour’s initial conditions\(^1\) outperformed Eisenberg et al’s\(^3\) is disconcerting because the MATLAB model was implemented directly from the model that Eisenberg et al created\(^3\) and used the very same equations and parameters that they tested and presented. Further, investigation is required to understand if there were changes made to the equations since they were used in the model which may have affected the initial conditions.

The final goal of the MATLAB project was to convert it to Systems Biology Markup Language (SBML). It appears that the MATLAB code cannot be directly translated into SBML, but must first be implemented in SimBiology, MATLAB’s own computational systems biology language. This must be completed before the model can be made accessible to the public in SBML format.
References

Achour S. Python Model, 2011

Eisenberg M., Samuels M., DiStefano J.J., Thyroid, 16, 1279-1292, 2006

Eisenberg M., Samuels M., DiStefano J.J., Thyroid, 18, 1071-1085, 2008
I decided to major in Cognitive Science because I was fascinated by the brain and mystified by its abilities to create the reality that exists today for each individual that is lucky enough to have one. I joined the reasoning lab at University of California, Los Angeles, run by Holyoak, last quarter in order to get more hands-on experience in the subject and dig deeper into the secrets of the mind. I gained a lot of valuable experience running subjects in experiments, cleaned data extracted from experiments, and qualitatively analyzed data from subjects' responses. I continued working in the reasoning lab this quarter alongside my previous adviser, Vendetti. I helped with the next phase to the Mapping Induced Forgetting project, worked on a new project running subjects in an MRI machine, and gave a good proportion of extra work into helping out the reasoning lab during my busiest and final quarter as an undergraduate.

Markman and Gentner (2000) described that comparisons are created between objects when: they display a sense of alignability with respect to one another, the semantic properties intrinsic to each of the objects are removed, and the major commonality between the two objects is allocated. Alignable objects share powerful characteristics and allow for intuitive comparability. Non-alignable items possess no primary attribute for which both can easily compare to and are therefore more difficult to pin-point a common similarity (Markman & Gentner, 2000). Their experiments showed increases in subjects' accuracy on a recognition test following tasks where alignable objects were seen (Markman & Gentner, 1996, 2000; Gentner & Markman, 1997).
The Mapping Induced Forgetting experiment I helped out with last quarter posed to expand Markman and Gentener's research by examining what occurred to the memory of non-alignable items. Did the alignable objects simply allow for the weights between their synapses to increase (Markman & Gentner, 1996, 2000; Gentner & Markman, 1997)? Would the memory for alignable objects decrease and memory for non-alignable objects increase for accurate responses on a recognition task if the objects being compared were transformed so much that subjects had to think more abstractly to make the comparison (Gick & Holyoak, 1983)?

The Mapping Induced Forgetting experiment cycled through 9 sets of photos. The base photo remained constant among all subjects and appeared first in each set. The secondary photo shown was one of two possible pictures. Each secondary photo possessed similarities to the base photo that resided in the same set. Each secondary photo, however, contained a different mapping of the items within its contents, alignable or non-alignable, to that of the base image. These alignable and non-alignable objects were then taken and shown to the subjects in a recognition test at the end of the study. I qualitatively analyzed data from the previous quarter along with another undergraduate student in the reasoning lab. We continued where we had left off from the previous quarter, and because we had great agreement in our analysis, data was split in half and we each graded a different set of subjects. Each response was read and analyzed to determine how many of the visible objects in the photo were mentioned in the response, as well as how many of those objects mentioned were determined alignable from the previous base image. It was decided that a third group was necessary to add to the experiment upon completing the analysis.
The original groups were the compare and describe group. The describe group wrote a description for each image seen. The compare group compared the base image to the particular secondary photo they were shown for each set (the describe and compare groups’ responses were that which I qualitatively analyzed for each set). The third group added this quarter was an object group; subjects in the object task were told to count the number of objects in each image and enter that number. This created a much more primitive group in order to see if a better understanding could be derived for how subjects were performing on the recognition task based off of the experimental group they fell within. Vendetti requested me to help him with a different project, so I put the Mapping Induced Forgetting on hold and jumped over to MRI studies.

When people used to study the anatomy of the brain, they had to either open up the skull while the person was under some traumatic injury to perform a miraculous surgery, they took a brain from a body that was already dead, or simply derived non-scientific conclusions by scanning the surface of the scalp. As technology has progressed, safer ways to view the brain have been developed and continue to be perfected. Functional Magnetic Resonance Imaging machines, or fMRI machines, are one of those techniques that have drastically improved the science that deals with solving the mysteries of the brain. Running subjects in an fMRI machine is completely safe and non-pervasive (if properly executed since radio waves cause no damage). Damage does not have to be done to the skin in order to catch a glimpse of the mind at work.

MRI machines have been used in the past to determine which parts of the brain are specifically recruited for certain activities. One such example of this is the extensive
research regarding the fusiform gyrus, found in the temporal lobe, and its abilities to determine if the subject is looking at a face or not. Meng et al. (2012) reports that the fusiform gyrus functions in various ways; the function the fusiform gyrus serves depends on the hemisphere it lies in. The left fusiform gyrus apparently fires when any face is being reflected onto the surface of the retina (human, abstract, cartoon, and even animal faces). The right fusiform gyrus supposedly then takes this “raw face or no face calculation” and determines if the face being examined is a human face or not. (Meng et al., 2012).

While the previous study did a great job showing how different stimuli can trigger certain parts of the brain to be recruited more or less depending on the type of stimuli the subject is being presented, would such affects be easily documented with stimuli that are more abstract than a picture? Tom Mitchell et al. (2008) placed subjects inside an MRI machine and showed them different nouns. It was found that a lot of the nouns shown to the subjects produced a brain activity that was predictable for the meaning of the noun (Mitchell et al., 2008). This is an interesting finding because if subjects were shown an unfamiliar language, the brain activity would show as a static, never-changing signal each time (more or less with insignificant change); each stimuli presented was a word in black font and understood by the subjects. If the language was foreign, these words would have simply produced changes in brain activity mainly in the occipital lobe. However, subjects read the words and examined the semantic meanings of the particular words they were looking at (Mitchell et al., 2008). Vendetti proposed to test whether a more abstract set of stimuli could enable for similar predictions to be made; could subjects be given an analogy, and based off of the brain activity seen, could the current analogy set be accurately predicted?
I was the pilot subject in this study; Vendetti ran me through the entire process, scanned to test for any potential flaws in the experiment, and got a glimpse of what he should expect when the real subjects were brought in. I had to get certified in order to run subjects in the MRI machine. While the machine is essentially safe, there are many things that may potentially go wrong, and it is important to be prepared and knowledgeable of possible dangers that exist near the giant magnetic field that is always on. I helped Vendetti run subjects once I was certified; subjects came in, filled out the safety forms, underwent a practice experiment with a computer simulation, were stripped of metal objects, got strapped onto the machine's runway, and were sent away into the magnetic field to begin the experiment.

Meetings were held every Friday with members of the reasoning lab. Each member took a turn to lead a discussion in an article. I led a meeting and discussed an article that was similar to the MRI experiment I was helping Vendetti with. Adam Green et al. (2011) described an experiment that created a way to measure creativity in analogies made by subjects. The gravity of the creativity was successfully measured by looking at the brain activity of the frontopolar cortex noted in the MRI blood oxygen level-dependent (BOLD) signal during the task. More of the frontopolar cortex was recruited when making a more creative analogy (Green et al., 2011). Results from this study indicated possibilities that predictions were possible in determining the creativity of the analogy made based off of looking at brain activity levels in the frontopolar cortex.

Having specialized in computing, I was able to successfully aid Vendetti in consolidating his program in Matlab to work more properly and efficiently in gathering all
the necessary info required. I showed him a few programming techniques: using for-loops and functions to greatly expedite in his finishing the program and compacting the code. His future programming will greatly benefit from incorporating these methods. I then double checked and created new trials for Vendetti’s People Pieces experiment; I went through Vendetti’s original Superlab file and changed trials for certain blocks matching the specifications he gave to me.

I also created a set of stimuli for another project Vendetti was planning to begin. Vendetti gave me eight sets of photos; each of the photos contained two images placed one over the other (one filled the top half of the page, and the other filled the bottom half of the page). These original images were blurry and poorly drawn. I touched up each of these photos using Photoshop; I redrew the outlines for the images for parts that were blurry or hard to see. I then removed parts of the images that were poorly drawn and redrew them either from scratch or from an image found from Google Images, giving it my own style and touch while sticking to the original theme. I completed editing the images in about ten days. Vendetti began using these images in a new experiment soon after I fixed them.

I really enjoyed working with Vendetti in the reasoning lab. Using the experience I gained from the previous quarter I was able to add on to and enhance it even more this quarter. I continued using previously obtained skills of running subjects in computer-based experiments and expanded to help run subjects inside an MRI machine. I gained hands on experience in creating stimuli used for experiments and helped mold blocks of trials for another experiment. The reasoning lab proved to be an exciting way to spend my last two quarters before graduating from the Psychology department at University of California, Los
Angeles. I met and worked with a lot of intelligent members of the reasoning lab and am glad I was given a change to contribute my knowledge and skills.
References


The Effect of Social Exclusion on Social and Cognitive Working Memory Tasks

Caitlin M. Black

Introduction

The human experience is primarily an intellectual one enveloped in a social world. Humans, as intelligent creatures, have interwoven intellectual and social aspects into their lifestyles, and it has been suggested that intelligence is a tool that developed to facilitate social functioning (Baumeister et al., 2002). Interestingly, although research in this area has generally assumed that social intelligence is a derivative of cognitive intelligence, more recent research has found that social intelligence may rely more on at least partially distinct neural mechanisms than general intelligence (Meyer et al., 2012). Thus, an alternative account on the relationship between general and social intelligence is that they are distinct processes that should be differently influenced by various psychological factors. The purpose of this study is to examine whether one psychological factor, rejection sensitivity, corresponds with reduced general cognitive ability but enhanced social cognitive ability.

Past research has shown that the experience of rejection reduces cognitive ability. For example, in a study examining the effects of social rejection and intelligence, Baumeister et al. (2002) illustrated a decline in cognitive intelligence performance after participants experienced a social exclusion manipulation compared to participants who experienced a control, non-exclusion manipulation. Additionally, in the social exclusion condition (relative to a control condition) participants attempted fewer questions during the intelligence test, illustrating that social rejection not only decreases the number of
correct answers, but also the number of attempted answers. Social rejection clearly had an impact on cognitive ability in this study (Baumeister et al., 2002).

Baumeister et al. (2002) suggest that this finding reflects the fact that cognitive capacity (whether general cognitive or social cognitive) relies on one system. Thus, social exclusion, which leads to enhanced rumination about the social environment, directly exhausts cognitive capacity, in turn depleting cognitive resources necessary to perform general cognitive tasks. However, an alternative explanation, supported by neurocognitive research (Meyer et al., 2012), is that the experience of social rejection turns attention toward social cognitive resources supported by a social working memory system and away from general cognitive ability supported by a cognitive working memory system. This account would also predict reduced cognitive ability in response to social exclusion. However, it would also lead to the novel hypothesis that social exclusion leads to enhanced social intelligence.

To test this alternative account, the current study examines the relationship between cognitive working memory and social working memory by measuring participants’ performance (reaction time and accuracy) on social and cognitive working memory tasks. In addition, to examining whether social rejection differentially affects social and general intelligence we measured participants’ sensitivity to social rejection with the rejection sensitivity questionnaire (Downey & Feldman, 1996). If social working memory and cognitive working memory rely on distinct resources, then participants’ tendency to be sensitive to rejection may bias them towards processing social information. Thus, consistent with Baumeister (2002), rejection sensitivity may negatively correlate
with cognitive working memory, but may also positively correlate with social working memory.

Methods

This study includes 52 undergraduate students at the University of California, Los Angeles recruited by the Sona Systems database. After the initial sign-up, participants were sent an email and asked to complete a survey prior to their participation in the study. This survey asked for the names of ten of the participant’s friends; the participant’s rated their friends on 36 trait dimensions on a sliding scale from 1 to 100. A rate of one meant that the friend did not embody the trait at all and 100 meant that the trait was extremely characteristic of the friend. Using a pre-set algorithm, this data was then compiled into four separate MatLab programs, two encompassing the friends trait dimensions in a social memory exercise and two utilizing the names of friends in a cognitive memory exercise. The “Thinking about Traits” task was the social memory task in which participants were given a list of their friend’s names on the first screen, a trait dimension shortly after, and then a true or false question about the order of the friends on the provided trait dimension. The two cognitive working memory tasks were named the “Thinking about Names” task in which participants again were given two to four of their friends’ names, provided with an alphabetizing cue word (ascending for Run 1 and descending for Run 2), and finally asked a true or false question based on the order of the names according to the alphabetizing cue word. The four trials were conducted in two alternating orders. The first order presented the two social working memory tasks first (SWM Run1 and SWM Run2) followed by the cognitive working memory tasks (CWM Run1 and CWM Run2); this was the arrangement
of tasks presented to participants with an even number ID. On the other hand, the order of trials for participants with an odd number ID was reversed: CWM Run1, CWM Run2, SWM Run1, and SWM Run2. Prior to the experiment, the participants were asked to complete a consent form. They were then given verbal instructions on how to complete the task and completed three practice trials utilizing celebrity names. All trials were performed on a research assistant’s laptop. Each participant completed four trials – two “Thinking about Traits” trials and two “Thinking about Names” trials. Reaction times were gathered from both the cognitive working memory tasks and social working memory tasks. After these four trials, the participants were asked to complete a variety of online surveys, including the rejection sensitivity questionnaire. This questionnaire assesses the tendency to anxiously expect, readily perceive, and intensely react to rejection.

Predicted Results

From previous literature, it seems likely that we will find a negative correlation between rejection and cognitive ability and a positive correlation between social rejection and social working memory. Rejection, as a form of social exclusion, should produce similar results from other studies. However, instead of strictly testing cognitive ability, this study presented social aspects, the names of friends to the participant, in both the social and cognitive working memory tasks. Individuals who scored higher on the rejection sensitivity questionnaire should have relatively quicker and more accurate responses on the social working memory task than average. Additionally, these individuals should demonstrate a similar but opposing relationship with the cognitive working memory task. Participants who scored higher on the rejection sensitivity questionnaire should have poorer
performance on the cognitive working memory task than those with average sensitivity to rejection. The discrepancy should be even more evident and in support of our hypothesis, if social and cognitive working memory relied on distinct resources. It is also expected that those with average rejection sensitivity scores should show a difference between the social and cognitive working memory task. If the two working memory systems rely on distinct resources, then there should be an apparent disparity in the response time and accuracy of the participants on the two tasks. Due to the effects of social rejection on cognitive ability, we expect the data to show a separation between the two memory systems.

Conclusion

This study hopefully serves to illustrate a characteristic difference between social working memory and cognitive working memory. The central question here surrounds whether or not these two separate uses of working memory actually function differently and whether or not one typically functions better than the other. If there is significant difference between cognitive working memory and social working memory, future studies could attempt to address the underlying mechanisms of these systems. Cognitive working memory has been extensively studied, yet the system of working memory related to social aspects has not been fully expanded. Future studies may also delve into more understanding of the mechanisms behind why social rejection could enhance social intelligence and impair cognitive ability, especially under which conditions, if any, these results do not occur. Additionally, functional magnetic resonance imaging (fMRI) could be used to compare the brain regions active in social working memory tasks under normal conditions versus active regions under conditions in which the participant has just
experienced social rejection.
References


Predicting Point Light Action during Binocular Rivalry

Chun Ye Eun

Introduction

When a man drives a car on the road, he has to use his all possible perceptions to drive safely. For example, he would use his motion perception and visual perception to see cars driving beside his car and people walking on the street. However people do not simply perceive moving objects when they are driving. They predict actions of others to prevent any possible accident. If so, is perception and prediction occurring separately or together?

According to article “Predicting point-light actions in real-time”, authors said that visual perception was not simply reconstructed from visual input, but was a predictive activity (Markus et al., 2007). This idea has been supported by the results from many previous studies of neuro-imaging. Brain areas, such as MT which processes real motion, were activated when static images with implied motion were presented (Kourtzi and Kanwisher, 2000). Other researchers also found that superior temporal sulcus (STS) coded biological motion when it was implied from static postures (Jellem and Perrett, 2003b). However there was not much evidence found about the timing of these predictive aspects, so a study, which investigated whether the prediction of human actions involves real-time simulation processes, was conducted in the article (Markus et al., 2007).

In order to show a number of human actions to subjects, they used the technique known as point-light animation of biological motion. With point-light animations, the human action is portrayed by small point lights placed on the head and the joints of the body (Randolph, 2007). The great advantage of using point-light actions is that observers can judge sex and the emotional implication of a point-light defined walker when viewing
the animation of the whole body. With this technique, experimenters in the article motion-captured a number of human actions and made them as point-light action animations. During the experiment, subjects watched sequences of point-light actions, followed by an occluder and then a static test posture. After that subjects were asked to judge whether the test posture depicted a correct continuation of the action before the occluder. Occluder time and the movement gap, which is the time between the endpoint of the action and the static test posture, were varied; occluder time (100ms vs. 400ms vs. 700ms) and movement gap (100ms vs. 400ms vs. 700ms). Error rates were lowest when occluder time and movement gap corresponded. However they found that this pattern of results was destroyed when the test postures were flipped over. These findings suggest that action prediction involves a real-time simulation process but this process breaks down when the test posture which observers are not familiar with is presented.

In our experiment, we assumed that observers would predict the correct action continuation even when two conflicting test human figures shown to the observers, one figure to each eye. This unusual experiment setting can be accomplished by using binocular mirror system. During binocular rivalry two images compete for access to consciousness and low-level inhibitory interactions and high-level excitatory influences will promote perceptual grouping and selective attention (Frank, 2006). Therefore our assumption is that action prediction involving a real-time simulation process will work on visual perception during binocular conflict and make observers see the correct continuation human figure more clearly than the other when it is in upright condition.

Method

Participants
The participants in this experiment were 18 UCLA undergraduate students, who were participating in this experiment to get course credit.

**Design**

A 2x3 within-subjects design was used for this experiment. The first independent variable (IV1) was represented by two levels of occluder time; 200ms or 800ms. The second independent variable (IV2) was represented by three levels of matchness of common figure in probe movie before occluder; Upright Matched vs. Upside down vs. Mismatched. The dependent variable (DV) was the percentage of correct responses out of all responses.

**Materials and Apparatus**

The experiment consisted of five parts: a learning session, a practice for block 1, actual experiment of block 1, a practice for block 2, and actual experiment of block 2. Learning session consisted of 6 movie sequences (A pulls B and B resist; A pulls B, B resists, but loses; Scramble for last seat, A loses and stands up; High-five; A sits, B pulls up A; A picks up stool, threatens to strike B) and they were repeated once again. Recognition test consisted of 24 trials that subjects judged whether the figure on the screen was old or new. In the beginning of each real experiment block, ten practice trials were given with upright common figure before occluder. There were two blocks totally, one with 200ms occluder time and the other with 800ms occluder time. Each block consisted of 72 trials and three matchness conditions (Upright vs. Upsidedown vs. Mismatched) equally distributed. For each trial, it consisted of two tasks. One was to judge which human figure depicted the correct action continuation. The secondary task was to choose the facing direction of the green figure.
The correct figure in the 200ms occluder time condition had a movement gap of 200ms, while the incorrect figure in the 200ms occluder time condition had movement gap of 800ms. Vice versa for the 800ms occluder time block. Correct figure was always presented with higher contrast (10 cd/m^2) and another incorrect figure was with lower contrast (3 cd/m^2) and they were counterbalanced for two eyes (left or right) and two colors (red or blue).

The size of the frame was 350 pixels and actor size was 132.9985 pixels in length by 107.7591 pixels in width. The size of each point-light was 5 pixels. The luminance level for common actor (green) was 8 cd/m^2.

Procedure

During the experimental session, a participant sat in front of the computer. First the participant was informed the instruction of the experiment and then placed the face on the chin rest. When the participant looked through the binocular mirror to see the monitor, the instructor turned off the light of the testing room so that the participant could fuse the two images on his/her each eye easily. When the instructor let him/her start, he/she started from the learning session. The participant was required to complete the recognition test with above 75% correct out of 24 trials to continue on prediction rivalry task. After all rivalry task completed, the participant was required to answer to 50 questions about the subjects’ autistic traits, and then the experiment finally ended.

Results

Figure 1 indicates the average accuracy rate. Looking at the pattern of results showed in Figure 1, it presents that average accuracy rate, in general, is higher when
occluder time is 800ms than 200ms. In addition, it presents some level of differences in the average accuracy rate throughout the three different matchness conditions.

To test these effects, the data were analyzed by using a 2 x 3 (Occluder time [200ms, 800ms] x Matchness [Upright, Updown, Mismatch]) within subject analysis of variance (ANOVA), which presented a significant main effect of type of matchness, such that accuracy rate was significantly higher when in Upside-down conditions ($M = 0.75$) than when in Mismatched conditions ($M = 0.71$), regardless of occluder time, $F(1, 18) = 3.342, MSE = 0.024, p = 0.047$. On the other hand a significant main effect of occluder time was not revealed, such that accuracy rate was not much different between time when occluder time was 200ms ($M = 0.68$) and when occluder time was 800ms ($M = 0.78$), regardless of matchness, $F(1, 18) = 1.090, MSE = 0.019, p = 0.311$. Additionally, the interaction between occluder time and matchness was not revealed to be significant, $F(1, 18) = 1.769, MSE = 0.006, p = 0.186$.

**Conclusion and Discussion**

From the results, we only came up with one main effect of matchness. The average accuracy for Upside-down condition was significantly higher than Mismatched condition but the difference between Upright matched and Mismatched was not significant because there was large variance across subjects in the Upright matched condition. These findings are not actually corresponding to our prediction of the experiment. As according to our hypotheses, the average accuracy for Upright matched condition should be the highest but we get Upside-down condition as the highest variable. Since the secondary task for Upside-down condition was more difficult than other two so it might have gathered more
attention on figures than other two. Therefore we should consider this condition as a confounding variable and should find a way to eliminate this attention effect.
References


Figure 1. Effects of occluder time (200ms vs. 800ms) and matchness (Upright vs. Updown vs. Mismatch) on accuracy of predicting human action.