Economic Value and State Representations: Preparations to Decode the Role of the Human Orbitofrontal Cortex

by Riley MacAulay
Advised by Yael Niv
Princeton Neuroscience Institute, Princeton University
2018

Table of Contents:

Table of Figures:

Acknowledgments:

Thank you to my post-docs:

I would like to thank Nico Schuck for his encouragement in bringing me onto this project and for his continued help along the way. Nico's 2016 paper sparked my excitement about this research during my lab interviews, and I was so fortunate to be able to work with him for a semester. Additionally, I am grateful to Nico for providing us with his unpublished fMRI data to include in the present study.

Thank you to Angela Langdon for walking me through this project, fielding my day-to-day questions, working through the behavioral stimulus with me, and for providing comments on my draft. Angela has been an excellent guide for my research experience, and I really appreciate her willingness to help me.

Thank you to Sarah Dubrow for her input on the project and her help with coding and analysis for this thesis. Sarah's patience and impressive email response times made this work much easier.

Thank you to my advisor:

I would like to thank Yael Niv for providing the vision and direction for this thesis and this project. Her guidance and feedback have been essential to the present work. Beyond her duties as an advisor, it has been a pleasure to be a part of Yael's lab. She is an inspiring woman and does a great job empowering the people in her lab and promoting women in STEM. I am so grateful for her help and guidance.

Thank you to my friends and family:

Thank you to family for their interest and support of this project. Thank you to my team for always lifting me up, and to the WCrew freshman – thank you for the thesis snacks. Thank you to the Science Girls for spicing up our summer of research. Thank you to my roommates and Bible Study group for the prayers. Thank you to Emerson Solms for her continued support and friendship. You all have made this year so much more fun.

Abstract:

The orbitofrontal cortex is implicated in state representations, value signaling, and value-based decision making. Previous research on the orbitofrontal cortex (OFC) has suggested that it encodes task-relevant economic values, but it remains unknown whether those values are included as part of the task structure or as a separate entity. Based off a recent finding that OFC encodes a cognitive map of task structure, we predict that relevant economic values will be encoded as part of the task representation, and as such, will be decodable from OFC activity using neuroimaging techniques. In this study, behavioral measures from a complex decision-making task support the given neural hypothesis and prepare future researchers to scan this experiment using fMRI. Assuming these predictions hold true, the common characterization of OFC as a value-signaling region operating under a goods-based model would be challenged and replaced by a new characterization of OFC as a structure representing task space.

Introduction:

How do humans decide between various options? How do we incorporate the value of upcoming rewards into our representation of the task at hand? These questions, and the endeavor of answering them, require us to decipher the role of the orbitofrontal cortex (OFC) – a brain region in ventral frontal cortex involved in learning and decision making. To begin addressing these questions, we start by understanding foundational concepts about the process of learning brought forth by psychologists and early neuroscientists. Rescorla and Wagner proposed that learning only occurs when events violate an animal's expectations (1972). Sutton and Barto developed the concepts of reinforcement learning (RL), or the process of learning what to do and how to map situations to actions in order to maximize a reward signal (1998). Central to RL theories is the concept of the 'reward prediction error' – a signal carried by dopamine neurons that encodes the discrepancy between expected and actual outcomes and can be used to learn the subjective value of different options for different task states (Schultz et al., 1997; Niv, 2009). These task states are integrated and encoded in neural activity as an "abstract representation of the task that describes its underlying structure, the different states of the task, and the links between them" (Wilson et al., 2014). These concepts of learning have provided a foundation for more recent inquiries into the neural basis for learning and decision making.

As RL theory suggests, "value is derived from knowledge of the states that define a task and the transition functions that link them together" (Schoenbaum et al., 2011). Accordingly, OFC is thought to play a role in RL, as it represents information regarding the features of expected outcomes – including the current value of immediately relevant rewards – using a state-based framework. When considered alongside the amygdala, OFC activity has been shown to encode the current value of reward representations attributed to predictive cues and to update the relative reward magnitude more frequently than the amygdala (Gottfried et al., 2003; Saez et al., 2017). Others have suggested that the OFC's role is to guide responses for delayed rewards, with a lesser role in guiding immediate rewards (Roesch et al., 2006). Neural activity in OFC has been shown to modulate with recent reward history such that time-discounted rewards are encoded independently from absolute reward values in OFC (Saez et al., 2017; Roesch et al., 2006). While the exact function of OFC is not yet agreed upon, taken together, these results provide that OFC is involved in representing rewards and outcomes.

The OFC is a highly interconnected region, and its connectivity with the hippocampus has been considered in two key studies. The first of these studies found that OFC links events to reward values while hippocampus links events to their context (Farovik et al., 2015). In particular, neural ensembles in OFC represent distinct value-based schemas and spatial contexts that define the mappings of stimuli to actions with or without rewards. These schemas are suggested to allow for more rapid memory consolidation in neocortex (Tse et al., 2007). The second of these studies found that the hippocampus contributes low-level features of expected outcomes and inferred or abstract properties of structure - such as task state - to OFC encodings (Wikenheiser et al., 2017). These contributions from the hippocampus might aid OFC in its representation of a cognitive map of task space (Wilson et al., 2014, Schuck et al., 2016). While OFC itself does not support cognitive functions such as response inhibition, credit assignment, prediction errors or value signals, many of these functions are supported by the cognitive map represented in OFC (Stalnaker et al., 2015). These studies provide a wider view of OFC's role as an interconnected region, while others take closer look at functional subregions in OFC.

Many studies have shown that OFC is not a functionally, or even structurally, homogeneous region. In a lesion study, Noonan et al. found that monkey analogues to medial OFC (mOFC) and lateral OFC (lOFC) have distinct functions (2010, 2012). In particular, mOFC is involved with reward-guided decision making while lOFC is concerned with reward-guided learning (2010). In their 2012 meta-analysis, Noonan's group looked at mOFC, lOFC and ventromedial prefrontal cortex (vmPFC), and suggested that lOFC is responsible for credit assignment, while vmPFC and mOFC are responsible for evaluation and choice maintenance over successive decisions, in addition to reward-guided decision making. Choice perseveration has long been considered a major effect of lesions to either OFC subregion, but more recent findings show that lOFC lesions cause animals to stop choosing previously successful options instead of perseverating on those options (Noonan et al., 2012; Rudebeck & Murray, 2011). In another study, mOFC was found to be critical for discriminating ambiguous states by abstracting away unnecessary information, adding unobservable information, and crafting a task state (Bradfield et al., 2013). Other experimenters have divided the OFC into a medial-caudal region and a rostro-lateral region, which are thought to be involved in reward-identity representations independent and dependent of predictive stimuli respectively (Klein-Flügge et al., 2013).

OFC has also been shown to play a computational role in value-based decision making. OFC activation has been found to reflect the difference in subjective value between available options (FitzGerald et al., 2009). Further, OFC activation alternates between states associated with the value of two alternatives, suggesting that subjective decision-making involves dynamic activation of OFC states (Rich & Wallis, 2016). A functional imaging study has shown that OFC computes Bayesian log-transformed posterior distributions over latent - or unobservable - causes (Chan et al., 2016).

These findings have been complimented by a body of work suggesting that OFC represents economic values of offered and chosen goods (Padoa-Schioppa and Assad, 2006). In their foundational study, the experimenters found that offer value and chosen value were amongst the three variables that accounted for almost 80% of OFC neural responses. This finding was the basis for their theory that OFC neurons operate under a goods-based model, such that economic choice is a choice between goods, as opposed to a choice between actions (Padoa-Schioppa and Assad, 2006; Padoa-Schioppa, 2011). They found that, in OFC, values of different goods are individually computed at the time of choice and are independent of the sensorimotor contingencies of the choice

(Padoa-Schioppa, 2011). Upon further inspection, the value representations in OFC adapt to different behavioral conditions and to the range of available values in a given condition (2009). Their results suggest that OFC instantaneously provides preference transitivity in the form of menu invariance, as well as computational efficiency in the form of range adaptation (Padoa-Schioppa, 2009; Padoa-Schioppa and Assad, 2008). In their 2017 study, Padoa-Schioppa and Conen make a case for a neural circuit in OFC as the locus for economic decisions. These conclusions from Padoa-Schioppa and colleagues have culminated in a major theory in the debate on OFC function, but it is not the only theory.

An influential alternative theory proposes that the OFC represents a cognitive map of the task state representation (Wilson et al., 2014, Schuck et al., 2016, Schuck et al., 2017). In agreement with Bradfield et al. (2013), Wilson and colleagues suggested that OFC is necessary for representing task states in partially-observable scenarios. Partially-observable scenarios are those in which components of states are not perceptually distinct from one another or are not readily observable to the subject at the moment of judgment. From their meta-analysis, Wilson and colleagues also determined that the state representations encoded in OFC are memory based, and that OFC does not directly encode expected value, but rather represents the partially-observable task states which are then used for calculating or learning values. Critically, OFC only encodes the partially-observable task states and components that are relevant to the task at hand.

To test Wilson et al.'s theory of OFC, Schuck et al. (2016) aimed to decode partially-observable task states from human OFC using functional magnetic resonance imaging (fMRI) during a decision making task. This study found that several task-relevant state components could be reliably decoded from OFC activity patterns, including those that were unobservable in the current trial. They further found that different neural activation patterns in OFC were related to the identity of different task states, and that more similar task states had more similar neural activation patterns. From these results, Schuck et al. proposed a new interpretation of Padoa-Schioppa and Assad's (2006) goods-based model: the value signals in OFC comes from the relationship between expected rewards and partially-observable components of the task, rendering relevant economic values part of the task state. The present study seeks to support this proposal in humans through neuroimaging techniques.

In this study, we aim to address the following question: how does the human OFC encode economic values - or rewards - relative to the task state representation during decision making? We hypothesize that economic values will be encoded in OFC activity as part of a task-state representation only when those values are relevant for the decision making task at hand. To test our hypothesis, will use a decision making task with an ambiguous behavioral stimulus and block-wise relevant or irrelevant rewards. While is not in the scope of this study to scan the human OFC using fMRI, it is our intent to establish behavioral trends that support our neural hypothesis and prepare a stimulus that is ready to be scanned. In a future neuroimaging study with this stimulus, we would expect to be able to decode the task state representations from OFC activity during the task and find that it contains task-relevant economic values.

Chapter 1: General Methods

This study used a behavioral stimulus with a 16-state space (Figure 1), adapted from Schuck et al. (2016). Each state contained four components: current image category, current image age, previous image category, and previous image age. The categories consisted of face images and house images. All images were equally sized and shown in grayscale. Participants were tasked with judging the images as either old or young, indicating their judgment with keypresses. Of the four state components, three were partially-observable and the fourth – current image age – was fully observable. In order to track which category to judge, participants needed to keep these four components in working memory. In future imaging research with this task, we expect that the three partially-observable components of each state, in addition to the task structure, will be decodable from human OFC using fMRI. The fourth component – current age – was not decoded above chance in Schuck et al. (2016), presumably due to its observable nature, so we would expect that finding to carry over into future work with this task.

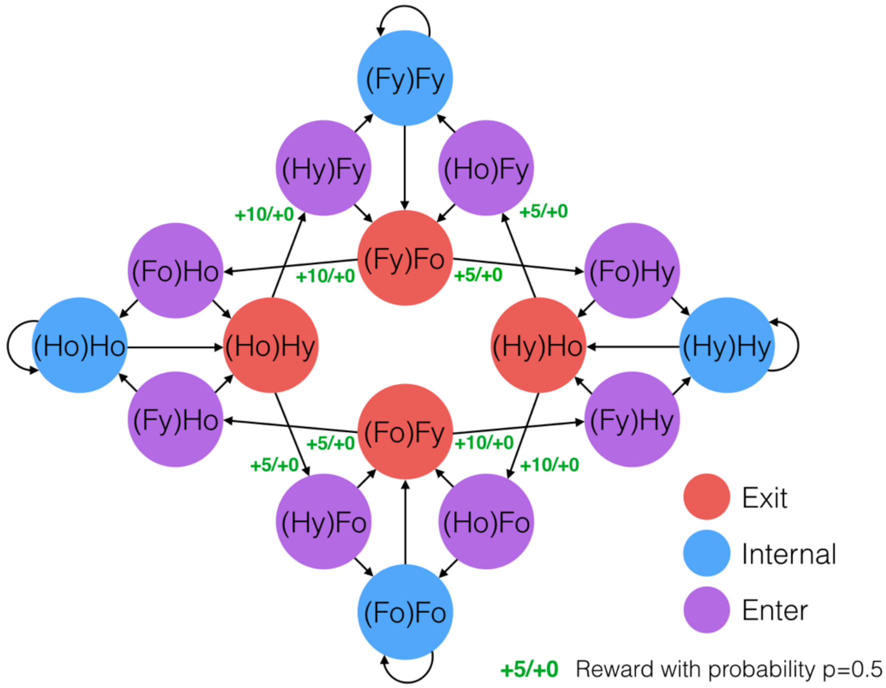


Figure 1: Task State Space – Each circle represents a single state of the 16-state space. Capital letters indicate category (face or house), lower-case letter indicate age (old or young). The category and age inside the parentheses reflect the previous state, which impacts the judgment in the current state. The category and age outside the parentheses reflect the current state. The +5 and +10 rewards are displayed on the screen between Exit and Enter states, while the screen remains blank during transition for the +0 rewards.

The task was structured into blocks (Figure 2a) and mini-blocks (Figure 2b). The blocks contained 20 (training) or 49 (experimental) trials with a consistent reward schema. The mini-blocks contained a series of stimuli of the same category – faces or houses – and were between two and six trials in length. In Block 0 (training), participants were acquainted with the stimuli images and practiced making the key-press age judgments. They viewed single images of either a face or a house assembled into mini-blocks of the same category. After 20 trials, participants with an error rate under 20% continued on to the experimental blocks, and participants with an error rate above 20% looped back for another round of training trials.

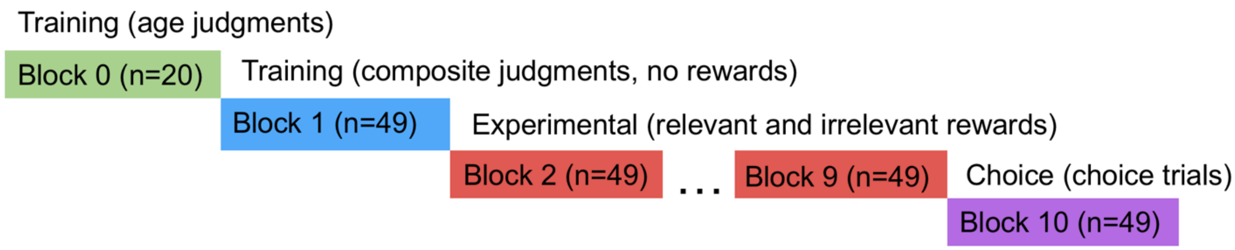


Figure 2a: Task Block Structure – eleven blocks of the behavioral task. Block 0 is a training blocks with single images. Block 1 is a training block comprised of composite images. Blocks 2-9 are the experimental blocks with composite images and rewards during transitions. Block 10 uses choice probes during transitions.

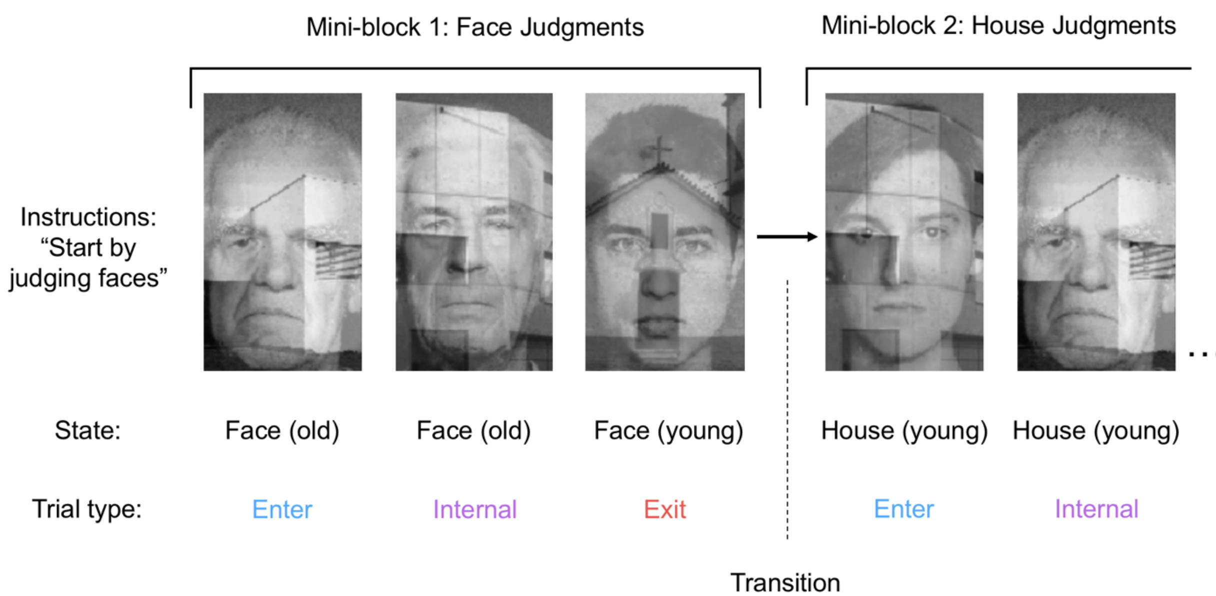


Figure 2b: Task Mini-block Structure – a sample series of trials at the beginning of an experimental block using composite images from the task.

For the subsequent blocks, the stimuli changed from single images of either a face or a house to composite images of a face and a house superimposed on top of each other. With the two-dimensional stimuli, subjects continued to make age judgments, but it was essential that they tracked which category they were judging. At the beginning of each block, an instruction screen indicated which category to judge first. As is shown in Figure 2b, if the instructions read "start by judging faces", then the participant would make their age judgment on the face dimension of the composite image. As long as the judgment stayed the same, the participant would continue judging the same category. If the first judgment was "old", then the participant would judge faces until they were presented with a young face. After indicating their switched judgment with a key press response, the participant would switch categories to judge houses. If the first house presented after the category switch – or transition – was young, then they would keep judging houses until they saw an old house, at which point they would switch back to judging faces. Put more simply, the task rule was to continue judging a category until the judgment changed, then to switch to judge the other category. It is important to note that a participant ignores the category dimension that she is not judging. Accordingly, the dimension not being judged is referred to as the irrelevant dimension, while the category being judged assumes the relevant dimension.

A mini-block, as shown in Figure 2b, contained Enter, Internal, and Exit trials with the same relevant dimension. The Exit trials were that in which the participant would switch their age judgment. The Enter trials were the first trials of the relevant dimension and could be either at the beginning of the task or following an Exit trial from a previous mini-block. The trials between the Enter and Exit trials were called Internal trials.

In Block 1, participants practiced implementing the task rule as they judged the composite images. As in all blocks, incorrect age judgments were followed by an error message that reminded the participant which category to judge next. If an error was made one trial before an Exit trial, the error message indicated the switch in the trial after next. This block did not have any rewards. In the following experimental blocks (Blocks 2-9), rewards were shown on the screen during 50% of mini-block transitions. In half the blocks, the reward points were indicative of the age judgment of the incoming Enter trial (CLUE blocks). In the other half of the blocks, the reward points had no relationship with the age of the incoming Enter trial (NO CLUE blocks). In order to collect the points, participants had to correctly judge the age of the Enter trial. The reward point values were +5 and +10. The indicative (relevant) rewards were associated with same-age transitions – in which the age of the Exit trial matched the age of the Enter trail – or different-age transitions – in which the age of the Exit and Enter trials did not match. Regardless of the relevancy of the rewards, the reward point values were evenly distributed across each block, and there were equal numbers of high- and low-value rewards shown. The associations between the high-value (+10) and low-value (+5) rewards and the same- or different-age transitions were counterbalanced across participants, as were the order of relevant and irrelevant blocks. In older versions of the task, the reward associations were probabilistic ($p = 0.80$), but in the final version, the associations were deterministic, yet shown in only half the transitions between mini-blocks.

The choice block - Block 10 - used a modified structure featuring choice trials on 50% of the Enter trials. On a choice trial, participants were shown two composite images side by side and asked which image they would rather see next. After eight blocks of experience with the reward relationships, the choice probes were aimed to determine if subjects understood which Enter trial image would correspond with the high-value reward. For example, a subject for whom the high-value reward was associated with different-age transitions would presumably choose the choice trial option that had a different age from that of the Exit trial. Like the previous experimental blocks, Block 10 contained 49 trials, but it uniquely included approximately 4.88 choice trials.

Chapter 2: Experiment 1

Introduction

To begin to address the question of how the human OFC encodes economic values with regard to the state representation, we started with a low-level behavioral question. Using the task described above, we sought to understand how the reward cues during mini-block transitions impacted participant reaction times (RT's). This inquiry, broadly referred to as the effect of reward cues on screen, was carried out as a behavioral pilot experiment. We expected that reward cues would decrease RT's relative to unrewarded Enter trials in the relevant blocks. In the irrelevant blocks, we expected that the reward cues would not predict the age of the incoming Enter state, such that the irrelevant rewarded RT's would be consistent with the irrelevant unrewarded RT's. Additionally, we predicted that both reward values (+5 and +10) in the relevant condition would evoke RT's that would be similarly faster than the baseline established by the unrewarded Enter trials.

Methods

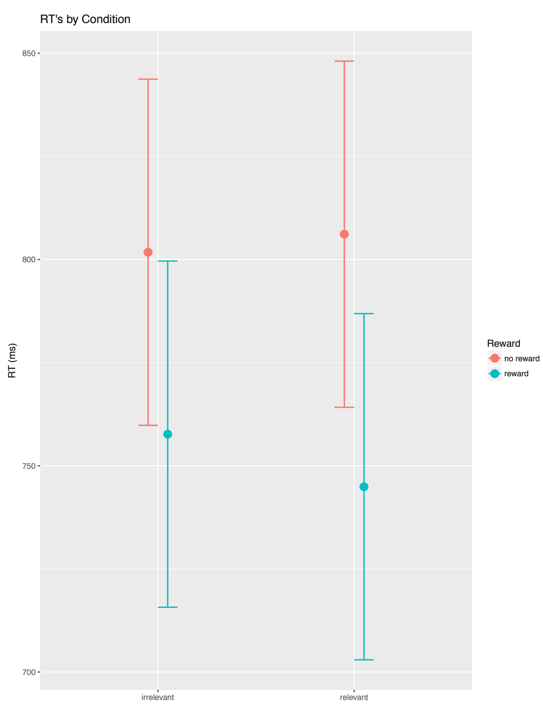
The task was simplified from its original form in Schuck et al. (2016) in two ways to better address the reward cueing effect. First, the relationship between the reward value and the age of the Enter trail following reward presentation was explicitly stated in on-screen instructions. For relevant blocks, participants were informed at the start of the block which transition type – same-age or different-age – corresponded with the high-value reward (+10) and which with the low-value reward (+5). Similarly, for the irrelevant blocks, participants were informed of the random relationship between reward value and age of the incoming Enter trial. This explicit instruction of the reward relationships replaced an implicit learning schema in which participants were expected to derive the relationships themselves over the course of the experiment using probabilistic rewards. We expected that this simplification would reduce the load on working memory and require less critical thinking to determine the relationships on the part of the subjects.

Second, this modified version of the Schuck et al. (2016) task used a fixed key mapping schema instead of a random alternating key mapping schema. Whereas the original task employed an age-response key mapping that was re-drawn on each trial, with this change, participants always knew which key corresponded to which age judgment for the duration of the experiment. The key mapping was counterbalanced across participants. The fixed mapping also allowed for participants' predictions about the age of the next trial to be more salient in their RT's. In particular, during a transition, an engaged participant would know which key to press to make the correct age judgment as soon as she saw a relevant reward. Considering this, we expected that if subjects were tracking the task structure, they would have faster RT's on an Enter trial following a relevant reward. Additionally, the fixed mapping minimized the complexity of the task by removing the variable of the mapping.

To test our predictions about the effect of a reward cue on screen, we collected behavioral data from Princeton University undergraduate participants (n = 9, 5 females). The participants were 19.89 years old (sd = 1.37 years) on average, and were compensated with a $12 base pay, in addition to a performance bonus scaled to the number of reward points collected during the experiment. The bonus conversion rate was 1 point to $0.01, and the bonuses ranged from $3.75 to $4.50. No participants were initially excluded.

Results

Participants performed the task with high accuracy, with an average error rate of 4.89% (sd = 2.26%). The primary measure of this experiment was participant RT's, while error rate was also considered. Across the experimental blocks, the group data revealed a small reward effect for the relevant rewarded trials over the relevant unrewarded trials, yet there was a similar RT differential for the irrelevant rewarded trials and the unrewarded trials (Figure 3a). When broken down by reward value (Figure 3b), we saw that the RT's were clustered for irrelevant trials, with the +10 RT's being slightly faster than the unrewarded RT's and +5 RT's. In the relevant condition, the +5 RT's were almost significantly faster than both the +10 and the unrewarded RT's (Welch Two Sample T-Test, $p = 0.0527$). This was a surprising result as only one reward value evoked a reward effect while the other did not. In fact, the relevant high-value rewards were the slowest on average for any condition in the experiment.



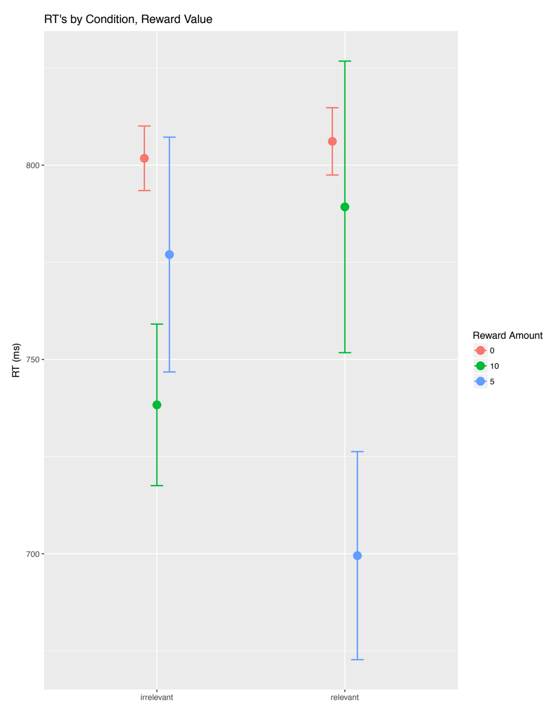
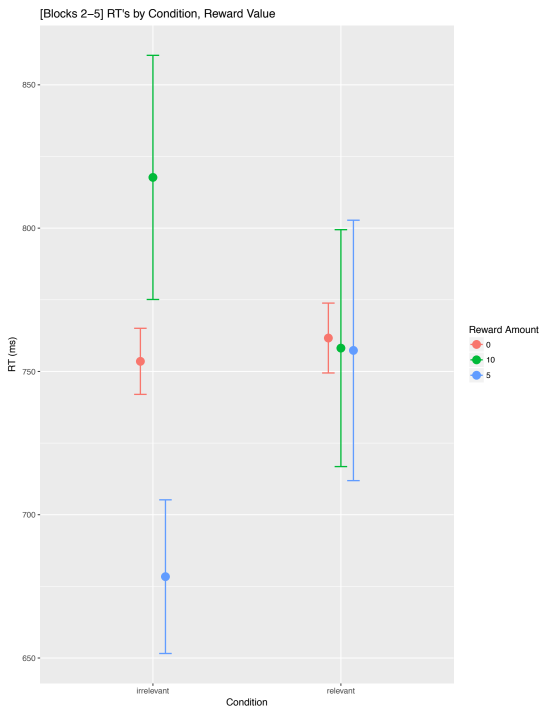


Figure 3a: RT's by Condition – (left) group RT averages over experimental blocks 2-9. Error bars represent standard error computed across reward value conditions (reward vs. no reward). While the relevant rewarded RT's are faster than the other RT subdivisions, there is also a reward effect in the irrelevant condition. **Figure 3b – RT's by Condition, Reward Value** – (right) RT's by Condition, Reward Value – group RT averages over experimental blocks, broken down by condition and reward value. Error bars represent standard error across participants. There is a clear reward effect in relevant +5 RT's, but not for relevant +10 RT's.

While the existence of a reward effect, even for one reward value, suggested our hypothesis could hold water, we wanted to understand the basis of the RT discrepancy between +5 and +10 trials. To start, we broke the experiment into Blocks 2-5 (FirstHalf) and Blocks 6-9 (LastHalf). Using the same analyses, FirstHalf showed a reversed trend from what was expected and what was seen from the whole-experimental data (Figure 3c). The relevant RT's were tightly clustered, and the +5 irrelevant RT's were significantly faster than the +10 RT's ($p = 0.006671$), which were slower than baseline. Further analysis did not reveal a clear cause for this reversed trend, but it is possible that participants were paying more attention to the relevant rewards – accounting for the clustering – but were not yet comfortable with the reward-transition relationships to predict the age of the incoming Enter trial. Perhaps on the irrelevant trials, with less focus, the participants had a wider range of RT's because they knew they had no relationship to the Enter trial age.



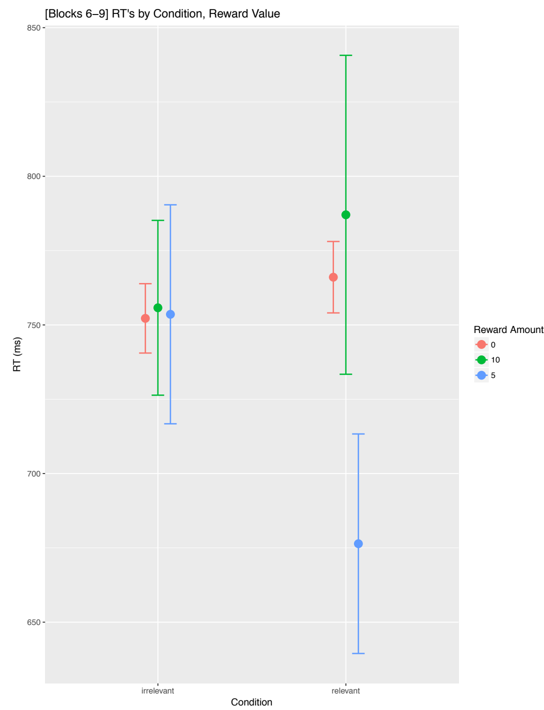


Figure 3c: [Blocks 2-5] RT's by Condition, Reward Value – (left) FirstHalf group RT averages show reversed trend from Figure 3b, with a split reward effect in the irrelevant condition and clustered relevant RT's. **Figure 3d: [Blocks 6-9] RT's by Condition, Reward Value**– (right) LastHalf group RT averages show reward effect for +5 only.

In the LastHalf data, the trends matched the whole-experiment data (Figure 3d). We saw tighter clustering for the irrelevant trials, which matched our prediction, but again there was only a reward effect for the relevant +5 RT's as compared to baseline ($p = 0.02363$). The +10 RT's were, on average, above baseline. We postulated that the predicted behavior became more visible in the later blocks because participants learned to use the relevant reward cues to predict the next judgment over the course of the experiment. Under this theory, they also would have learned not to pay attention to the irrelevant reward cues, which would account for the clustering at baseline for that condition in the LastHalf data.

To measure whether participants were learning over the course of the experiment, we looked at the average error rates over the 9 experimental blocks (Figure 4a). By comparing the relevant condition error rates to the baseline error rates established by the irrelevant condition, we saw that there was a higher error rate on average for the second half of the experimental blocks than for the first half (3.57% for blocks 2-5; 3.94% for blocks 6-9; p = 0.5427). These error rate trends do not support the idea that participants learned over the experimental blocks. However, it is important to note that performance was high with this modified task, making the error rates relatively low across the experiment, suggesting that learning was not a crucial element of this experiment. Participants are nearing a performance ceiling, so the small change is error rates across the experiment - in either direction - were unlikely to account for the RT behavior between the relevant reward values.

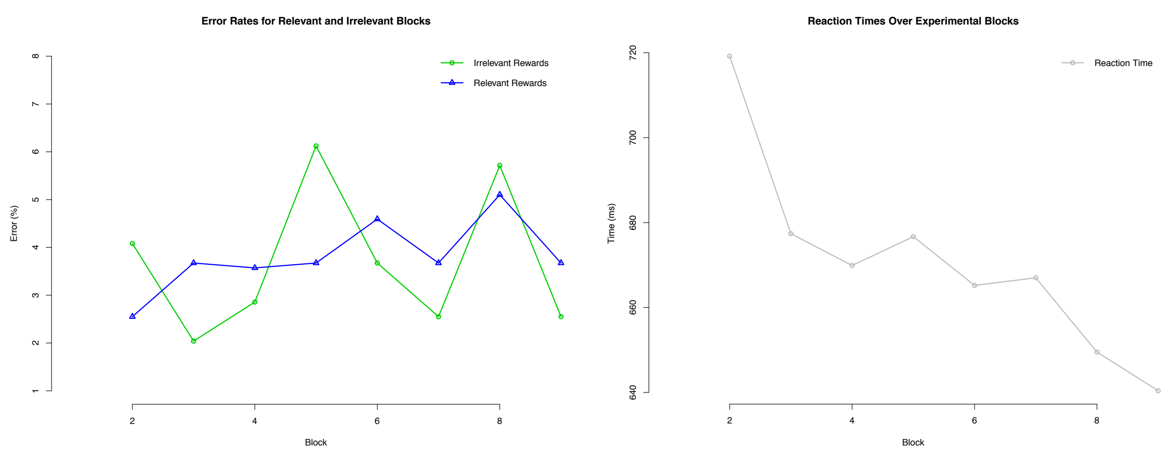


Figure 4a: Error Rates for Relevant and Irrelevant Blocks – (left) group error rate averages over experimental blocks 2-9 by condition. The average error rate for the last four blocks of the experiment were insignificantly higher than that of the first four blocks. **Figure 4b: Reaction Times Over Experimental Blocks** – (right) group RT averages over experimental blocks 2-9. This downward trend in RT's suggests learning over the course of the experiment.

In addition to error rates, we also considered the average RT's over all the experimental blocks (Figure 4b). This analysis showed significant change over the eight experimental blocks (p = 0.03686 for Block 2 to Block 9). Unlike the ceiling effect found in the error rate analysis, the decrease in RT's across blocks provided evidence for learning during the experiment. This could explain the reversal of RT trends from the first four blocks to the last four blocks, yet it does not account for the difference in reward effects for the relevant reward values. To better address that particular concern, we looked next at subject-by-subject data.

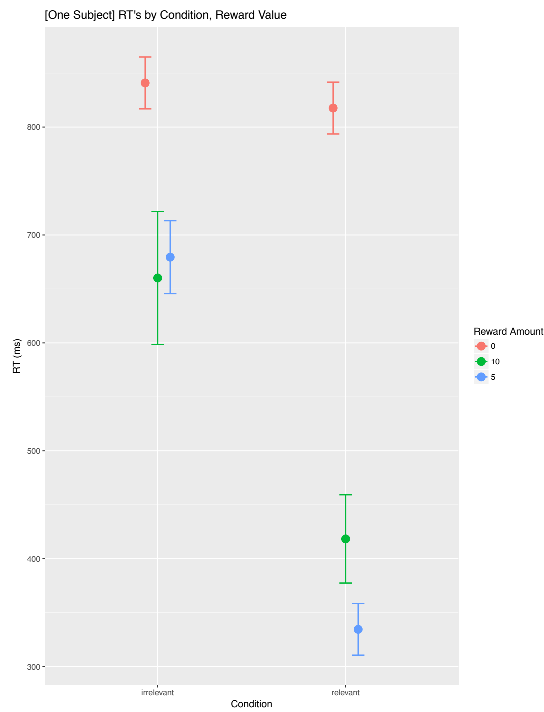
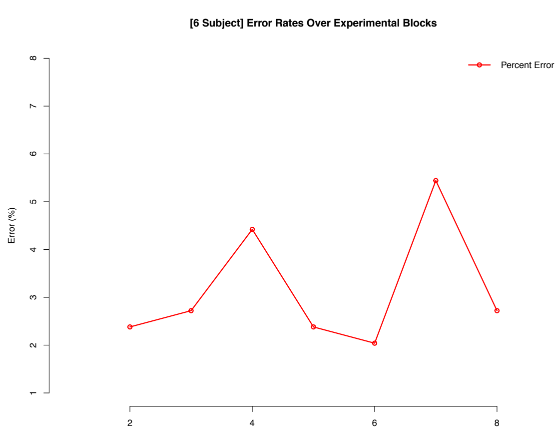


Figure 5a: [One Subject] RT's by Condition, Reward Value - Exemplary subject RT data across experimental blocks 2-9 with clear reward effect in relevant condition for both reward values. The irrelevant rewarded RT's are below baseline, but still significantly slower than the relevant rewarded RT's ($p = 1.04e\text{-}08$ for +5's, $p = 0.003748$ for +10's)

Figure 5a shows RT data from a single subject that closely resembles the expected behavioral results. While three of the nine participants produced results that exhibited a clear reward effect for both relevant reward values, three other participants had little to no reward effect, and three more participants had one relevant reward value with RT's well above baseline. This variety in the data calls into question the validity of the hypothesis, but the notion that one third of subjects were able to make judgment predictions using the reward cues gave us reason to believe that the reward signal we were seeking with this task was present and needed to be drawn out more clearly. After excluding the three participants with one reward value above baseline, RT and error rate trends were re-analyzed. Figure 5b shows the error rates over the experimental blocks with three subjects excluded. As compared to the relevant condition in Figure 4a, the peaks were distributed differently across the blocks, but were not significantly different over the eight blocks (p = 0.3123). Perhaps a more informative measure, Figure 5c exhibits RT's over the experimental blocks with three subjects excluded. This exclusion does not significantly affect the overall negative trend of the RT's over the course of the experiment, as compared to Figure 4b (p = 0.4134). Figure 5d shows the average RT analysis by reward value and condition after the exclusion, which appears more similar to the expected behavioral trends. The irrelevant RT's were more closely clustered than in the full group data, and both relevant reward values show a reward effect of faster RT's – although the magnitudes are disparate. By removing subjects from the dataset that were not tracking the task structure, some of the unexpected behavioral results – particularly the opposite directions of the relevant reward effects – were accounted for. However, the difference in reward effect magnitude remains unaccounted for.



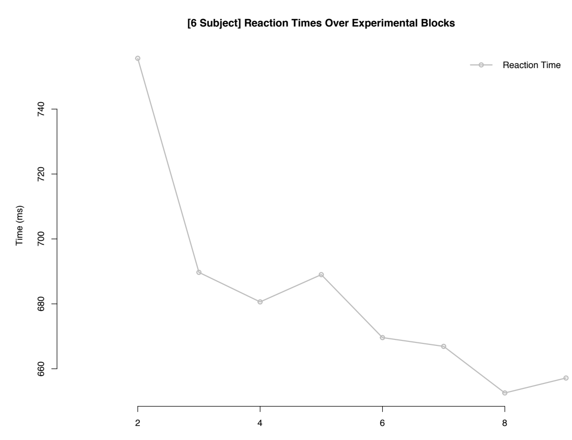


Figure 5b: [6 Subject] Error Rates Over Experimental Blocks – (left) group average error rates over experimental blocks with 3 subjects excluded. **Figure 5c: [6 Subject] Reaction Times Over Experimental Blocks** – (right) group RT averages over

experimental blocks with 3 subjects excluded. Neither figure is significantly different with the exclusion of three subjects.

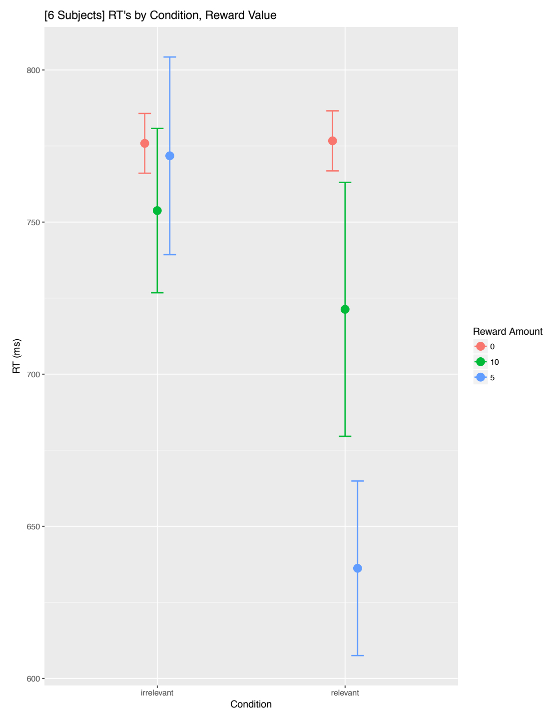


Figure 5d: [6 Subject] RT's by Condition, Reward Value – group RT averages after 3 exclusions by condition and reward. This data shows reward effect in both relevant reward values.

Participants were evenly distributed between the two reward-transition relationships. For different-age transitions paired with the high-value reward (+10), the relationship was called HVDifferent. For same-age transitions paired with the high-value reward (+10), the relationship was called HVSame. Importantly, in the above exclusion of participants, equal numbers of each reward-transition relationship were preserved. To address the reward effect magnitude phenomena, we proposed that one of these relationships could have been more challenging to learn and implement than the other. In particular, we predicted that a task rule of similarity would be computationally easier to learn than a task rule of difference, as switching age judgments in the transition requires more cognitive power than continuing with the same age judgment.

Accordingly, we expected that the HVDifferent condition would be harder to learn considering the judgment switch.

To test this theory, we split the participants (n=9) by their reward transition relationships - HVDifferent and HVSame - and compared RT's between the groups. Figure 6a shows the Enter trial RT's following relevant high-value rewards by block for the HVDifferent (triangles) and HVSame (circles) groups. A comparison of the two revealed that the reward transition relationship did not significantly impact the relevant high-value rewarded Enter trial RT's ($p = 0.2066$).

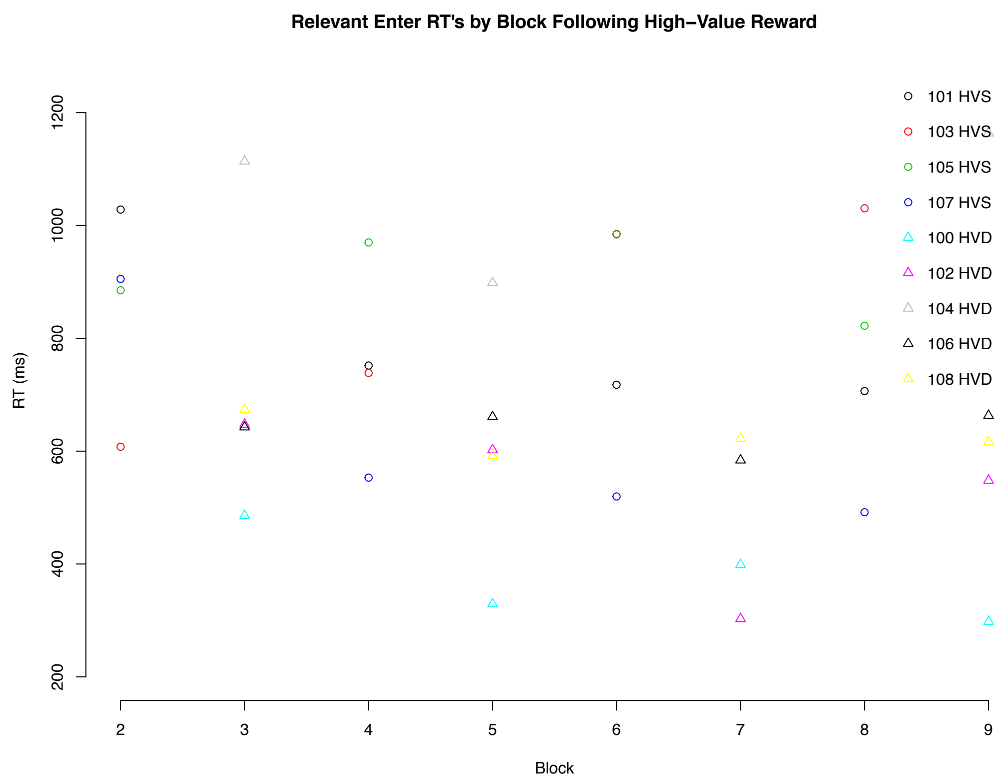


Figure 6a: Relevant Enter RT's by Block Following High-Value Reward – HVDifferent (HVD – triangles) and HVSame (HVS – circles) Enter RT's after high-value (+10) reward on relevant blocks. The two reward transition relationships do not elicit significantly different RT's.

We expected that we might find slower RT's for the HVDifferent relationship group, but this analysis suggests that one relationship is not significantly harder to learn than the other. Next, we considered the Enter trial RT's following low-value (+5) rewards in the relevant condition for both HVDifferent and HVSame relationships (Figure 6b). Again, the reward transition relationships do not produce significantly different RT's between the groups (p = 0.178).

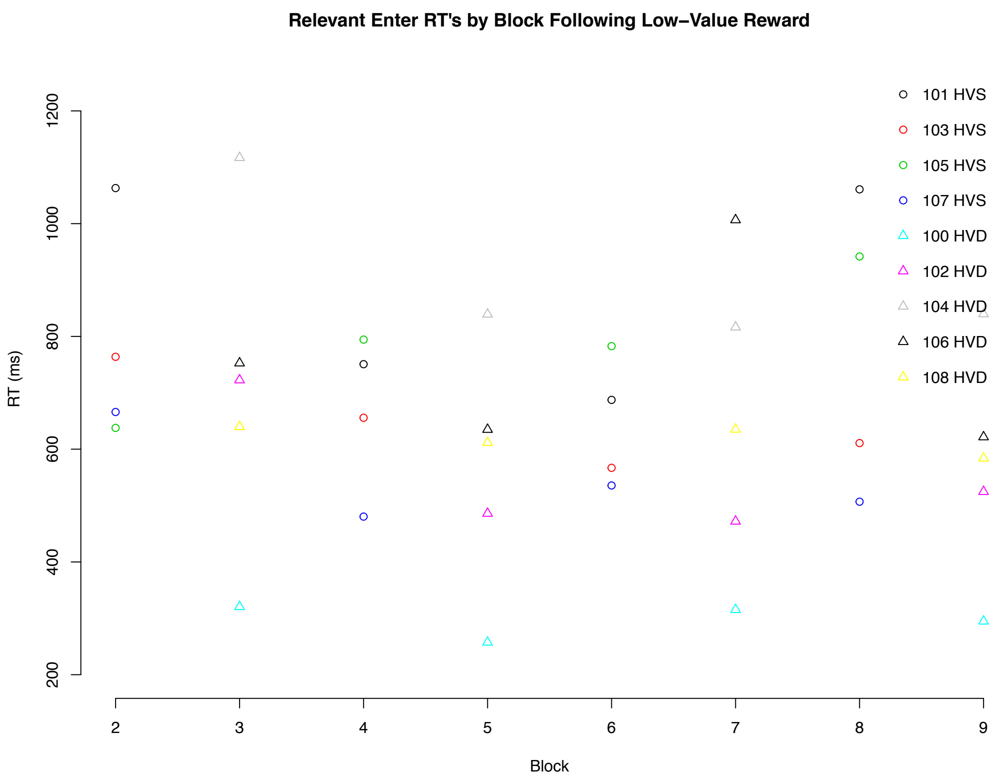


Figure 6b: Relevant Enter RT's by Block Following Low-Value Reward – HVDifferent (HVD – triangles) and HVSame (HVS – circles) Enter RT's after low-value (+5) reward on relevant blocks. The two reward transition relationships do not elicit significantly different RT's.

These quantitative analyses do not produce any compelling evidence for one relationship being more challenging to learn than the other relationship. In a final effort to understand why the relevant +5 Enter RT's were uncharacteristically faster than the relevant +10 Enter RT's, we compared the relevant Enter RT's to see if the low-value (+5) RT's for both relationships were faster than the high-value (+10) RT's for both relationships. Surprisingly, this relationship was not significant (p = 0.2863).

Additionally, the relationship between the HVSame and HVDifferent groups when considering both high- and low-value rewards was not significant ($p = 0.2177$). Lastly, we considered if the act of switching judgments produced significantly slower RT's than the act of staying in the same age judgment. For this analysis, we compiled the RT's for the HVSame group following a high-value reward and the HVDifferent group following a low-value reward, as in both of these cases, participants stayed with the same judgment. These RT's were compared to RT's from the HVDifferent following a high-value reward and the HVSame group following a low-value reward, in which participants switched their age judgment. The switching versus staying paradigm did not significantly account for the differences in RT's ($p = 0.617$).

To conclude these analyses, the singular reward effect for relevant +5 rewards cannot be accounted for by the reward transition relationships, exclusion of confused participants, or learning over the experimental blocks. Due to the small sample size of this experiment ($n = 9$), we attributed this trend to behavioral noise and sought to collect more data in hopes to see a stronger behavioral signal reflecting a reward effect for both relevant reward values.

Chapter 3: Experiment 2

Introduction

Our inquiry into the behavioral effect of reward cues on screen provided some support for our theory that participants were using the reward cues to predict the age on the incoming Enter trial. As the data from Experiment 1 was noisy, we felt it necessary to run another experiment to increase the sample size, while also changing the task slightly to answer another important question about the reward cues. In this experiment, we asked whether the reward cues on screen during a transition were used to predict the age of the incoming Enter trial or to indicate a category switch. Experiment 1 was based on the principle that reward cues, when present, would help participants to anticipate the next age judgment. Experiment 2 provided a necessary sanity check of that assumption. This step was necessitated by the possibility that participants were not fully tracking the task structure, such that the presence of a reward cue on screen would indicate a category switch, allowing them to make a faster age judgment on the subsequent trial. We predicted that the reward cues would provide information about the next age judgment in relevant experimental blocks instead of indicating a category switch.

Methods

To test this hypothesis, no-reward outcomes were added to all trials without reward cues in the task adapted from Schuck et al. (2016). These no-reward outcomes were in the form of a cyan-colored "--" cue on the screen following a trial. They indicated that there was no reward for the given trial. As 50% of Exit trials were followed by reward cues, these no-reward cues were presented in the remaining 50% of

Exit transitions, and after all other trials in experimental blocks 2-9, such that every trial in these blocks had some outcome. The rest of the task structure was consistent with Experiment 1.

The addition of no-reward cues minimized the possibility that subjects could use the reward cues as an indicator of a category switch, because all trials were followed by cues, not just the trials in which the category changed. With regard to Exit transitions, the addition of no-reward cues equalized the information provided in each trial, preventing the Exit trials from being visually distinct, as they were in Experiment 1. This allowed us to compare the RT's for Enter trials following no-reward cues and reward cues without any confounds of additional information from the task state. We expected that, in the relevant condition, the Enter RT's following a no-reward cue would be approximately equivalent to the no-cue Enter RT's from Experiment 1. Beyond that, we expected that the Enter RT's following a relevant reward cue would be significantly faster than the relevant no-reward Enter RT's. These forecasted behavioral results would support our hypothesis that participants used reward cues to predict the age of the incoming Enter trial. However, if we observed that the RT's following no-reward cues were statistically similar to the RT's following reward cues, this would support the conclusion that the reward cues primarily indicated a category switch.

In this experiment, Princeton University undergraduates and community members ($n = 10$, 7 female) completed the task for monetary compensation. The subject group had an average age of 21.89 years ($sd = 5.16$ years). They were paid $12.00 plus an addition performance bonus of $0.01 for each 1 point collected in the experiment. As before, in order to collect the reward points, participants needed to correctly judge the age of the Enter trial following the reward. The bonuses ranged from $2.75 to $4.55. On average, the group had an error rate of 4.99% over experimental Blocks 2-9 (sd =

0.23%). The percentage of missed trials – or trials in which no response was made in the requisite 2.5 second response period – was 0.38% over the experimental blocks.

Results

The experiment primarily addressed how participants used reward cues. With the data from Experiment 1 as a baseline comparison, we posited that the reward cues would either be used to predict the age of the incoming Enter trial, or to indicate a category switch between mini-blocks. The group data across the relevant experimental blocks revealed faster RT's for the Enter trials following a reward cue, as compared to the Enter trials following a no-reward cue (Figure 7). Consistent with our hypothesis, the no-reward RT's were significantly slower than the rewarded RT's in the relevant condition of Experiment 2 ($p = 3.286e\text{-}14$). When compared to Experiment 1 – which differed only in terms of the no-reward cues – the rewarded relevant RT's between experiments one and two were comparable, and the no-reward RT's were comparable with the no-cue RT's. From this data, we can infer that participants were not using the reward cues for information about a category switch because the no-reward condition elicited comparable RT's to the no-cue condition from the first experiment, but without the transition-indicating capability. This supports the premise on which the first experiment was founded – that participants used the reward cues to predict the age of the incoming trial. This result suggests that the noisy data from Experiment 1 might reveal the expected trends. A larger sample size is needed to establish the expected trend of relevant reward effects.

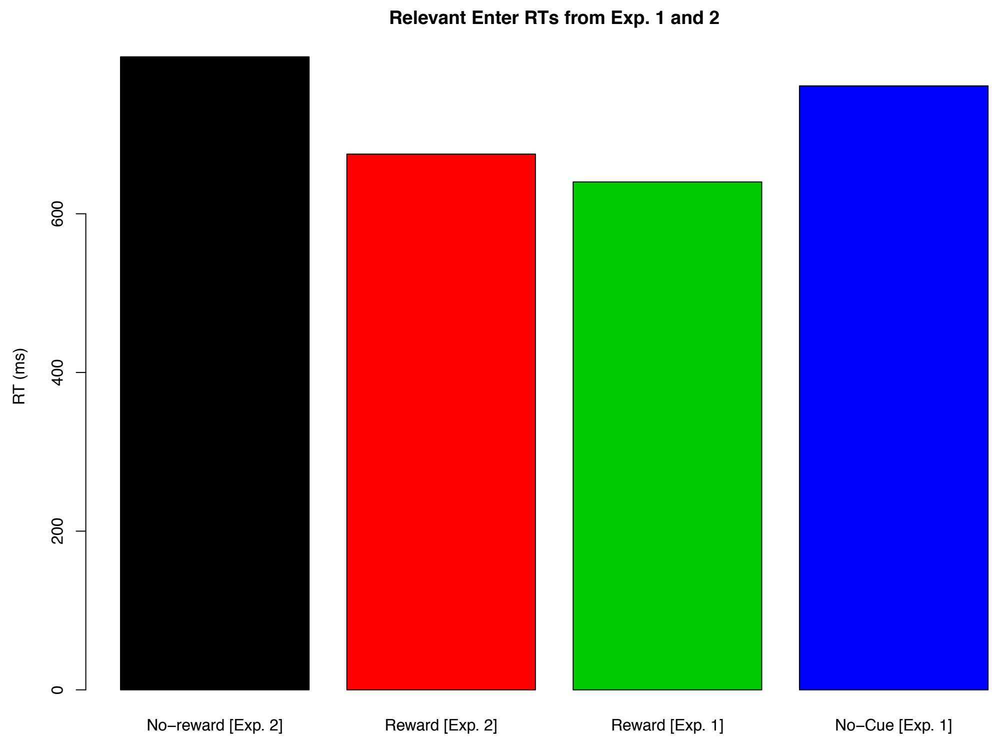


Figure 7: Relevant Enter RT's from Exp. 1 and 2 – Comparing RT's from Experiment 2's Enter trials following a no-reward "--" cue and following a reward cue to Experiment 1's Enter RT's following a reward cue and no cue. The comparable RT's between no-reward and no-cue, as well as the faster RT's following reward cues, suggest that participants did not use the reward cues to indicate a category switch and instead used them to predict the incoming age judgment.

While the primary aim of this experiment was to understand the role of the reward cue in Enter trial RT's, the high degree of similarity between these tasks allowed for this data to compliment Experiment 1's RT trends. We analyzed the RT's and error rates for this experiment as was done in for the previous experiment. To start, we looked at the behavioral effect of a cue on screen. The average group data showed that trials following reward cues had faster RT's in both the relevant and irrelevant conditions, as compared to trials following no-reward cues (Figure 8). The no-reward RT's of both conditions were consistent with one another, providing a steady baseline allowing for inter-conditional comparisons. In the relevant condition, the difference in RT's was significant for both reward values ($p = 2.553e-11$ for +5 versus no-reward; $p < 2.2e-16$ for +10 versus no-reward). In the irrelevant condition, only one reward value was significantly faster than baseline ($p = 0.02166$ for +5 versus no-reward; $p = 0.9806$ for +10 versus no-reward). Regardless of significance from baseline, the relevant reward values were significantly faster than the irrelevant reward values ($p = 4.322e-05$ for +5; $p = 5.044e-08$ for +10).

Visualized again in Figure 9a, the same average group data showed a clear reward effect between the relevant and irrelevant conditions, such that the relevant rewards elicited faster RT's while the irrelevant rewards brought about RT's consistent with the no-reward RT's (Figure 9a). In figure 9b, the group data is broken down by condition and reward value. We saw trends consistent with the original hypothesis: only relevant rewards had faster-than-baseline RT's because participants could predict the incoming age judgment based on the reward value and the given reward-transition relationship. In particular, both relevant reward values had a reward effect in the expected direction. The relevant baseline, or no-reward RT's, were similar to the RT's for both irrelevant reward values and the irrelevant no-reward RT's. Whereas in Experiment 1, the relevant +5 reward value produced faster RT's than the relevant +10 reward value, both relevant reward values in this experiment produced RT's that were significantly faster than baseline and that were consistent with one another. In fact, this data shows that the +10 relevant reward elicited slightly faster RT's than the +5 reward value. These behavioral trends support our hypothesis, but we needed to conduct a more rigorous analysis to determine if these trends were consistent across participants.

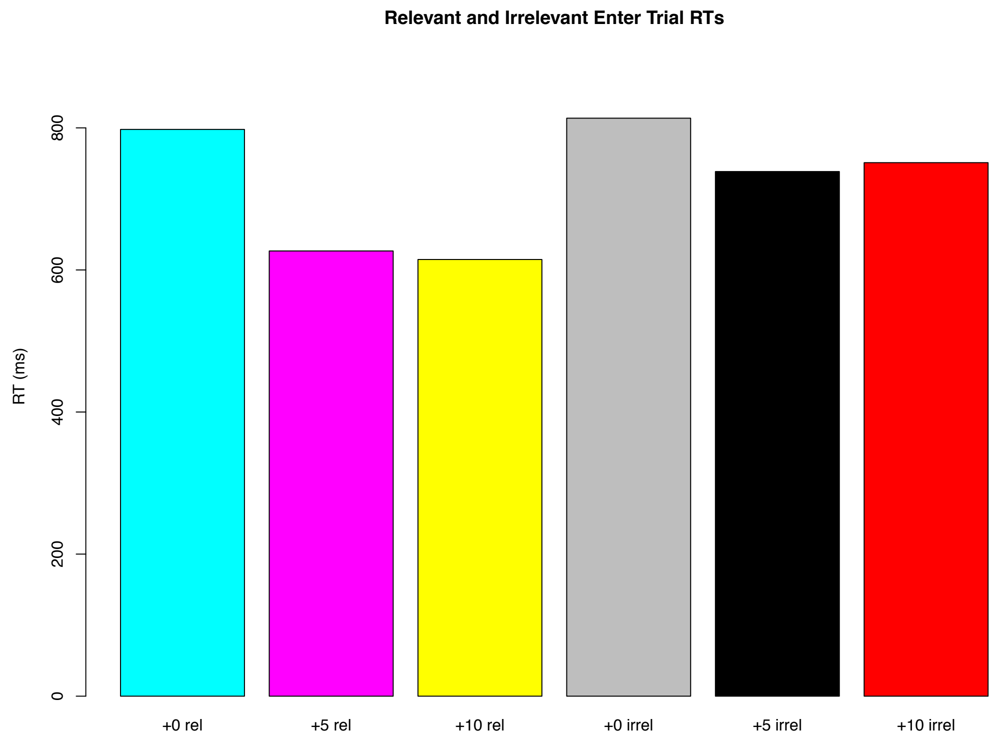
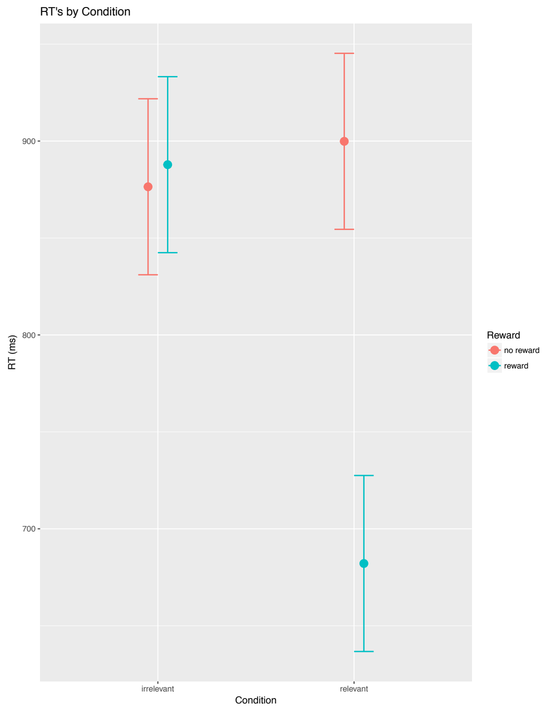


Figure 8: Relevant and Irrelevant Enter Trial RT's – group RT averages by condition and reward value on Enter trials in experimental blocks. Irrelevant rewarded RT's are faster than baseline, but relevant rewarded RT's are significantly after than irrelevant rewarded RT's. These reward effects are comparable between reward values.



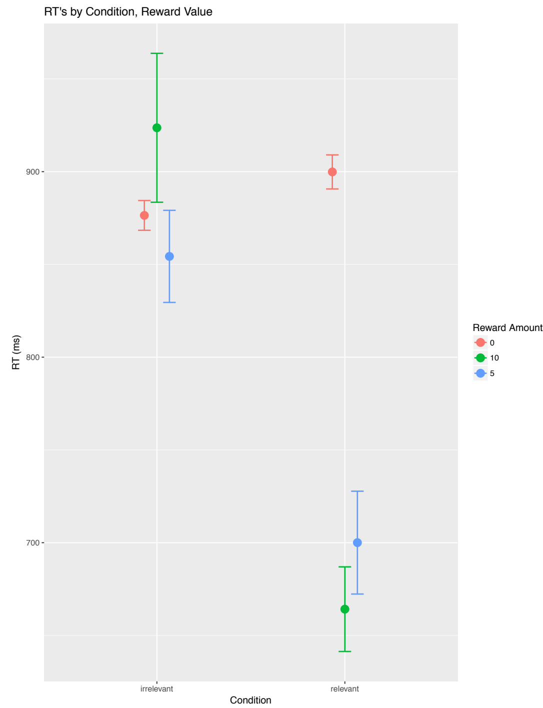


Figure 9a: RT's by Condition – (left) group RT averages by condition over experimental blocks 2-9, showing a clear reward effect in relevant condition. **Figure 9b: RT's by Condition, Reward Value** – (right) group RT averages by condition and reward value over experimental blocks 2-9. This data shows reward effects in both reward values of the relevant condition with clustering in the irrelevant condition.

Subject-by-subject analysis revealed two major behavioral trend classifications from the RT data. First, four of the ten participants produced data that showed a clear reward effect in the relevant condition only, with consistent RT's across outcomes in the irrelevant condition. These participants were able to track the task structure successfully in order to derive predictive information from the relevant reward cues and ignore the irrelevant reward cues. Exemplary data from one subject is displayed in Figure 9c. The second major behavioral trend classification encompassed four subjects that did not track the task structure as well as expected. In particular, these subjects displayed clear relevant reward effects, with rewarded RT's being faster than no-reward RT's, but they also showed a reward effect in the irrelevant condition. We could infer from this data that these participants were not able to implement the instructions regarding CLUE (relevant) versus NO CLUE (irrelevant) blocks as well as expected. Notably, three of the four in this classification exhibited the uncharacteristically fast RT's for the irrelevant +5 reward value. These RT's were comparable with the relevant reward value RT's. The fourth participant in this group had uncharacteristically fast RT's for the irrelevant +10 reward value instead.

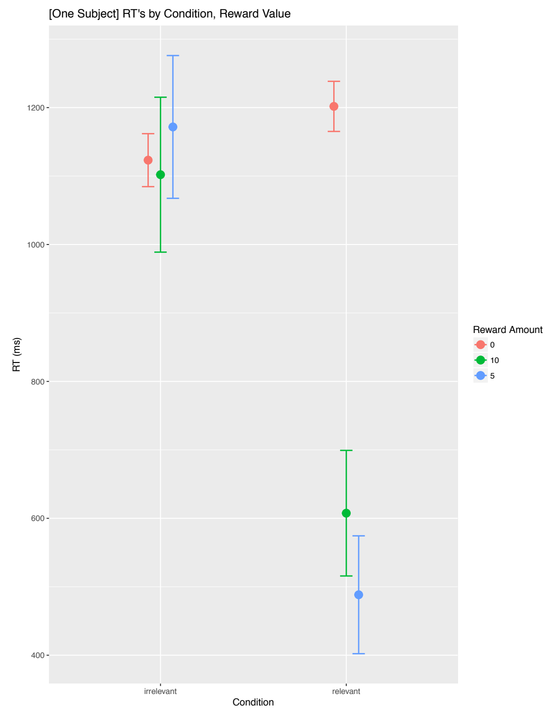


Figure 9c: [One Subject] RT's by Condition and Reward Value – exemplary RT data from a participant showing clear reward effect in the relevant condition and RT clustering in irrelevant condition.

The three subjects with fast irrelevant +5's were part of the HVSame condition, while the subject with the fast irrelevant +10's was in the HVDifferent condition. For the HVSame condition, the +5 rewards were shown in different-age transitions, and for the HVDifferent condition, the +10 rewards were shown in different-age transitions. Considering these relationships, it is possible that participants actively learned the different-age transition relationship and used it to infer the same-age transition relationship instead of learning both rules and implementing them in parallel. While this is a speculative exercise, these assumptions would provide that the different-age rule is easier to learn than the same-age rule, which is consistent with the reward-transition dynamics in Experiment 1.

To conclude the subject-by-subject analysis, the remaining two participants showed uncommon patterns of deviation from the expected behavioral trends. One participant had mirrored RT's across conditions, such that the relevant and irrelevant +10 reward values elicited below-baseline RT's, and the +5 and no-reward RT's in both conditions were clustered together. The other participant had a clear reward effect for the relevant +10 reward value, yet the relevant +5 RT's were slower than baseline. This trend echoes the single-reward-value reward effect observed in Experiment 1. Considered together, seven of the ten participants had a clear and equal reward effect for the relevant reward values. As compared to Experiment 1, the salience of this essential behavioral trend was much more pronounced and widespread across participants. While it is possible that individual variance from our small sample sizes drove the variation in behavioral trends, we looked to the difference in the tasks to account for the contract in trend clarity. The data from this experiment, however, does not support the notion that the no-reward cues would impact performance significantly. Instead, the data shows that on Enter trials following no-reward cues, the RT's were the same as on no-cue Enter trials from Experiment 1. Perhaps neural data would represent the impact the no-reward cues had on the task structure and the salience of rewards more accurately than RT data. Regardless, in further experimentation with this task, we would seek to make the CLUE (relevant) and NO CLUE (irrelevant) block instructions more clear and more visible. We expect that change could reduce the behavioral result of participants reacting faster to irrelevant rewards.

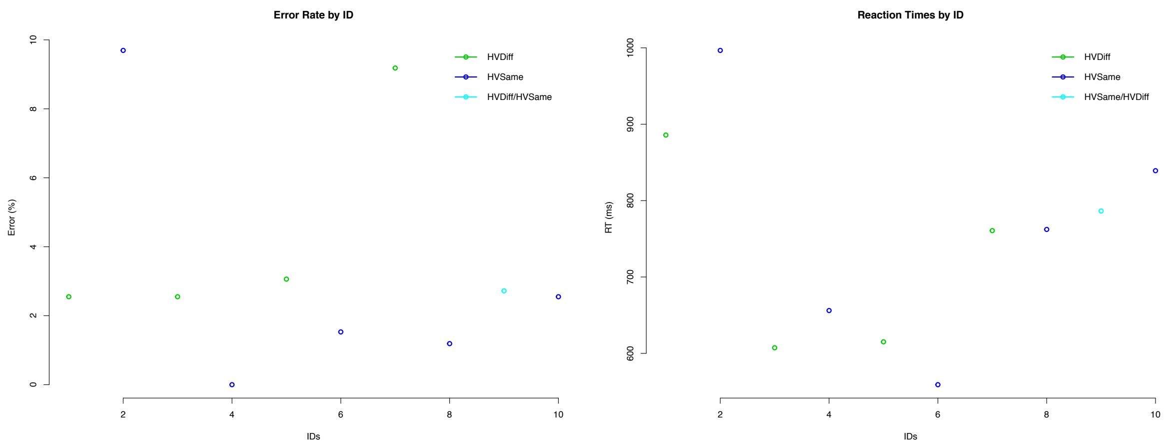


Figure 10a: Error Rates by ID – (left) average error rate for each participant by reward-transition relationship. HVDiff/HVSame category (cyan) was a participant that had both conditions in the last four experimental blocks in lieu of irrelevant blocks due to a malfunction in the stimulus code. HVDifferent participants had insignificantly higher error rates. **Figure 10b: Reaction Times by ID** – (right) Average RT's for each participant by reward-transition relationship. HVSame participants had insignificantly slower RT's.

As in Experiment 1, we analyzed the error rates in order to understand how participant performance improved or worsened over the course of this experiment. We started with a subject-by-subject error rate analysis. Figure 10a shows the error rates for each participant. Notably, the four subjects with the unexpected behavioral trend classification had below-average error rates. Their accuracy, we can assume, was not impacted by their lack of understanding of the instructions. The HVDifferent condition produced higher error rates than the HVSame condition, however not significantly ($p = 0.671$). In Figure 10b, HVDifferent had faster RT's than HVSame ($p = 0.6648$), so between the two measures, one reward transition condition does not appear to be more challenging than the other.

To verify that one reward transition relationship was not more difficult than the other, we looked specifically at the Enter RT's following relevant high-value rewards. Figure 11a shows that the HVDifferent and HVSame groups did not have significantly different RT's over the experimental blocks ($p = 0.2066$). Figure 11b shows the same analysis on the Enter RT's following relevant low-value rewards, and again, the groups did not have significantly different RT's over the experiment ($p = 0.178$). From this assessment, we concluded that the reward transition relationships did not disproportionately impact one group over another.

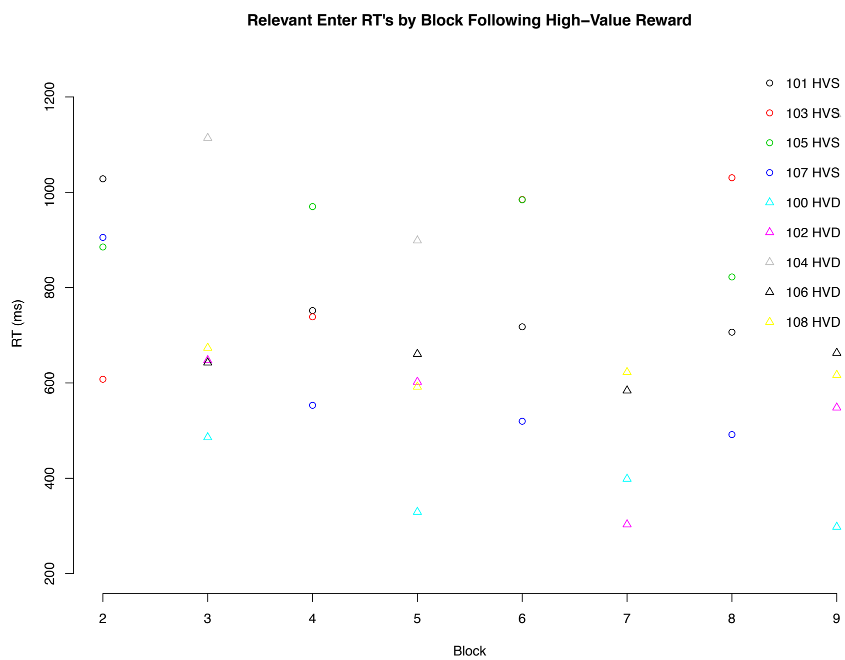
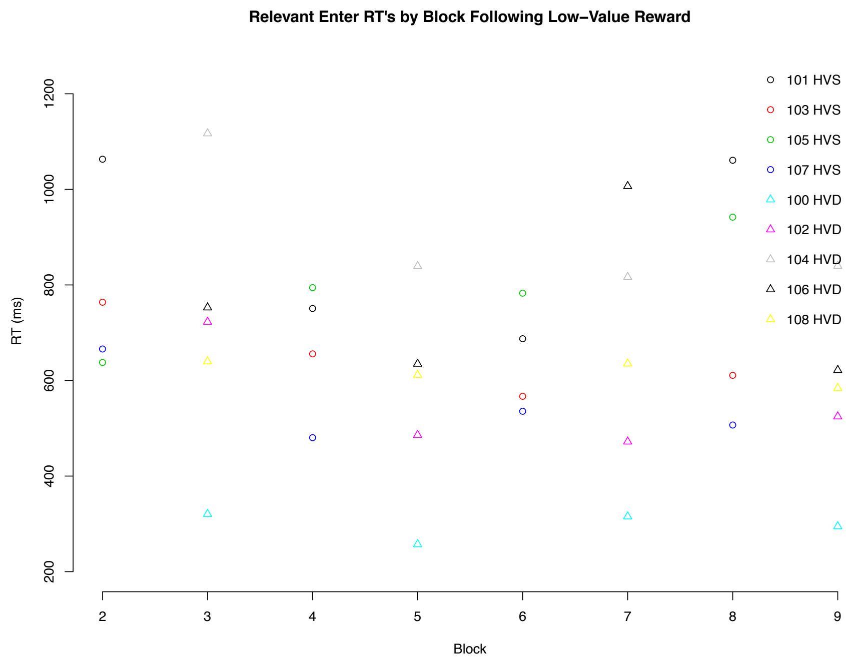


Figure 11a: Relevant Enter RT's by Block Following High-Value Reward – (top) HVDifferent (HVD – triangles) and HVSame (HVS – circles) Enter RT's after high-value (+10) rewards on relevant blocks. The two reward transition relationships do not elicit significantly different RT's. **Figure 11b: Relevant Enter RT's by Block Following Low-Value Reward** – (bottom) HVDifferent (HVD – triangles) and HVSame (HVS – circles) Enter RT's after low-value (+5) rewards on relevant blocks. The two reward transition relationships do not elicit significantly different RT's.



Next, we looked at the error rates over the experimental blocks. Figure 12a shows the average group data as it generally increased over the course of the experiment. It is important to note that in the verbal instructions delivered prior to the experiment, participants in this experiment were told to take small breaks between blocks to "stand up and shake it out" in order to maintain focus in the final blocks of the experiment. While the error rates did reach an experiment high around 5% in the final three blocks, the plateau in error rates exhibited in those last three blocks might be the result of a concerted effort to maintain focus at the end of the experiment. Also of note, the error rates were higher across this experiment than in the first experiment. As the only change to the task was the no-reward cues, it appears that these cues made the task more challenging. While the primary finding of this experiment supports the notion that the reward cues were used for more than a reminder of the category switch during a transition, this comparison suggests that the no-reward cues reduce the clarity or simplicity of the task structure by providing the same outcome on all the Enter and Internal trials and on half of the Exit trials. Figure 12b shows an inverse relationship between the error rates of the relevant and irrelevant blocks. These differences, however, were not significant and could likely be attributed to random subject variation, not a systematic difference in conditions (p = 0.697).

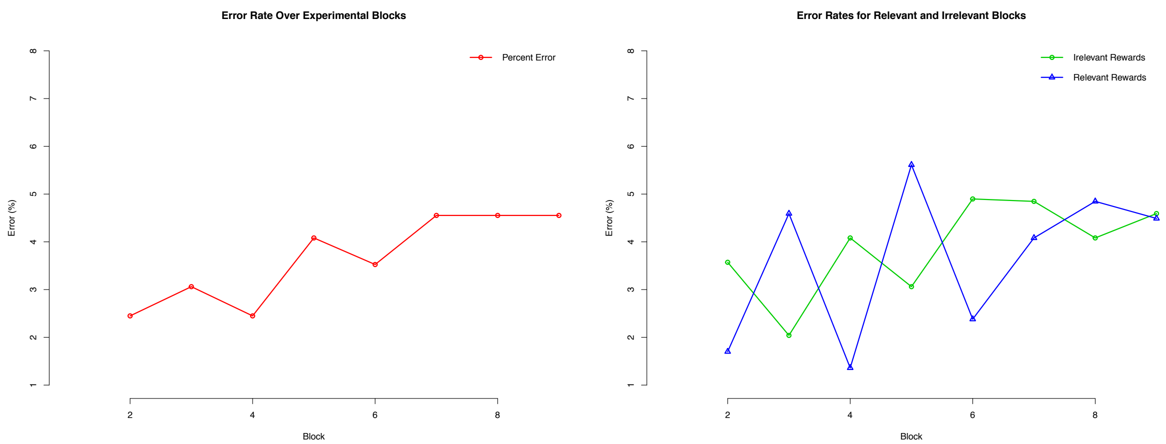


Figure 12a: Error Rates Over Experimental Blocks – (left) average error rate over experimental blocks 2-9. The last three blocks have the highest error rates, and these error rates are slightly higher than those in Exp. 1. **Figure 12b: Error Rates for Relevant and Irrelevant Blocks** – (right) average error rate over experimental blocks 2-9 by reward condition. The inverse effect between irrelevant and relevant error rates is likely due to noise as it is not significant.

Overall, the behavioral results from Experiment 2 support our hypothesis that participants used the reward cues to predict the age of the next trial, and the overarching behavioral hypothesis that relevant rewards elicit faster reaction times than non-rewarded trials and irrelevant rewarded trials. Taken together, these findings provide a foundation of support for our larger neural hypothesis that the human OFC encodes task-relevant values as part of the task structure. We will next look at fMRI data collected while participants played an earlier version of this task in order to see how brain activation is affected by this task structure.

Chapter 4: fMRI analysis

Using an earlier version of the Shuck et al. (2016) task, Schuck conducted an fMRI experiment with human participants. The following is unpublished imaging data from that scanning experiment. Importantly, the task used for these scans did not include economic values in the form of rewards. Additionally, this version of the task had random age-judgment key mapping, and subjects indicated their response using two buttons on a controller held in their right hand. The rest of the task structure matched the task outlined in the present work.

The scanner used for this experiment was a 3T Siemens Magnetom Skyra (Seimens). The voxel size was 3x3 mm, and the slice thickness was 2 mm with a gap of 50%. The acquired sequences had 46 axial slices each, with a TR of 2.4 seconds, TE of 27 milliseconds, FOV of 196 mm and a flip angle of 71 degrees. Slices were oriented 30 degrees backwards relative to the anterior-posterior commissure axis.

In this post-hoc analysis, we asked how the category – face or house – of the stimuli's attended dimension affected brain activation. Face images have been shown to activate the fusiform face area (FFA) in extrastriate visual cortex above baseline (Kanwisher et al., 1997), and house or building images have been shown to activate the parahippocampal place area (PPA) above baseline (Epstein et al., 1999). Considering these findings, we predicted that the participants' brain activation would increase at FFA when the attended dimension was a face image and would increase at PPA when the attended dimension was a house image. Additionally, we expected that when the attended dimension of the stimuli was a face, PPA would not activate significantly because of the selective attention focused on the face image. Although the face and the house images were equally visible in every composite image, we predicted that FFA would not activate significantly when the house was the attended dimension.

Figure 13 shows statistical brain images from 6 subjects attending to either faces or houses. From the face attended stimuli, bilateral FFA activation was clear, while PPA activation was not significant (Figure 13a). From the house attended stimuli, we observed bilateral PPA activation with a greater degree of activation in the right hemisphere over the left hemisphere, with no significant FFA activation (Figure 13b). These imaging results are consistent with our hypothesis that participants' selective attention to the attended dimension of the composite images drives activation for that dimension only. Additionally, this analysis provides a sanity check that subjects running through a version of the 16-state space task are representing the task structure to the extent of attending to the appropriate dimension. Future imaging studies with this task should be able to rely on this sanity check and should structure their scanning to focus on reward encoding in OFC.

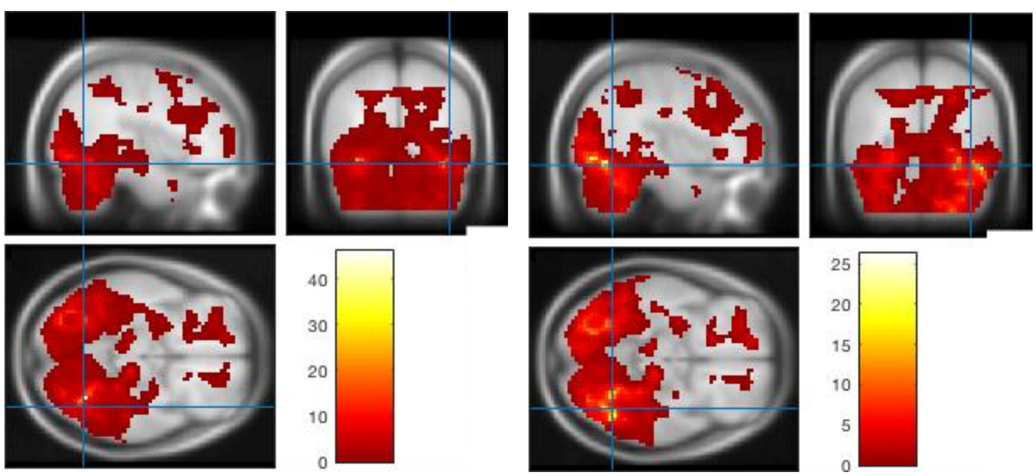


Figure 13a: Brain Activation for Face Attended Stimuli – (left) average brain activation over 6 subjects for all experimental trials in which the face image was the attended dimension. We expected FFA activation from the attended face images, and this data shows bilateral FFA activation. **Figure 13b: Brain Activation for House Attended Stimuli –** (right) average brain activation over 6 subjects for trials when house images were the attended dimension. We expected PPA activation from the attended house images, and this data shows strong PPA activation in both hemispheres, with slightly more activation in the right hemisphere.

Chapter 5: Discussion

The present work provides a foundation of behavioral results in support of our neural hypothesis and prepares future researchers to scan this task to decode OFC's activity. We predicted that reward cues would decrease RT's in the relevant condition and have no effect on RT's in the irrelevant condition. Despite behavioral noise – likely stemming from our small sample sizes – our reaction time analyses showed faster RT's evoked by both low- and high-value relevant rewards, as predicted. In the irrelevant condition, both reward values were clustered with the no-reward RT's, signaling the lack of a reward effect. Our average group data – particularly from Experiment 2 – and the majority of individual participant data provided support for our behavioral hypothesis. These behavioral trends, complimented by post-hoc fMRI analysis, suggested that our participants were successfully tracking the task structure and using the relevant reward cues to predict the subsequent age judgment. While we recognize that behavioral results cannot provide direct support to a neural hypothesis, these trends support our behavioral hypothesis, which, in turn, supports our prediction that task-relevant economic values are encoded in the task state representation in human OFC.

Future functional imaging research using this task should be tailored to decoding the task state representation from OFC activity. Additionally, we suggest incorporating the no-reward cues from Experiment 2 into the task for scanning, as this version of the task produced stronger behavioral results. We predict that the task's relevant rewards values - +5 and +10 – will be included in the task state representation, as suggested by Schuck et al. (2016). Hence, with an fMRI experiment, we expect that these relevant reward values would be decoded as part of the task state representation. Evidence supporting this hypothesis would also support a characterization of OFC as a region that encodes partially-observable task states and uses them to calculate or learn values. This characterization could possibly replace the characterization of OFC as a value signaling region under the goods-based model (Padoa-Schioppa & Assad, 2006).

More broadly, this research aims to understand how rewards impact human decision making. While the goods-based theory of OFC function would suggest that rewards impact decision making in a value-only space (Padoa-Schioppa & Assad, 2006), these results and neural predictions would imply that rewards are one of many factors that impact decision making in a larger task space. Future research in this vein will provide insights into neuroeconomics and will create a better understanding of how we flexibly use rewards in every-day decisions. There are still unanswered questions about the OFC, but we hope this study brought the field one step closer to understanding the role of the human OFC.

References:

Bradfield, L. A., Bertran-Gonzalez, J., Chieng, B., & Balleine, B. W. (2013). The Thalamostriatal Pathway and Cholinergic Control of Goal-Directed Action: Interlacing New with Existing Learning in the Striatum. *Neuron*, *79*(1), 153–166. https://doi.org/10.1016/j.neuron.2013.04.039

Chan, S. C. Y., Niv, Y., & Norman, K. A. (2016). A Probability Distribution over Latent Causes, in the Orbitofrontal Cortex. *Journal of Neuroscience*, *36*(30), 7817–7828. https://doi.org/10.1523/JNEUROSCI.0659-16.2016

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The Parahippocampal Place Area: Recognition, Navigation, or Encoding? *Neuron*, *23*(1), 115–125. https://doi.org/10.1016/S0896-6273(00)80758-8

Farovik, A., Place, R. J., McKenzie, S., Porter, B., Munro, C. E., & Eichenbaum, H. (2015). Orbitofrontal Cortex Encodes Memories within Value-Based Schemas and Represents Contexts That Guide Memory Retrieval. *Journal of Neuroscience*, *35*(21), 8333–8344. https://doi.org/10.1523/JNEUROSCI.0134-15.2015

FitzGerald, T. H. B., Seymour, B., & Dolan, R. J. (2009). The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects. *Journal of Neuroscience*, *29*(26), 8388–8395. https://doi.org/10.1523/JNEUROSCI.0717-09.2009

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding Predictive Reward Value in Human Amygdala and Orbitofrontal Cortex. *Science*, *301*(5636), 1104–1107.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Klein-Flügge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J., & Behrens, T. E. J. (2013). Segregated Encoding of Reward–Identity and Stimulus–Reward Associations in Human Orbitofrontal Cortex. *Journal of Neuroscience*, *33*(7), 3202–3211. https://doi.org/10.1523/JNEUROSCI.2532-12.2013

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., & Rushworth, M. F. S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences*, *107*(47), 20547–20552. https://doi.org/10.1073/pnas.1012246107

Noonan, M. P., Kolling, N., Walton, M. E., & Rushworth, M. F. S. (2012). Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *European Journal of Neuroscience*, *35*(7), 997–1010. https://doi.org/10.1111/j.1460-9568.2012.08023.x

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature; London*, *441*(7090), 223–226. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1038/nature04676

Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience; New York*, *11*(1), 95–102. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1038/nn2020

Padoa-Schioppa, C., & Conen, K. E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron*, *96*(4), 736–754. https://doi.org/10.1016/j.neuron.2017.09.031

Padoa-Schioppa, C. (2009). Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex. *Journal of Neuroscience*, *29*(44), 14004–14014. https://doi.org/10.1523/JNEUROSCI.3751-09.2009

Padoa-Schioppa, C. (2011). Neurobiology of Economic Choice: A Good-Based Model. *Annual Review of Neuroscience*, *34*(1), 333–359. https://doi.org/10.1146/annurev-neuro-061010-113648

Rescorla, R. A, & Wagner, AR. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non reinforcement. In AH. Black & W.F. Prokasy (eds.), Classical conditioning II: current research and theory (pp. 64-99) New York: Appleton-Century-Crofts.

Rich, E. L., & Wallis, J. D. (2016). Decoding subjective decisions from orbitofrontal cortex. *Nature Neuroscience; New York*, *19*(7), 973–980. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1038/nn.4320

Roesch, M. R., Taylor, A. R., & Schoenbaum, G. (2006). Encoding of Time-Discounted Rewards in Orbitofrontal Cortex Is Independent of Value Representation. *Neuron; Cambridge*, *51*(4), 509–520. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1016/j.neuron.2006.06.027

Rudebeck, P. H., & Murray, E. A. (2011). Dissociable Effects of Subtotal Lesions within the Macaque Orbital Prefrontal Cortex on Reward-Guided Behavior. *Journal of Neuroscience*, *31*(29), 10569–10578. https://doi.org/10.1523/JNEUROSCI.0091-11.2011

Saez, R. A., Saez, A., Paton, J. J., Lau, B., & Salzman, C. D. (2017). Distinct Roles for the Amygdala and Orbitofrontal Cortex in Representing the Relative Amount of Expected Reward. *Neuron; Cambridge*, *95*(1), 70–77.

https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1016/j.neuron.2017.06.012

Schoenbaum, G., Takahashi, Y., Liu, T.-L., & McDannald, M. A. (2011). Does the orbitofrontal cortex signal value? *Annals of the New York Academy of Sciences*, *1239*(1), 87–99. https://doi.org/10.1111/j.1749-6632.2011.06210.x

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron; Cambridge*, *91*(6), 1402–1412. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1016/j.neuron.2016.08.019

Schuck, N. W., Wilson, R. C., & Niv, Y. (2017). A state representation for reinforcement learning and decision-making in the orbitofrontal cortex. *bioRxiv*, 210591. https://doi.org/10.1101/210591

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nature Neuroscience; New York*, *18*(5), 620–627. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1038/nn.3982

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. MIT Press

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., … Morris, R. G. M. (2007). Schemas and Memory Consolidation. *Science*, *316*(5821), 76–82. https://doi.org/10.1126/science.1135935

Wikenheiser, A. M., Marrero-Garcia, Y., & Schoenbaum, G. (2017). Suppression of Ventral Hippocampal Output Impairs Integrated Orbitofrontal Encoding of Task Structure. *Neuron; Cambridge*, *95*(5), 1197–1207. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1016/j.neuron.2017.08.003

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron; Cambridge*, *81*(2), 267–279. https://doi.org/http://dx.doi.org.ezproxy.princeton.edu/10.1016/j.neuron.2013.11.005

This paper represents my own work in accordance with University regulations.

Riley MacAulay