

Spring 4-2014

# Familiarity-Based Object Recognition across Animacy and Real-World Size

Sheldon Hill

Western University, sheldon.james.hill@gmail.com

Follow this and additional works at: [https://ir.lib.uwo.ca/psychd\\_uht](https://ir.lib.uwo.ca/psychd_uht)

---

## Recommended Citation

Hill, Sheldon, "Familiarity-Based Object Recognition across Animacy and Real-World Size" (2014). *Undergraduate Honors Theses*. 13. [https://ir.lib.uwo.ca/psychd\\_uht/13](https://ir.lib.uwo.ca/psychd_uht/13)

This Dissertation/Thesis is brought to you for free and open access by the Psychology Department at Scholarship@Western. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of Scholarship@Western. For more information, please contact [tadam@uwo.ca](mailto:tadam@uwo.ca), [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

Familiarity-Based Object Recognition across Animacy and Real-World Size

Sheldon Hill

Honours Psychology Thesis  
Department of Psychology  
University of Western Ontario  
London, Ontario, CANADA  
April, 2014

Thesis Advisor: Stefan Köhler, Ph.D.

### **Abstract**

Previous evidence suggests that parahippocampal cortex (PhC) is associated with contextual encoding and retrieval while perirhinal cortex (PrC) is associated with item familiarity. However, this is inconsistent with recent evidence that PhC is associated with item familiarity for some objects (Martin et al., 2013). The familiarity of items seems to be represented in both PhC and PrC, with object categories being represented in one area more strongly than the other, or dually associated in both. Stimuli properties of items in these categories are thought to play a role in this differential association. The purpose of the current experiment is to create an effective experimental paradigm that will be used in a future imaging study to examine the representation of item-based familiarity signals in PhC and PrC, as a function of varying stimulus dimensions. Thirty-nine students were recruited through the Psychology Research Participation Pool and were compensated 1.5 research credits for their participation. A continuous recognition memory task was employed, where items from three stimulus categories, animals, tools, and appliances, were presented and recognition memory responses were recorded. Through data analysis and iterative refinement of the paradigm, it was found that the fourth version of the paradigm should be used for a future imaging study because responses are primarily familiarity-based, overall recognition is significantly greater than chance, and two of the three stimulus categories are matched in performance.

### **Familiarity-Based Object Recognition across Animacy and Real-World Size**

Mapping out regions of the brain associated with cognitive processes and memory is an important part of neuroscientific research and one that is being studied more frequently.

Memory is a complex process and although there have been numerous experiments studying the neural correlates of memory in the extant literature, there are many aspects of memory that still need to be explored. One particular memory process that has received much attention in the cognitive, neuropsychological, and neuroimaging literature is recognition memory.

#### **Recognition Memory: Recollection and Familiarity**

Recognition memory is the ability to recognize stimuli that have been previously encountered (Eichenbaum, Yonelinas, & Ranganath, 2007). People encounter countless stimuli every day and some of these stimuli may be more memorable than others, such that they are easier to recognize when they are reencountered. This may be differentiated into two types of recognition memory: recollection and familiarity. Recollection is described as the retrieval of contextual details when the stimulus is reencountered, while familiarity is described as a feeling of knowing of the reencountered stimulus, without any additional contextual information retrieved (Eichenbaum et al., 2007). For example, if an individual recognizes someone walking down the street because they work with that person, it would be recollection because they are retrieving contextual details about the building they work in, the job, and any number of details they have associated with that person. If the same individual recognized someone walking down the street, but did not recall from where they knew them or any contextual details about them, but they felt like they knew that person, it would be familiarity. Recognition memory has been studied in the extant literature from a neuroscientific approach and the areas of the brain correlated with recognition memory have been identified.

In a review of studies addressing the neural correlates of declarative memory, Manns and Eichenbaum (2006) proposed a highly influential view that recognition memory is supported by the medial temporal lobes; specifically, the hippocampus and parahippocampal region, which is a region that can be subdivided into parahippocampal, perirhinal and entorhinal cortex. Similar to the two types of recognition memory, it was proposed that the medial temporal lobes are functionally divided into cortex associated with recollection or familiarity. The hippocampus and parahippocampal cortex (PhC) are associated with recollection, while perirhinal cortex (PrC) is associated with familiarity (Manns and Eichenbaum, 2006). Although these regions are each associated with a single type of recognition memory, like many regions of the brain, there is an interplay between the different areas of the medial temporal lobe, characterizing a highly interconnected neural network.

### **Binding of Item and Context Model**

Ranganath (2010) proposed that episodic memory is processed within the medial temporal lobes. Episodic memory is memory associated with a person's life and their past experiences (Eichenbaum et al., 2007). Recognition memory is a type of episodic memory because recognition requires that one must have personally encountered the stimulus. Ranganath (2010) proposed a network of neural connections in the medial temporal lobes, called the Binding of Item and Context (BIC) model. In agreement with previous literature, Ranganath (2010) proposed that PhC was solely associated with the encoding and retrieval of contextual representations and that PrC was solely associated with the familiarity of items. The BIC model suggests that the contextual representations of a stimulus can strongly impact item representation, so PhC can thus have a strong effect on PrC. The BIC model also suggests that contextual information and the item representations converge in the hippocampus, where they

are “bound” together. This lends support to recollection memory, where the contextual details are retrieved with the reencountered item because the memories of the two are intertwined. Conversely, brain signal activation during familiarity responses are proposed to be localized to PrC. The BIC model suggests that context and item representations remain separate in the medial temporal lobes until they reach the hippocampus. However, the BIC model and the sectioning of PhC and PrC into different functional regions does not always hold amongst all stimuli used in recognition memory tasks.

### **Familiarity in Medial Temporal Lobes**

In a recent experiment, item-based familiarity was examined in the activation of PhC and PrC, across three different categories of stimuli (Martin, McLean, O’Neil, & Kohler, 2013). Martin et al. (2013) found that activation in the PhC is associated with some categories of objects; these categories show a correlation with familiarity signals in PhC. PhC was also more active than PrC, suggesting a preferential activation for some categories of stimuli. This suggests a difference in item-based familiarity signals across stimulus categories, where some stimulus categories are represented in PhC, but other categories are represented in PrC.

Martin et al. (2013) investigated how item categories are represented in brain regions during familiarity responses and sought to determine which characteristics of these stimuli categories mediate familiarity signals between PhC and PrC. This question was motivated by an imaging study by Martin et al. (2013) which examined if the PhC, which is anatomically adjacent to PrC, supports item-based familiarity signals.

Familiarity was examined across three categories of stimuli: faces, chairs and buildings. The experimental task started with an encoding stage, where face, chair, and building images were presented to the participant, who was asked to make an attractive, comfort or value

judgement, respectively. The encoding stage was completed prior to scanning, and approximately an hour after the encoding stage, the participants entered the fMRI scanner to perform the testing stage. In the scanner, they were presented with stimuli of faces, chairs and buildings. They were asked to make a judgement after the presentation of the stimulus, on a familiarity scale, where one was 'not familiar' and four was 'very familiar'. There was also a fifth response of 'recollection', if the participant happened to retrieve contextual details during the recognition response. Some of the stimuli presented were "old" and had been presented once in the encoding stage, while some stimuli were novel. This experiment focused on familiarity so the design intentionally minimized contextual details in the items. Specifically, the stimuli had all background contextual details removed and item analyses were completed in a pilot study to ensure that the items supported familiarity responses. The participants were instructed to make a fast judgement on the presented item and to attempt to avoid intentional recollection of details, to decrease the frequency of recollection responses.

The results of Martin et al. (2013) demonstrates that the precautions taken to avoid recollection responses were effective because the frequency of familiarity responses was greater than that of recollection responses. There were four regions of interest (ROI) that were studied in the fMRI analyses: PhC, PrC, anterior hippocampus and posterior hippocampus. Between the three item categories, the anterior and posterior hippocampus did not show a significant change in activation. Chairs elicited significant activation in PhC, as well as PrC. Buildings elicited significant activation in PhC, while faces elicited significant activation in PrC. In summary, chairs were associated with both PhC and PrC, buildings were only associated with PhC and faces were only associated with PrC, in terms of item-based familiarity signals.

These findings are contrary to the BIC model, which proposed that PhC is solely associated with contextual information while PrC is solely associated with item familiarity. These results suggest that the BIC model is too exclusive by defining a rigid differentiation between PhC and PrC; PhC is associated with item-based familiarity. The differences in stimulus properties of the three categories are likely the cause of the differences in location of the familiarity signals. Chairs are dually associated, so both regions must have a preference for whichever properties characterize chairs. As well, the differences in familiarity signals between PhC and PrC in relation to buildings and faces, respectively, lends support to the idea that different stimulus properties may be the underlying cause to the location of familiarity-based signals in the brain.

### **Animacy and Real-World Size**

The current study will examine the properties of different stimulus categories and how these properties affect the location of the familiarity signal. Stimuli can exhibit a number of different properties, more than can be explored in a single study, especially because the properties studied would ideally be differentially associated with PhC and PrC. The properties of the stimulus categories may be due to visual differences or semantic differences. Visual characteristics that may cause differences in familiarity signal localization include: colour, form, shape, distance and size. Semantic characteristics that may cause differences in familiarity signal localization may be: how the stimulus can be used, how someone interacts with the stimulus, and what the stimuli does in everyday life.

In a recent experiment focusing on the visual modality, Konkle and Caramazza (2013) used fMRI analyses to map the activation of the brain during image presented tasks. They varied the stimuli on two dimensions: animacy and real-world size. Specifically, they used images of



small animals, large animals, small objects and large objects to examine the organization of the neural activation from these stimuli properties. Although this was not a memory study, the fMRI data included part of the temporal lobes. The results showed that there is, what Konkle and Caramazza (2013) called, a tripartite organization, because small objects were activated in one network of the brain, large objects in a different network, and animals in a third network. It was found that the overall organization of the brain is associated with large objects being more posterior in the brain and small objects being more anterior. Specifically, it was found that PhC showed significant activation to large objects and since PhC is more posterior to PrC, there may be a differential representation between object sizes. There was no significant difference between animal sizes in brain activation so both PhC and PrC may become dually activated during recognition of animal stimuli.

### **Current Study**

The purpose of this experiment is to create a paradigm that will be modified through iterative refinement until it is ready to be used in an imaging study. Therefore, this acts as a pilot study by preparing stimuli and an experimental design to ensure that any findings in the future imaging study can be attributed solely to the underlying concepts being studied and not any flaws in the design. This experiment extends on results obtained by Martin et al. (2013), by focusing on the creation of a paradigm to study the representational structure of PhC and PrC in response to familiarity-based object recognition. The creation of this paradigm will ultimately lead to results that will show the differential activation of PhC and PrC, across two stimulus properties that have not been studied in terms of a memory experiment before. The two stimulus dimensions under investigation are animacy and real-world size, chosen because of the tripartite division displayed in Konkle and Caramazza's experiment (2013). This motivates the current

study by examining the composition of a tripartite division in the medial temporal lobes, in terms of item-based familiarity signals.

The current experiment will be performed with a similar design to that of Martin et al.'s (2013) study, with the important exception that the task will be continuous. This means that there will not be an encoding stage and a test stage, but only one stage where every stimuli presented will have a judgement decision required; the stimuli at the beginning will all be novel, but after a certain number of stimuli have been presented, some will repeat once and familiarity judgements can be made.

The current research is important to extend previous findings of familiarity memory and to lend support to the proposed mapping of the neural correlates of memory. Ultimately, understanding memory at a neural level will shed light on the networks that memory is involved in. By using stimuli that vary between animacy and real-world size, it may help elucidate the differences in object familiarity responses, based on differing properties of the stimuli.

Since the purpose of the experiment is the creation of the experimental paradigm, the design will be to run a number of subjects, analyze the results, then modify the design to ensure the components of the design are not skewing the data. The variation within and between stimulus categories is an anticipated problem, one that will be minimized and addressed through iterative refinement of this paradigm. Specifically, this refinement will ensure that recognition is primarily familiarity-based, that memory performance is significantly above chance, and that performance is matched between categories. This is an important aspect of the paradigm creation because the results should be attributable to differences in brain activation from the stimuli, and not because of differences in difficulty between categories. As well, this refinement

will ensure that some categories will not be too easy or too difficult to recognize, meaning that all data reflect what is supposed to be measured.

The refinement of this paradigm will be done by optimizing components of the design. Firstly, item analyses for stimulus categories will be completed to ensure that the stimuli presented are adequately matched. If some stimuli are innately easier or more difficult to remember, then the performance between categories will not be matched, so these stimuli will be replaced with items that do not cause this confounding effect. Secondly, the presentation time of the stimulus will be modified because a longer presentation time may lead to contextual encoding; for example, if there is an attention drawing event in the environment during the experiment or if the participants have a long time to study a stimulus, contextual details may be encoded with the stimulus. Thirdly, the amount of time given to make a judgement should be long enough for the participant to give an accurate judgement, but not so long that they recollect contextual details of whether they have seen the item before. Finally, the type of judgement is important to what data will be examined, whether it be by a familiarity rating scale or a judgement between novelty, recollection and familiarity. These parts of the design will start out as what was used in the Martin et al. (2013) experiment and then modified throughout the experiment based on the results.

## **Methods**

### **Participants**

Thirty-seven participants took part in the experiment (20 females, 17 males; age  $M = 19$ ,  $SD = 1.59$ ). Participants were recruited through the Psychology Research Participation Pool at the University of Western Ontario, which is a system that students use to find experiments where they can obtain research credits required for Psychology 1000. Participants were compensated

1.5 research credits for taking part in the experiment and were provided informed consent by the standards and procedures of the University of Western Ontario Health Sciences Research Ethics Board.

### **Stimuli**

Images varied according to three different categories (animate, large inanimate, and small inanimate) were the stimuli used in the current stimuli norming experiment. Specifically, animal images consisted of antlered animals and birds, large inanimate objects consisted of large appliances and furniture, and small inanimate objects consisted of tools, office supplies and kitchen utensils. For the animal stimuli, 300 images were obtained (150 antlered animals and 150 birds), while 150 images were obtained for large inanimate stimuli and 150 images for small inanimate stimuli. All images were obtained on the internet by using Google Image Search. All images had the backgrounds removed by using Adobe Photoshop and were fitted on a white background with no space-defining or contextual details present. The processed images were fitted and constrained to the white background template of 375 X 250 pixels to ensure stimuli were homogenous in size. For each of the three categories, 135 stimuli were selected and split amongst nine runs such that each run had an even mixture of the three stimuli types. Stimuli selected did not appear in more than one run. For each category, 15 stimuli were selected per run, 12 stimuli from each category would repeat, and the remaining three per category would act as filler items that would only appear once.

### **Procedure**

Participants were explained informed consent upon their arrival and were given a letter of information and signed the consent form. The experimenter would start the experiment on the computer and have the participant read the instructions on the first screen where definitions were

given for novelty, familiarity, and recollection. After the participant had finished reading, the experimenter would elaborate on the instructions by going into further detail about recognition memory, what novelty, familiarity and recollection are, and the theory motivating the experiment. This was all standardized for each participant by following a script. The experimenter would ask if there were any questions and have the participant explain the differences between novelty, familiarity and recollection, in their own words, before advancing to the next screen.

The second screen explained what the participant would be doing in the experiment and the scale they would be responding with. They were told that they would be presented with items that would be shown for 2.5 seconds and they would make a judgement on the keyboard, where one was 'sure novel', two was 'unsure novel', three was 'unsure familiar', four was 'sure familiar', and nine was 'recollection'. They were urged to make this judgement as soon as possible during the item presentation and to respond to every item. A crosshairs was presented in between items and a red asterisk would appear 1 second before the presentation of an item. The time between each item presentation was variable, ranging between 2.5-15 seconds.

There were nine runs, where any given item would only appear within a single run. Some of the items presented would appear more than once within the same run, while some items would only be shown once. Between runs, participants were given breaks that were as long as the participant wished. After the experiment was completed, the participant was given a debriefing form, asked if they had any questions, and thanked for their participation in the experiment. The experiment took approximately an hour and a half to complete.

Data analysis was performed in Python, an open-source computing program. One-way ANOVAs were performed for recognition performance in each category, in terms of overall

recognition, familiarity responses and recollection responses. Post-hoc tests were performed when data was significantly different. Overall response counts were recorded for each of the three stimuli categories, as well as the proportion of responses made across the three stimulus categories based on whether the item was novel or old.

### **Version One.**

Many of the things that were optimized throughout the experiment were decided on based on the design of Martin et al.'s study (2013). In terms of the four components to be optimized, it would be worthwhile to highlight them here. In this first version of the experiment, items were randomly generated from the existing stimuli sets. The presentation time of each item was 2.5 seconds and the judgement was asked to be made during that presentation, so response time shared the same amount of time. The judgement scale was one was 'sure novel', two was 'unsure novel', three was 'unsure familiar', four was 'sure familiar', and nine was 'recollection'.

### **Version Two.**

Based on the results from version one, the design was changed such that the presentation time of each item was 1.5 seconds, rather than 2.5 seconds. Within the animate stimuli, the antlered animals were removed from the stimuli set and only bird stimuli were used to characterize the animate test category. Within the large inanimate stimuli, each run had a subcategory that the stimuli were ordered into. For example, all large inanimate in the first run were all fridges, in the second run were all stoves, in the third run all couches, etc.

### **Version Three.**

Based on the results from version two, the design was changed such that four of the runs of large inanimate stimuli were different. The stimuli for these four runs were removed and replaced with completely new sets of ordered large inanimate stimuli.

### **Version Four.**

Based on the results from version three, the design was changed such that the large inanimate stimuli were ordered with two subcategories. For example, the first run was half fridges and half tables for the large inanimate stimuli for that run.

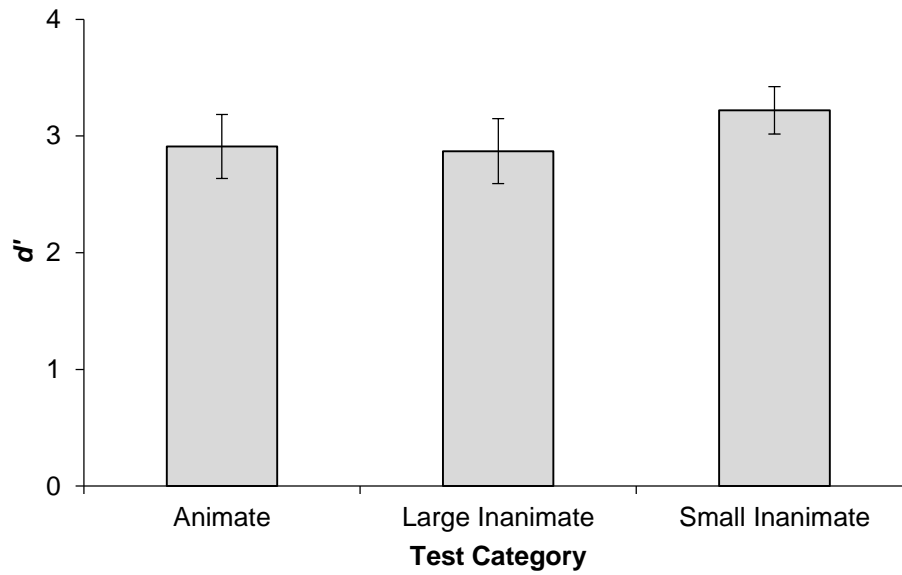
## **Results**

### **Version One**

There were ten participants tested for this version of the experiment. One-way ANOVAs were performed for overall recognition, familiarity, and recollection. Overall recognition performance was not significantly different between categories,  $F(2, 7) = 0.53, p = .60$ , Figure 1. Overall recognition memory performance is significantly above chance, because  $d'$  is greater than zero, Figure 1. Familiarity recognition performance was not significantly different between categories,  $F(2, 7) = 0.81, p = .46$ , Figure 2. Recollection recognition performance was not significantly different between categories,  $F(2, 7) = 0.06, p = .94$ , Figure 3. The average number of responses for each category and response is displayed in Figure 4.

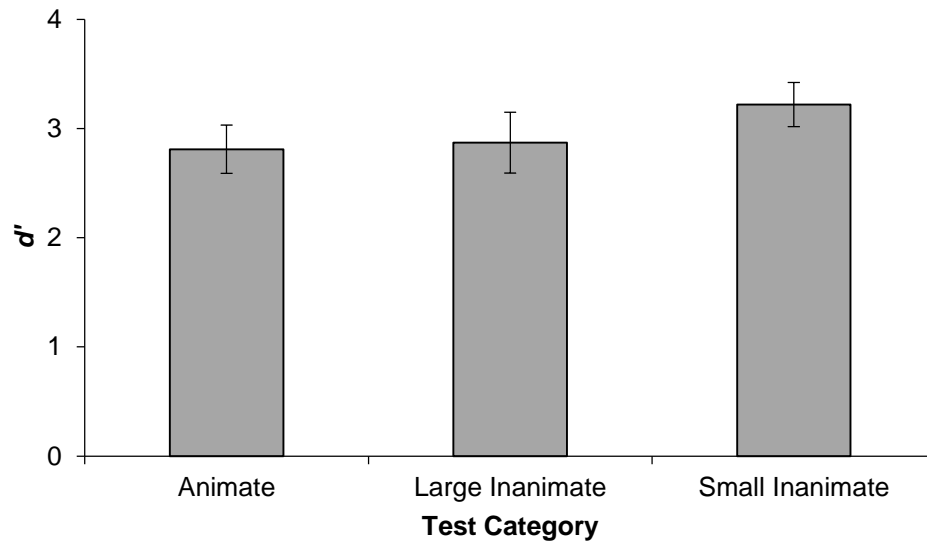
### **Version Two**

There were nine participants tested for this version of the experiment. One-way ANOVAs were performed for overall recognition, familiarity, and recollection. Overall recognition performance was not significantly different between categories,  $F(2, 6) = 3.28, p = .06$ , Figure 5. Overall recognition memory performance is significantly above chance, because  $d'$  is greater than zero, Figure 5. Familiarity recognition performance was significantly different between categories,  $F(2, 6) = 3.46, p = .04$ , Figure 6. A Tukey's HSD test was performed at  $p < .05$ , showing that there was a significant difference between large inanimate and small inanimate performance, and no difference between animate and large or small inanimate performance (Figure 6).

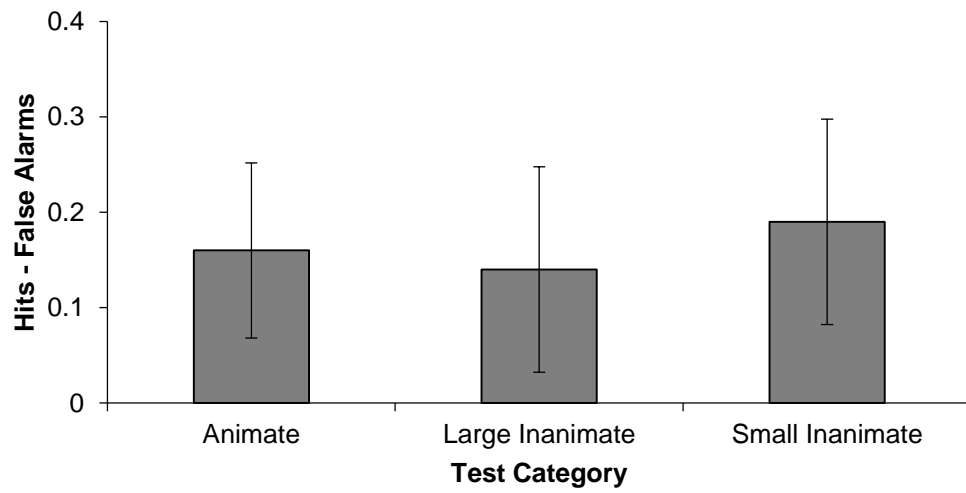


*Figure 1.* Overall recognition performance ( $d'$ ) between stimulus categories for version one of the experiment,  $n = 10$ . Error bars represent standard error.

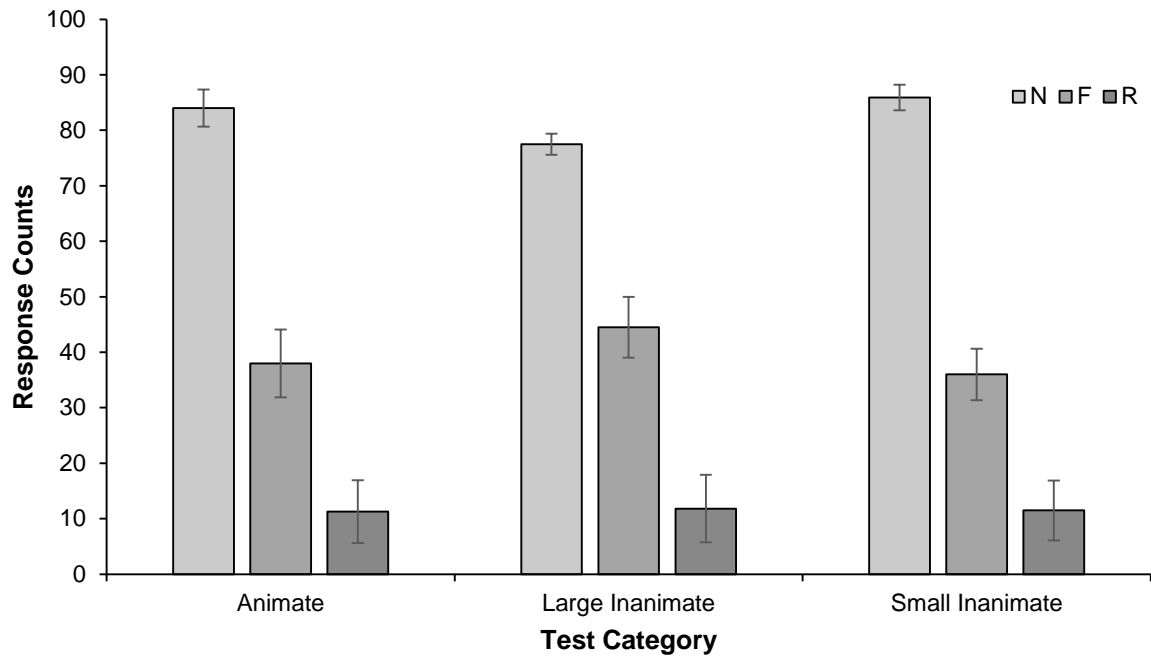




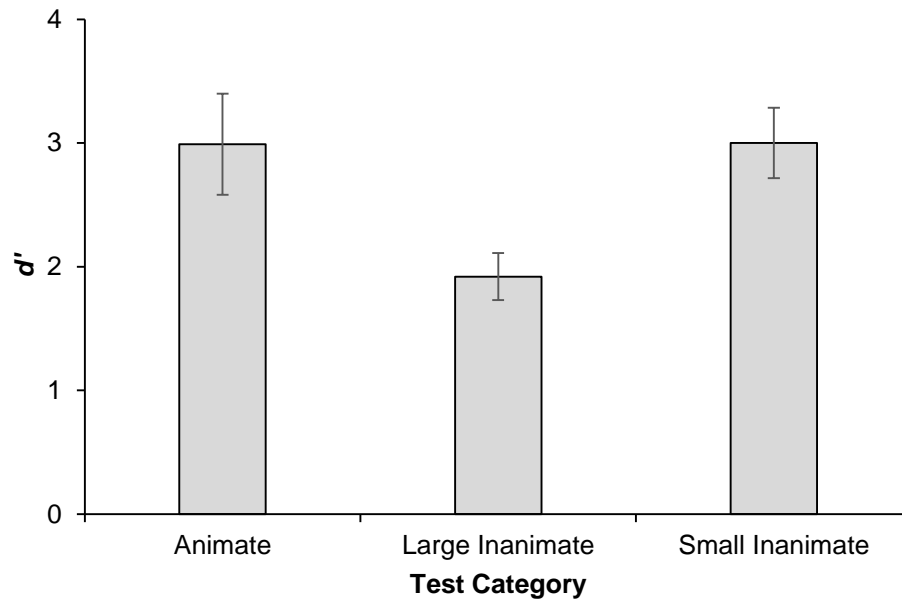
*Figure 2.* Familiarity recognition performance ( $d'$ ) between stimulus categories for version one of the experiment,  $n = 10$ . Error bars represent standard error.



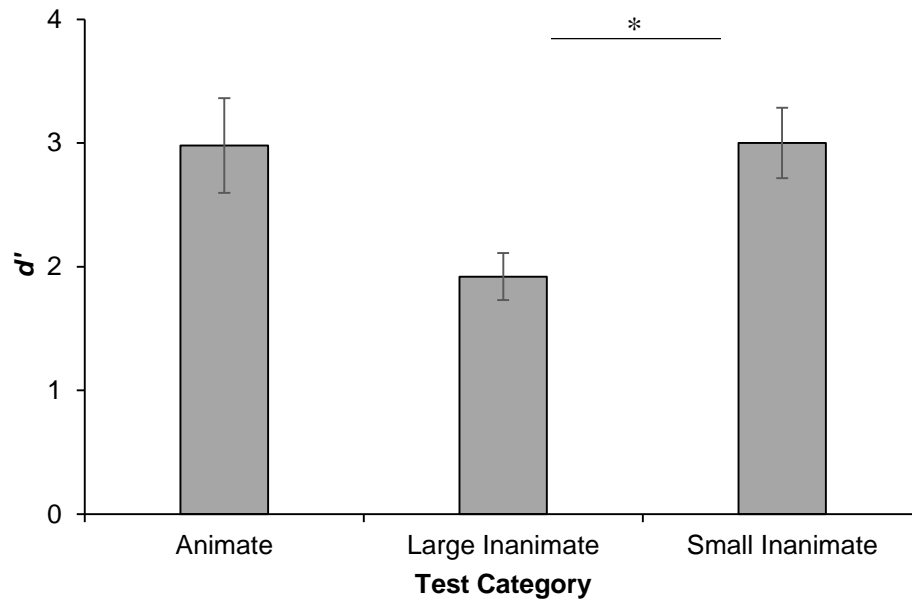
*Figure 3.* Recall recognition performance as a function of hits minus false alarms between stimulus categories for version one of the experiment,  $n = 10$ . Error bars represent standard error.



*Figure 4.* Average number of responses for each recognition response type, between stimulus categories for version one of the experiment,  $n = 10$ . Error bars represent standard error.



*Figure 5.* Overall recognition performance ( $d'$ ) between stimulus categories for version two of the experiment,  $n = 9$ . Error bars represent standard error.



*Figure 6.* Familiarity recognition performance ( $d'$ ) between stimulus categories for version two of the experiment,  $n = 9$ . Error bars represent standard error.  $*p < .05$ .

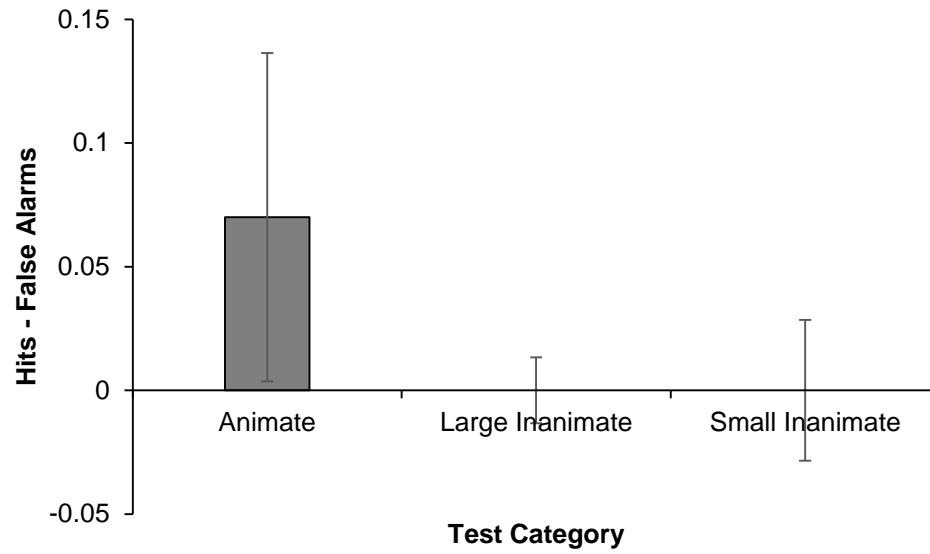
Recollection recognition performance was not significantly different between categories,  $F(2, 6) = 0.67, p = .52$ , Figure 7. The average number of responses for each category and response is displayed in Figure 8.

### **Version Three**

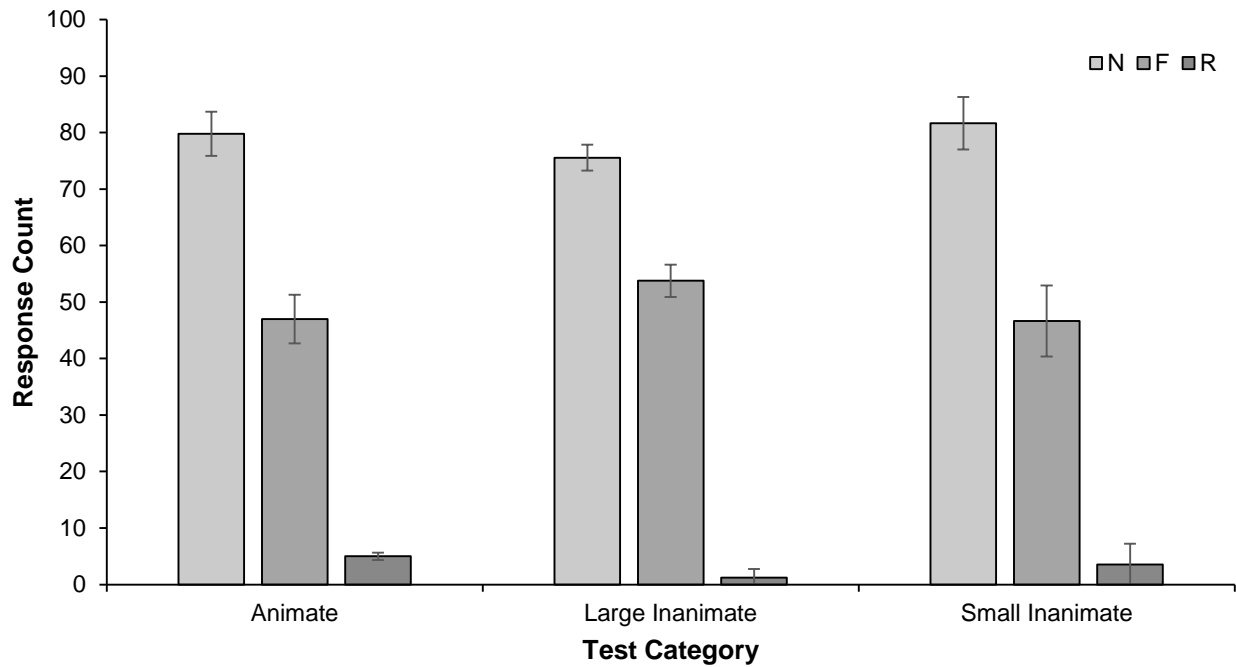
There were eight participants tested for this version of the experiment. One-way ANOVAs were performed for overall recognition, familiarity, and recollection. Overall recognition performance was significantly different between categories,  $F(2, 5) = 5.12, p = .02$ , Figure 9. A Tukey's HSD test was performed, showing that there was a significant difference between large inanimate and small inanimate performance, and no differences between animate and large or small inanimate. Overall recognition memory performance is significantly above chance, because  $d'$  is greater than zero, Figure 9. Familiarity recognition performance was significantly different between categories,  $F(2, 5) = 5.05, p = .02$ , Figure 10. A Tukey's HSD test was performed at  $p < .05$ , showing that there was a significant difference between large inanimate and small inanimate performance, between animate and small inanimate performance, but not between animate and large inanimate performance (Figure 10). Recollection recognition performance was not significantly different between categories,  $F(2, 5) = 0.20, p = .82$ , Figure 11. The average number of responses for each category and response is displayed in Figure 12.

### **Version Four**

There were ten participants tested for this version of the experiment. One-way ANOVAs were performed for overall recognition, familiarity, and recollection. Overall recognition performance was not significantly different between categories,  $F(2, 7) = 3.21, p = .06$ , Figure 13. Overall recognition memory performance is significantly above chance, because  $d'$  is greater than zero, Figure 13. Familiarity recognition performance was significantly different between

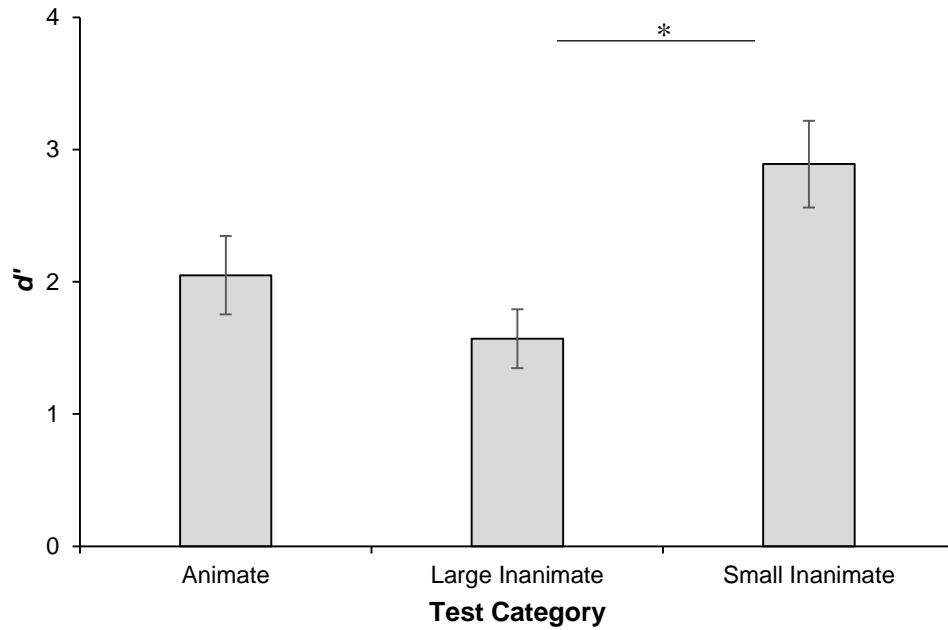


*Figure 7.* Recall recognition performance as a function of hits minus false alarms between stimulus categories for version two of the experiment,  $n = 9$ . Error bars represent standard error.

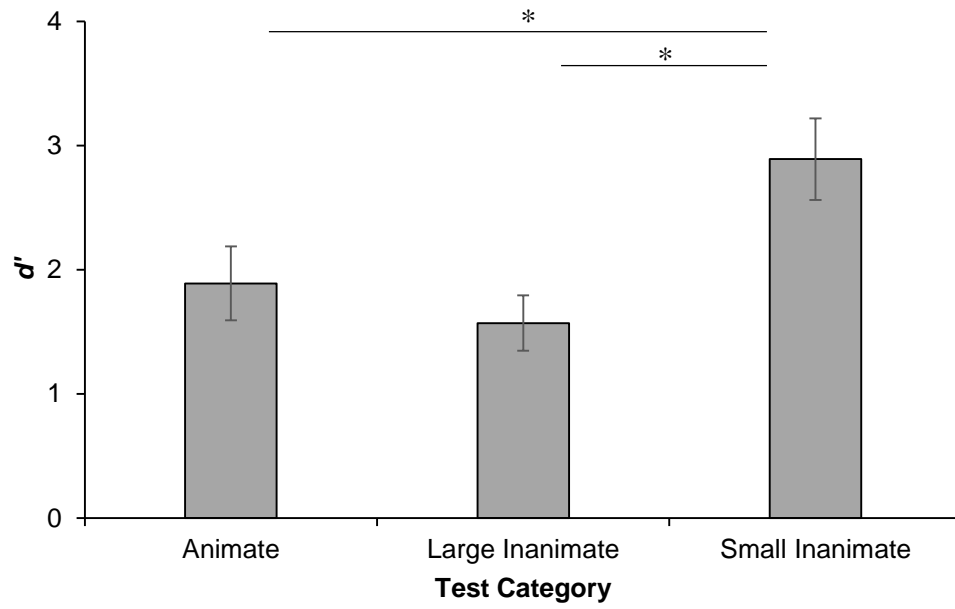


*Figure 8.* Average number of responses for each recognition response type, between stimulus categories for version two of the experiment,  $n = 9$ . Error bars represent standard error.

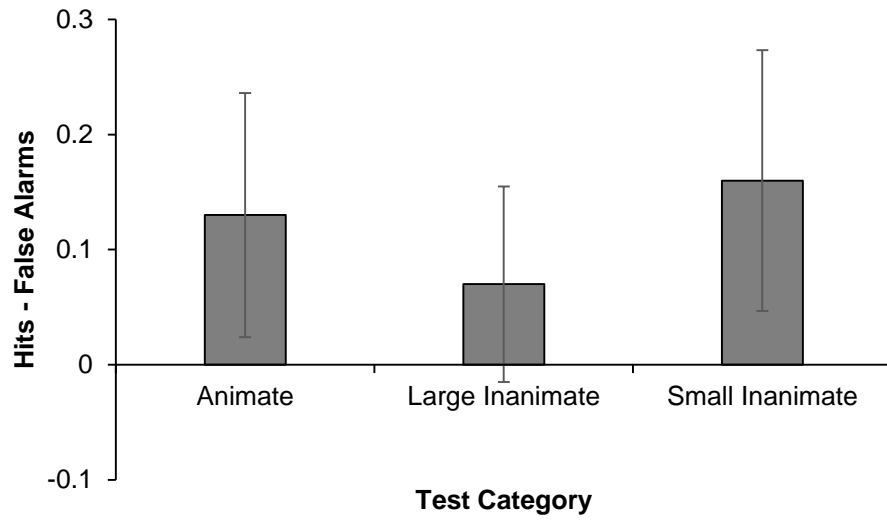




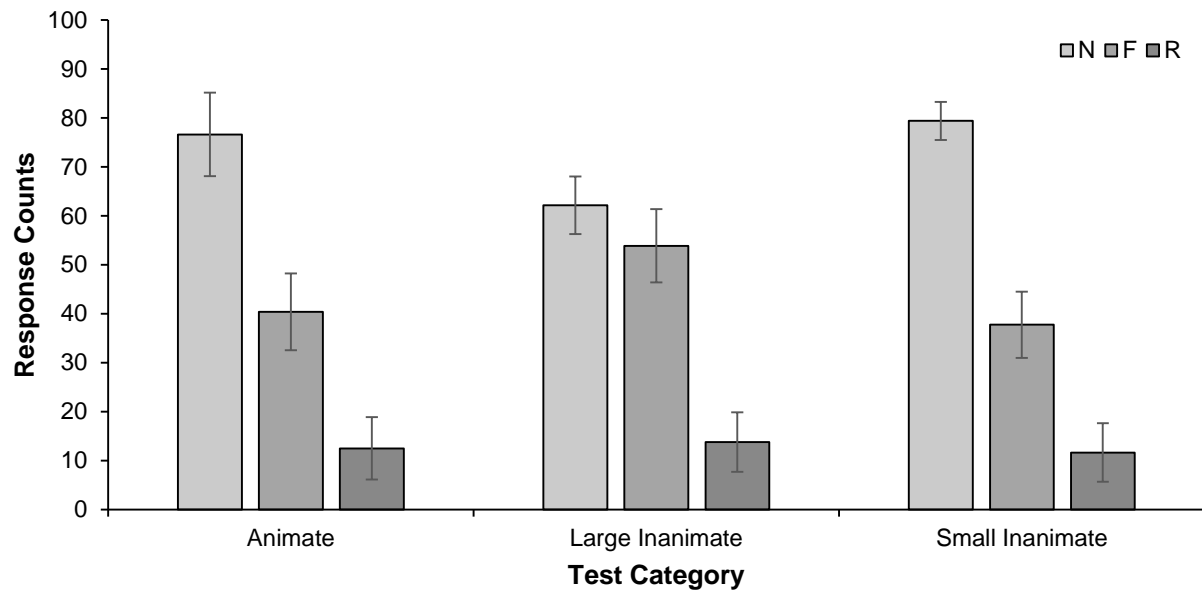
*Figure 9.* Overall recognition performance ( $d'$ ) between stimulus categories for version three of the experiment,  $n = 8$ . Error bars represent standard error.  $*p < .05$ .



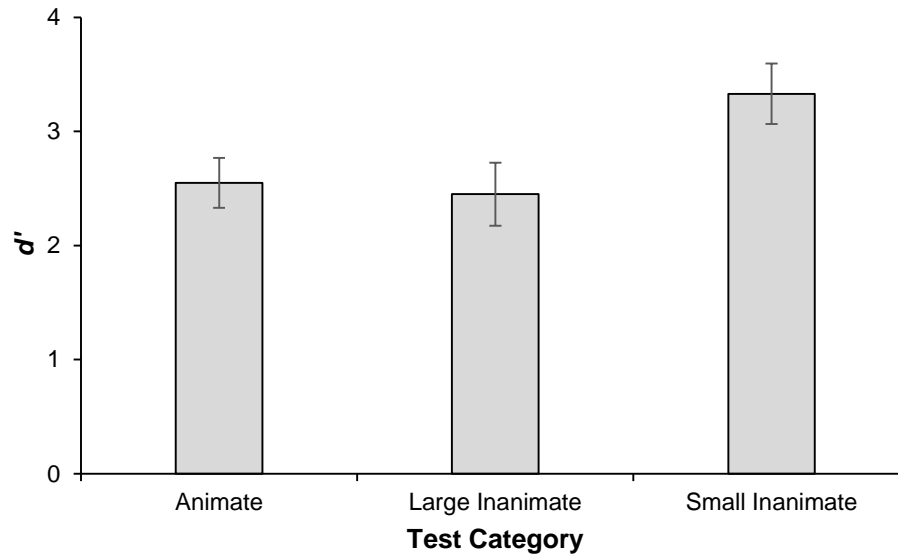
*Figure 10.* Familiarity recognition performance ( $d'$ ) between stimulus categories for version three of the experiment,  $n = 8$ . Error bars represent standard error.  $*p < .05$ .



*Figure 11.* Recall performance as a function of hits minus false alarms between stimulus categories for version three of the experiment,  $n = 8$ . Error bars represent standard error.



*Figure 12.* Average number of responses for each recognition response type, between stimulus categories for version three of the experiment,  $n = 8$ . Error bars represent standard error.



*Figure 13.* Overall recognition performance ( $d'$ ) between stimulus categories for version four of the experiment,  $n = 10$ . Error bars represent standard error.

categories,  $F(2, 7) = 3.55, p = .04$ , Figure 14. A Tukey's HSD test was performed, showing that there was a significant difference between animate and small inanimate performance, between large inanimate and small inanimate performance, but not between animate and large inanimate performance (Figure 14). Recall recognition performance was not significantly different between categories,  $F(2, 7) = 0.05, p = .95$ , Figure 15. The average number of responses for each category and response is displayed in Figure 16.

### **Discussion**

The aim of the current study was to create a recognition memory paradigm that could be used in a future imaging study to characterize the activation of PhC and PrC, in terms of item-based familiarity. Four different versions of the experiment were created by analyzing data and results after each version, and altering the paradigm in relation to what was found. Each version will be discussed in relation to the three main objectives of this study: ensure recognition memory is primarily familiarity-based, ensure overall memory performance is significantly above chance, and ensure performance is matched between stimuli categories.

The first version of the experiment was primarily familiarity-based, memory performance was significantly above chance, and performance was matched between categories. This version fulfilled the three objectives sought in this experiment, however, upon inspection of the data, some changes needed to be made. Firstly, the experiment was too easy, such that some participants scored perfectly on memory performance. Although this may not be a problem in the current study, in a future imaging study, this may have a confounding effect. That is to say, any differences in localization of familiarity signal may be affected by the ease of the study, and by removing the explanation that the ease of the study is associated with the activation of the brain, will make any results found much stronger. Additionally, the ease of the study may be

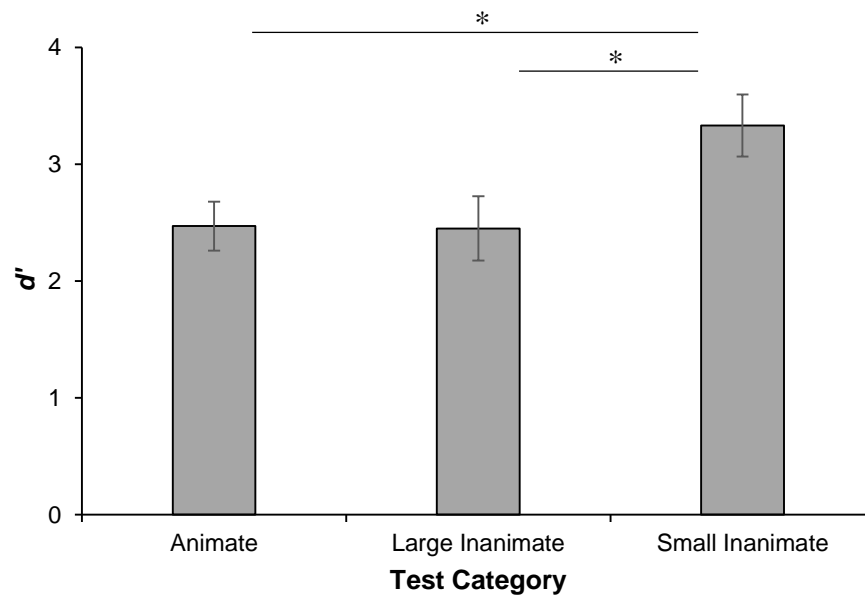
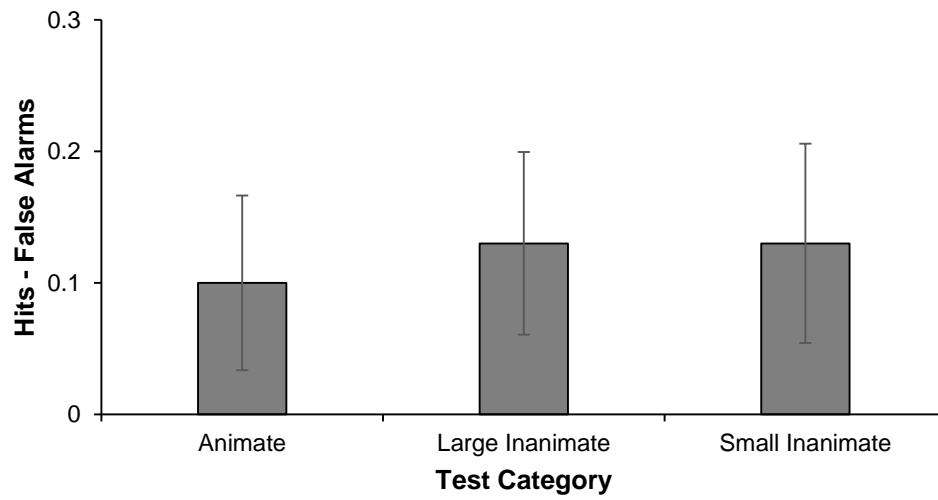
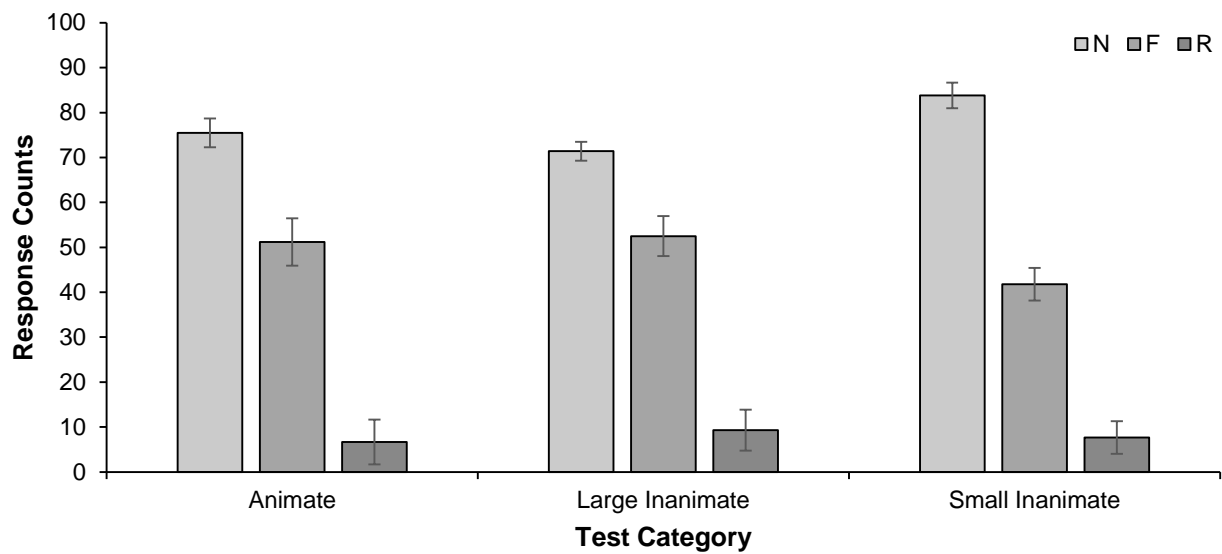


Figure 14. Familiarity recognition performance ( $d'$ ) between stimulus categories for version four of the experiment,  $n = 10$ . Error bars represent standard error.  $*p < .05$ .



*Figure 15.* Recall performance as a function of hits minus false alarms between stimulus categories for version four of the experiment,  $n = 10$ . Error bars represent standard error.





*Figure 16.* Average number of responses for each recognition response type, between stimulus categories for version four of the experiment,  $n = 10$ . Error bars represent standard error.

masking differences in stimulus categories because of the high performance. As well, the recollection responses, while still smaller than that of the familiarity responses, was still larger than anticipated, and larger than the number in the three following versions of the experiment.

These conclusions lead to changes in the paradigm that focused on decreasing memory performance and decreasing recollection responses. The stimuli in two of the categories were changed to make it more difficult. The animate stimuli, which were originally composed of a mix of antlered animals and birds, was changed such that birds were the only stimuli composing the animate category. This was motivated by Konkle and Caramazza's study (2013) which suggested animate items did not differ across real-world size in activation signals in the brain. Having the two subcategories for the animate category was meant to act as a control for size, however, it may have contributed to the ease of recognition because antlered animals and birds are visually, quite different. Birds were chosen to be the animate stimuli because birds appear to have more commonalities amongst them, rather than antlered animals who can appear quite different. The stimuli in the large inanimate category were grouped in each run, into objects that are the same, but were still different items. For example, all instances of inanimate stimuli in the first run were fridges, the second run were all stoves, the third run all couches, etc. This was thought to make memory performance decrease because as the stimuli became more similar, it was harder to differentiate between them.

A final change to the paradigm from the results of the first version was to reduce item presentation time from 2.5 seconds to 1.5 seconds. This was meant to decrease overall performance because the participants had less time to attend, react, and recognize the item, making it more difficult to be sure when an item was novel or recognized. As well, this change was meant to decrease the number of recollection responses because there is less time to see the

item and think about any contextual details in which they remembered the item. This decrease in time would drive the responses to familiarity which is the focus of the study.

The second version of the experiment was primarily familiarity-based, memory performance was significantly above chance, but performance was not matched between categories. Specifically, the large inanimate stimuli category became more difficult and memory performance for familiarity decreased, leading to small inanimate performance being better than large inanimate. This decrease in familiarity memory performance is likely due to the changes done to the paradigm after the first version of the experiment, but the changes may have been too effective, making the memory performance of large inanimate stimuli too difficult. Interestingly, animate stimuli was slightly more difficult, but still matched with small inanimate performance. As well, the change in item presentation time was effective in reducing the number of recollection responses, likely because there was less time to actively recollect when the item was presented, driving the participant to answer based on a feeling of knowing, which is a familiarity response (Eichenbaum et al., 2007).

The changes made from this version were based on the finding that large inanimate stimuli performance for familiarity had become too difficult. This was addressed by changing four of the runs that had appeared unmatched between categories. They were changed by replacing the four runs with entirely new stimuli that was still ordered, like the changes from version one, but the new stimuli appeared more variable and distinct. It was thought that these changes could correct for the decreased performance of the large inanimate category.

The third version of the experiment was primarily familiarity-based, memory performance was significantly above chance, but performance was not matched between categories. Specifically, overall recognition performance was less for large inanimate than small

inanimate stimuli. As well, performance for familiarity recognition was less for animate and large inanimate than small inanimate stimuli. Comparing data from the second version to the third was thought to show an increase in large inanimate recognition performance for both overall and familiarity responses. However, the large inanimate performance still remained quite low and the animate performance decreased, as well. On the basis of familiarity responses, animate and large inanimate were matched categories with performance, and lower performance than in the first version. This is an ideal outcome, however, that some difference still exists between animate and large inanimate categories, with the solution to make large inanimate easier. Once animate and large inanimate are closely matched, small inanimate stimuli can be changed to be matched with the other two categories.

Large inanimate runs were modified by creating subcategories that consist of two different types of large inanimate stimuli. For example, for one run, half of the large inanimate stimuli may be fridges and the other half, tables. This was done for all nine runs. The rationale for this change was that it would remain difficult enough because there was still stimuli that were very similar in each run, but it also had a second subcategory of stimuli that would contrast it. For example, if a participant saw a fridge and a table, and both of them repeated shortly after, it would be easier to recognize them as different and recall seeing them, rather than two fridges or two tables presented closely together.

The fourth version of the experiment was primarily familiarity-based, memory performance was significantly above chance, but performance was not matched between categories. Specifically, animate and large inanimate categories were lower in familiarity-based recognition performance than the small inanimate category. However, animate and large inanimate categories are perfectly matched, therefore, the ideal sets of stimuli to use for both of

these categories has been found for the paradigm. As well, the recollection responses remained quite low, despite the large inanimate category increasing in performance.

The fourth version is the final version of the experimental design, although further refinement is necessary to finish matching across categories. For a fifth version of the experiment, it would be beneficial to decrease overall and familiarity-based recognition performance for the small inanimate category. This may be achieved in a similar manner to that of decreasing performance for the large inanimate category. By creating subcategories among the small inanimate stimuli, with half and half compositions per run, the performance would likely decrease. Once small inanimate recognition performance is matched with the other categories, the paradigm will be ready to be used in an imaging study to characterize the representational structure of PhC and PrC for item-based familiarity recognition signals.

In conclusion, the fourth version of the paradigm created is the best starting point for the future imaging study, and with some minor refinement that will decrease overall and familiarity-based recognition performance in the small inanimate category, the paradigm will be ready to be used. This behavioural study was necessary to ensure that recognition is primarily familiarity-based, that overall recognition is significantly above chance, and that stimuli categories are matched in performance. By refining this experimental paradigm, any conclusions drawn during the imaging study will be completely attributable to the familiarity-based recognition in PhC and PrC, and not attributable to any flaws in the design.

### References

- Eichenbaum, H., Yonelinas, A.P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30, 123-152.
- Konkle, T. & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience*, 33(25), 10235-10242.
- Manns, J.R. & Eichenbaum, H. (2006). Evolution of declarative memory. *Hippocampus*, 16, 795-808.
- Martin, C.B., McLean, D.A., O'Neil, E.B., & Köhler, S. (2013). Distinct familiarity-based response patterns for faces and buildings in perirhinal and parahippocampal cortex. *Journal of Neuroscience*, 33(26), 10915-10923.
- Ranganath, C. (2010). A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus*, 20, 1263-1290.