**Study Information**

1. **Title**

Feature-Binding Errors in Associated Objects (Expt 4)

1. **Authorship**

Paul Scotti, Yoolim Hong, Andrew Leber, & Julie Golomb

1. **Research Questions**

How are features bound to object identities given capacity constraints in visual working memory?

In Experiment 3, we observed that repulsion was significantly greater when the retention interval between study and test was 3 seconds, compared to 1 second. One potential explanation for these results is that the longer delay led to noisier memory representations, which led to greater reliance on relational memory mechanisms, and therefore larger repulsion effects for the longer delay were observed. However, another possible explanation is that the longer delay led to more time for the two memory items to compete in working memory. This active competition account suggests that it was not the fidelity of the memory representations that led to decreased repulsion in the shorter delay. Instead, the amount of competition induced from active maintenance during the retention interval could also account for these results. Finally, a third explanation is that the total duration of the retention interval was critical in the observed repulsion biases present in Expt 3. That is, the delay time may have been important and not the active competition or noise of the memory representations.

In Experiment 4, we propose an experimental manipulation to test these alternative explanations. Specifically, we adopt the same design as Experiment 2 but introduce a filler task (size comparison task) on half of the trials. One second in to the 3-second retention interval, 50% of trials will be interrupted for exactly 2 seconds to complete a filler. After the filler task, the subject is immediately tested using the continuous color report. The remaining half of the trials will be the blank delay condition (total 3s delay before the report). We then aim to test whether reports on the continuous color report similarly exhibit repulsion biases in these two conditions.

The total time between study and test is constant across conditions, but the predicted results differ hypotheses. A filler task during the retention interval is predicted to worsen memory performance (i.e., noisier memory representations). Therefore, the memory noise hypothesis predicts that repulsion would be greater on trials where the participant completed the filler task. At the same time, if we assume that the two memory items were actively competing in the retention interval, a filler task is predicted to add interference to the limited capacity working memory system. Hence, the competition present between study items would be interfered with, resulting in noisier memory representations but without as strong a systematic bias in a certain color direction. The third possibility is that repulsion is equivalent between conditions because the retention interval remains constant throughout the experiment, and it is this delay time that is the critical factor that accounts for increased repulsion observed in Expt 3.

1. **Hypotheses**

Hypothesis A: The filler task condition should increase memory noise (due to increased memory interference & engaging central attention) and hence repulsion should be greater thantheblank delay.

Hypothesis B: The filler task condition should interfere with relational mechanisms active between study items (due to new interference introduced from filler task), and hence repulsion should be reduced relative to the blank delay.

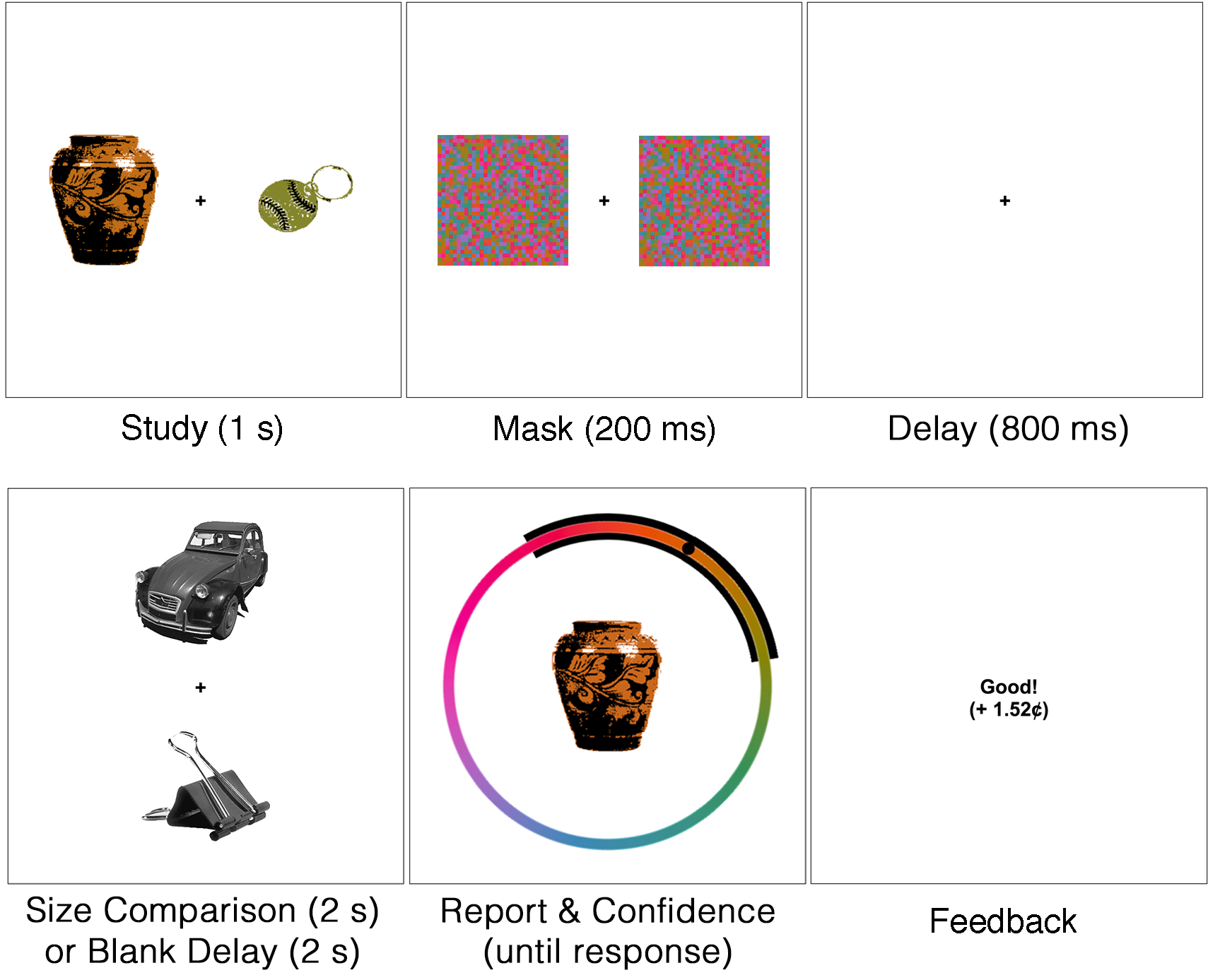
Hypothesis C: The increased repulsion observed in Expt 3 for the 3-seconds condition, compared to the 1-seconds condition, was due to the increased duration of the retention interval. Given that this experiment keeps the delay between study and test constant between conditions, there should be equal repulsion in the filler and blank delay conditions.

**Design Plan**

1. **Study Type**
   1. Psychophysics experiment
2. **Blinding**
   1. All participants that pass exclusionary criteria will be included.
   2. MTurk: Participants’ worker IDs will be discarded prior to analysis to prohibit the ability to trace an individual to their Amazon Mechanical Turk account.

**Design**

1. Below is an example trial procedure. There will be 280 trials total, split into 14 blocks of 20 trials each. Every trial, two unique real-world objects with a single associated color (360 deg. circle in CIE L\*a\*b\* space with coordinates L\*=70, a\*=20, b\*=38, radius 60; Zhang & Luck, 2008) will be displayed. Unknown to participants, the color distance between objects will be either 45 deg. or 90 deg. The retention delay will be 3 seconds total: on half of the trials there will be a blank delay for 3s (blank delay condition); on the other half of trials, a 2-second filler task will onset after a one second delay (filler task condition). The filler task involves two real-world grayscale objects presented above and below the central fixation cross. The participant will be instructed to click on the larger object (according to their real-world size). They will have 2s to complete this task, and then the continuous color report is displayed. Each of these four critical conditions (no-filler/45deg, no-filler/90deg, filler/45deg, filler/90deg) will occur for 70 trials each, and their presentation order will be randomized within each subject. Participants will be probed with a grayscale representation of one of the two objects, and have to select the original color of the object and specify a confidence range (the smallest range of colors they believe contains the correct color; see Chen, Leber, & Golomb, under revision). As the mouse moves around the color wheel, the initially grayscale object will dynamically change to the color closest the mouse pointer. Feedback and bonus are then presented for 750 ms. The next trial then begins after a 500 ms inter-trial interval.
2. The bonus is composed of two parts, with each part awarding a max of 1 cent. The first part is degrees of error (distance from report to the original color). If x is absolute degrees of error, then cents awarded equals *(1 – x/45)*, such that more fractions of a penny are awarded for less deg. of error but nothing is awarded if x>=45. The 2nd part is based on the confidence report, specifically, if y is the confidence range (360 being a highlight of the entire color wheel), then cents awarded equals *(360 - y)/359*, such that smaller intervals award more money (359 because the smallest highlight possible encompasses 1 degree). Except, this only occurs if the highlighted region contains the true original color, if it doesn't, then no bonus is awarded for this part. A negative bonus is never awarded. If the trial included a filler task and the participant responded incorrectly or not at all, then the bonus will be halved.



**Sampling Plan**

1. **Existing Data**
   1. As of the time of submission of this research plan for preregistration, no data have been created, collected, or analyzed.
2. **Data Collection Procedures.**
   1. *Timeline.* The study will take place during the Spring 2019, with the aim to finish the proposed experiment by Summer 2019.
   2. *Participants.* Subjects will be recruited through Amazon Mechanical Turk and will receive $6.00 an hour as compensation, plus a bonus of up to $5.60 depending on performance.
3. **Sample Size**
   1. Expt. 4: Amazon Mechanical Turk: 81 participants.
4. **Sample Size Rationale**
   1. We adopt the same sample size as Experiment 3 (see preregistration for Experiment 3).

**Exclusion Criteria**

1. **Subject exclusions**
   1. Exclusion will depend on individual subject fits to a model that includes all memory reports. We plan to exclude subjects if their proportion of random guessing exceeds .50. These exclusion criteria are based off the exclusion criteria in Golomb et al. (2014) and Golomb (2015). Participants will also be excluded if their accuracy for filler tasks was below 75%.

**Analysis Plan**

1. **Probabilistic Mixture Modeling**
   1. Using probabilistic mixture modeling, memory responses will be characterized according to one of four underlying distributions: a target distribution (responses around the correct color), a swap distribution (*S*: responses around the distractor object’s color), a swap-comparison distribution (*C*: responses around the color in the opposite direction to the distractor object’s color), and a random guessing distribution (uniform responding across all colors). Errors made towards the distractor object’s color will be assigned a positive sign, and errors made away from the distractor object’s color will be assigned a negative sign. In this way, we may observe a mean shift in the target distribution wherein responses are either towards or away from the distractor object’s color. We can therefore characterize shift errors (attraction or repulsion from the distractor color; 0) as well as swap errors (misremembering the original color as the distractor color; *S* > *C*).
   2. The probability distribution can be expressed as

where is the difference between the reported and correct color values, is the proportion of trials on which the participant responded at random, is a von Mises distribution with mean , , or , and concentration or (standard deviation ), is the proportion of trials on which the participant responded around the distractor object’s color (von Mises distribution with mean , the distance from the original color to the distractor color, and concentration ), and comparably is the proportion of trials on which the participant responded around the color in the opposite direction to the distractor object’s color.

* 1. For each participant, we will separately model responses for trials containing objects that were 45 deg. apart in color space, and trials containing objects that were 90 deg. apart in color space. We will also model responses for all trials regardless of distance in color space. We will use standard t-tests and ANOVAs to compare maximum a posteriori estimates between conditions.
  2. In addition to the above model fits, we will separately model memory errors across subjects after splitting each subject’s data into their most and least confident memory responses, comparing model parameters between confidence levels. For example, for the 45-degrees model, we will measure each subject’s median confidence range across trials that had objects 45 deg. apart in color space. Each subject’s trials will be divided into more and less confident halves, and a model will be separately fit to each half.
  3. We may, in addition, use a Bayesian hierarchical model to fit memory responses. We can then attain both group-level and subject-level parameter estimates and use 95% highest posterior density intervals (HDIs) to determine significance (Kruschke, 2011).

1. **Follow-up Analyses**

We may conduct supplementary analyses not listed in this document, for instance, to explore subjects with poor fits, the importance of color category, and alternative explorations into how confidence reports relate to different types of memory distortions. We will report any additional analyses as exploratory.