# The Influence of Ensemble Statistics and Focused Attention on Feature Perception

Undergraduate Research Thesis

Presented in Partial Fulfillment of the Requirements for graduation "with Honors Research Distinction in Neuroscience" in the undergraduate colleges of The Ohio State University

by
Jane Elyse Beaufore

The Ohio State University
April 2017

Project Advisor: Dr. Julie Golomb, Departments of Psychology and Neuroscience

**Acknowledgements**

I would like to thank Dr. Julie Golomb, Jiageng Chen, and everyone in the OSU Vision and Cognitive Neuroscience Lab for all their help and guidance in the pursuit of this thesis and in my undergraduate education. Thank you to Dr. Angela Brown and Dr. Alex Petrov for serving on my thesis committee and providing valuable feedback on this project. Additional thanks are due to all of my teachers, friends, and family for their ongoing support and encouragement.

## Abstract

Ensemble statistics are a means by which our brains can take the average across features such as color (De Gardelle & Summerfield, 2011), size (Chong & Treisman, 2003), and texture (Alvarez & Oliva, 2009) to help us retain the gist of a scene without relying on effortful processing of every object. Previous research has also shown that features such as color appear to be mixed together when attention is split (Golomb et. al, 2014). We wanted to examine how focused attention versus split attention would influence the use of ensemble statistics and feature mixing. To test this, subjects fixated on the center of a computer screen and three squares appeared in their periphery. Either all three of the squares became bold (a split-attention trial) or just one became bold (a focused-attention trial), and subjects were asked to direct their attention to the bold square(s). Each square was then briefly filled with a different color, and subjects were asked to report the color in one of the squares by choosing between two response colors (Experiment 1) or clicking along a color wheel (Experiment 2). Experiment 1 did not show significant effects of the ensemble mean on either attention condition, but this may have been due to a less sensitive experimental design. In Experiment 2 we could fit the distribution of responses with probabolistic models, allowing us to examine differences in guess rate, precision, swapping errors, and shifts toward the ensemble mean. Focused-attention trials showed a decrease in overall noise, with lower standard deviations (precision) and guessing rates. Focused-attention trials also showed a decreased frequency of "swap" errors (reporting one of the distractor colors instead of the target color). We found some evidence that subjects tended to report a color shifted toward the ensemble mean, but there was no significant difference in this measure for focused-attention trials and split-attention trials. These results suggest that ensemble statistics can bias the perception of features by drawing them toward the ensemble mean, and focusing spatial attention does not appear to modulate this effect. Future experiments will test whether this trend persists for larger ensembles containing more objects and different distributions of distractor colors.

## Introduction

When we look at a scene, we must perceive and encode a multitude of different objects and their features. However, it is difficult to accurately encode every feature of every object, so we must employ a mechanism that allows us to retain the overall gist of a scene without relying on effortful processing of every object. One such mechanism is called ensemble statistics, and it is a means by which our brains can take the average across different features such as color (De Gardelle & Summerfield, 2011), size (Chong & Treisman, 2003), and texture (Alvarez & Oliva, 2009) when observing a scene so that our perceptions of a feature of a single object are influenced by the features of the objects around it. For example, if we view a flowerbed containing several red-ish, blue-ish, and yellow-ish flowers, our perception of an individual flower's color may be slightly shifted from its actual color because we take into account the colors of the surrounding flowers.

One seminal study of this effect asked subjects to observe a scene containing circles of different sizes. The circles then disappeared, and subjects were asked to report the size of one individual circle within the scene. Subjects showed a tendency to report a size that was shifted toward the average size across all circles (Brady & Alvarez, 2009). This suggests that ensemble statistics affect feature perceptions when attention is divided across objects in a scene.

But what about when our attention is focused on a specific object instead of divided across the entire scene? Will the features of surrounding objects still influence our perceptions, and how?

Previous research has described spatial attention as an important factor for perceiving objects. The less attention we pay to objects, the harder it becomes to encode their features (Palmer, 1990). Spatial attention acts as a potential key to The Binding Problem, which asks how our brains are able to correctly bind features to objects to obtain a coherent depiction of our environment (Triesman & Gelade, 1980). The more attention an object receives, the easier it becomes to correctly perceive and report its features. When objects are outside our frame of attention, or when we split our attention across objects, we make more mistakes.

For example, when we divide our attention across two objects in a scene, we are prone to “mixing errors” when reporting the color of one of those objects (Golomb, L’Heureux, & Kanwisher, 2014). This type of error means that the reported color is a blend of the target color and the distractor color. If we are splitting our attention across several objects in the scene and our brains are enacting an ensemble representation of the features, we might expect these mixing errors to be biased toward the *average* across all colors in the scene.

When we attend to a *single* object, on the other hand, we are more likely to temporarily exclude stimuli from other locations. It has been shown that when an object in a scene is cued before the scene is displayed, subjects make fewer binding errors when asked to report a feature of that object later on (Treisman, 1998). This indicates that focused attention can reduce the frequency of feature-binding errors. Thus, we might expect to see a decrease in mixing errors (ensemble bias) when attention is focused on a single location in advance.

We studied how the interaction between spatial attention and ensemble statistics could influence feature perception. When attention is split across objects in a scene, would the perception of a single target color be biased toward the average across all colors in the scene? When people are told in advance to focus on a certain location within the scene, would they be able to effectively suppress the colors of other objects in the ensemble, thus reducing the colors' influence on the person's perception?

To address these questions, we ran two experiments. In the first experiment, subjects were briefly presented with an ensemble consisting of three differently colored objects and were asked to attend to either all three of the objects or just one. Subjects were then asked to report the color of one of the objects by choosing between two response colors. The second experiment followed the same setup, but subjects were asked to report the color of one of the objects by choosing a color along a color wheel (this allowed us to see the direction and degree of any biases). A proposed third experiment will test the effects of focused versus split attention on feature perception when subjects are exposed to an ensemble of five objects.

## General Methods

**Experimental Setup.** Experiments were coded using MATLAB and PsychToolbox (Brainard, 1997; Pelli, 1997) with assistance from graduate student Jiageng Chen. Subjects were taken from the Ohio State University Research Experience Program and ranged from 18-40 years of age with normal or corrected-to-normal vision.

Each subject was tested individually in a dimly lit eye-tracking room. Subjects were first led to the eye-tracking room where the experiment was explained to them and they were given the chance to ask questions about the procedure. They then took a seat and rested their head on a chin rest positioned 60 centimeters from the computer monitor to reduce their head movements, after which their eye position was calibrated with an

Eye-Link 1000 (SR Research) infrared eye tracker. The eye-tracker tracked the subjects' eyes throughout the experiment to ensure that they stayed fixated on a central point. If subjects failed to maintain their gaze on the fixation point, a pop-up message with the phrase “fixation error” appeared, reminding them to return to the fixation point, at which point the trial would restart.

Further details for each experiment are outlined in subsequent sections.

## Experiment 1

In the first experiment, subjects fixated on the center of a computer screen and three squares appeared in their periphery equally spaced from the fixation dot. Either all three of the squares became bold (a split-attention trial) or just one became bold (a focused-attention trial), and subjects were asked to direct their attention to the bold square(s). Each square was then briefly filled with a different color, and subjects were asked to report the color in one of the squares by choosing between two response colors Figure 1 shows a sample trial.

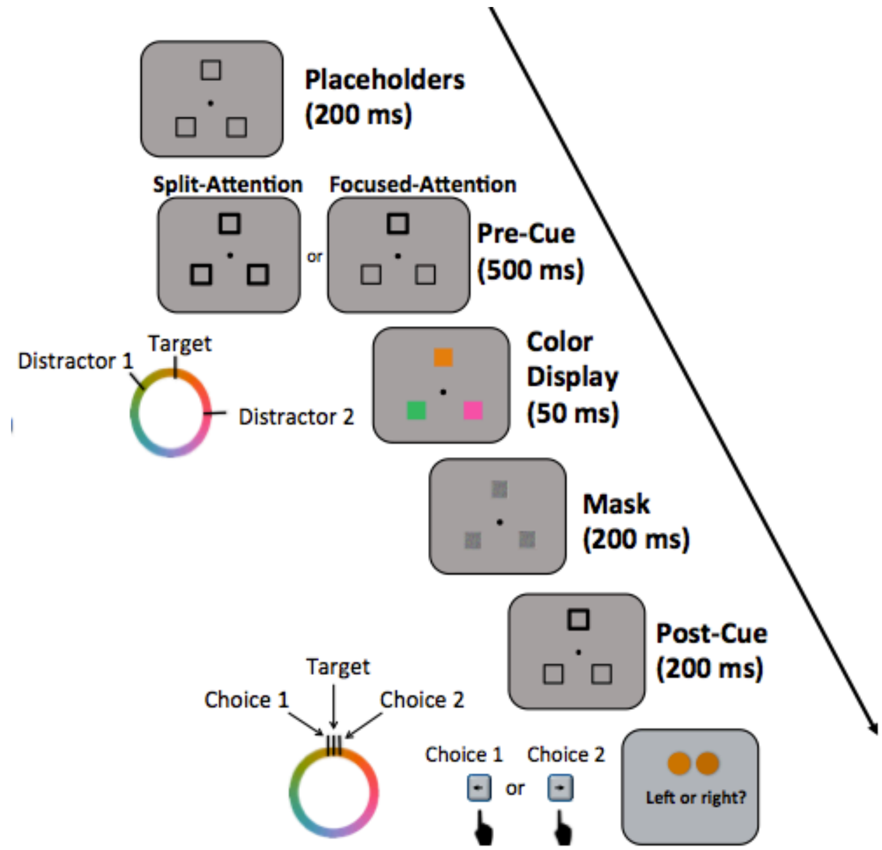


Figure 1: Sample trial for Experiment 1

**Participants.** Data was collected from 8 subjects for this experiment (mean age 25 years, age range 19-40 years, 3 females, 5 males). All subjects were recruited from the Research Experience Program at The Ohio State University and were compensated with course credit. All subjects reported normal or corrected-to-normal vision.

**Methods.** Subjects first ran through a practice block consisting of 96 trials to get them used to the task. Once it was determined that subjects understood the task and the eye-tracker was effectively tracking their eye movements, they began the experiment. The experiment consisted of 8 blocks with 96 trials per block. For each trial, subjects fixated on a small black dot in the center of the computer screen (Figure 1). Once their fixation was steady, three empty black placeholder squares (2° x 2°) appeared in their periphery evenly spaced from the fixation dot (3.5° eccentricity). After a 200 ms delay, either all three of the squares became bold (split-attention pre-cue), or a single square became bold (focused-attention pre-cue) for 500 ms. Each square was then simultaneously filled with a different color, and this color display persisted for 50 ms. One of these three colors was considered the target color (the color that subjects were asked to report later on). The target color was chosen at random for each trial and was selected from 180 possible colors, which were evenly distributed along a circle in Commission Internationale de l'Éclairage (CIE) L*a*b* color space (Zhang & Luck, 2008). One of the other two colors (distractor colors) was shifted by ±30° from the target color in color space (with positive being clockwise and negative being counterclockwise on the color wheel). The remaining distractor color was shifted 90° from the target color in the opposite direction of the other distractor color. In each block, half of the trials displayed distractor colors positioned at -30° and +90° from the target color, and the other half displayed distractor colors positioned at +30° and -90° from the target color. The colors then disappeared and were replaced with a mask for 200 ms. Following the disappearance of the mask, a post-cue was presented for 200 ms in which one of the three squares became bold, indicating the target color to report. Subjects were then presented with a response choice that contained two colors, and were instructed to choose which of the two colors they thought was most similar to the target color.

The response choice was centered on the location of the post-cued square (to minimize the Simon Effect - see Appendix for more details), and contained two options: one color was shifted slightly in one direction from the target color in color space, and the other was shifted by the same amount in the opposite direction. Neither was an exact match to the target color. This forced subjects to choose a color that was slightly shifted in one direction; the goal was to determine in which direction (if any) a bias existed. I.e., if the true color was red, but subjects were biased toward the ensemble mean color (blue), they might be more likely to report the slightly purple-red choice over the slightly orange-red choice. Both choices were shifted 4° in color wheel space from the target color, and subjects were instructed to make their best guess if they felt unsure.

There were two different attention conditions for this experiment. The first condition was a *split-attention* condition in which all three squares became bold during the pre-cue, and subjects were instructed in advance to try to remember the colors in all three squares, since they would be asked to report the color in one of these three squares later on. In this case, the square that became bold in the post-cue was randomly chosen. The second condition was a *focused-attention* condition in which just a single square became bold during the pre-cue, and subjects were instructed that they would need to focus their attention on that square since that square would always be the one that was post-cued (contain the target color to report).

**Analyses.** For each trial, we recorded whether the subject chose the response color shifted *toward* the ensemble mean (average across all colors) or the response color shifted *away from* the ensemble mean. We then used this data to calculate the proportion of trials in which subjects reported the color shifted toward or away for the split-attention condition and for the focused-attention condition. We ran one-sample t-tests for each condition to determine if subjects were reporting a color shifted toward or away from the ensemble mean at a rate significantly different than chance (50%). We also ran paired t-tests to determine if there was a significant difference between split-attention trials and focused-attention trials.

**Results.** The experiment was designed so that on each trial, the average color across the three colors lied at a color value that was slightly shifted in either the clockwise or counterclockwise direction (along a color wheel) compared to the target color. We expected that subjects would be sensitive to this "ensemble" information and would consistently choose a color that was shifted toward the ensemble mean for split-attention trials but would not be as influenced in focused-attention trials (resulting in a roughly 50/50 split for which response color was chosen). However, paired t-tests showed that responses for focused-attention versus split-attention trials were not significantly different ($t(7) = 0.988$, $p = 0.283$). When comparing the results of focused-attention trials to chance, the result was not significant ($t(7) = 0.988$, $p=0.356$). The same was true when comparing the results of split-attention trials to chance ($t(7)= -0.773$, $p=0.465$). Figure 2 shows the average proportion responses reporting a shift toward or away from the mean across all three colors for both trial types.

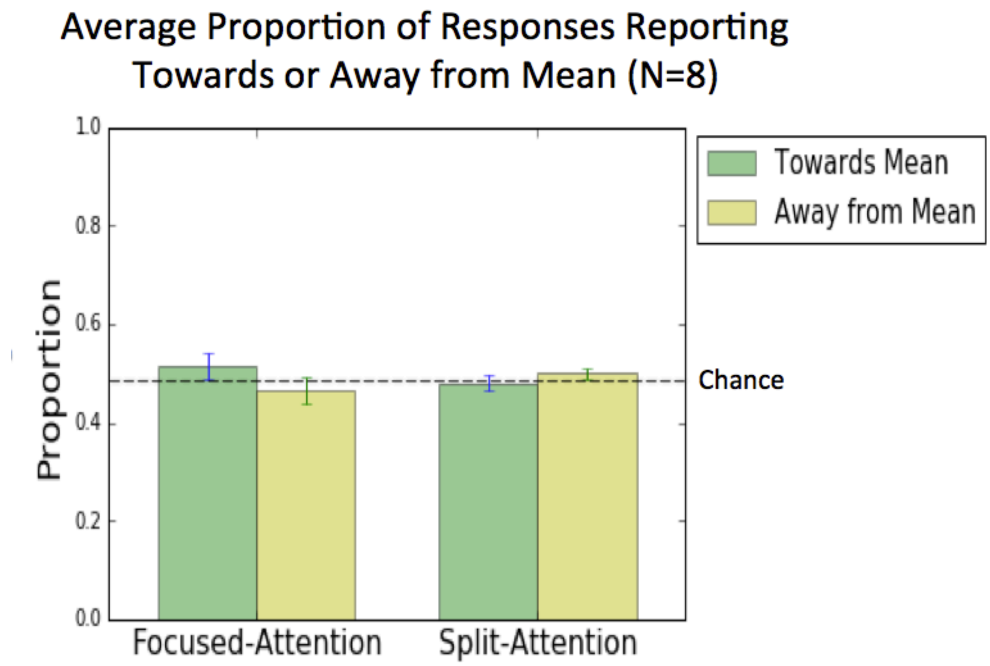


Figure 2: Average proportion of responses reporting a shift toward or away from the mean for each trial type

Figure 3 shows the difference between the proportions of responses that reported a shift toward the ensemble mean versus away from the ensemble mean for each individual subject in both experimental conditions (split-attention and focused-attention). Positive values indicate a shift toward the ensemble mean, and negative values indicate a shift away from the ensemble mean. Three subjects (1, 6, and 8) showed a larger proportion of focused-attention trials in which they reported the color shifted toward the ensemble mean. Notice how the red dots (signifying focused-attention trials) lay at larger values for these subjects.

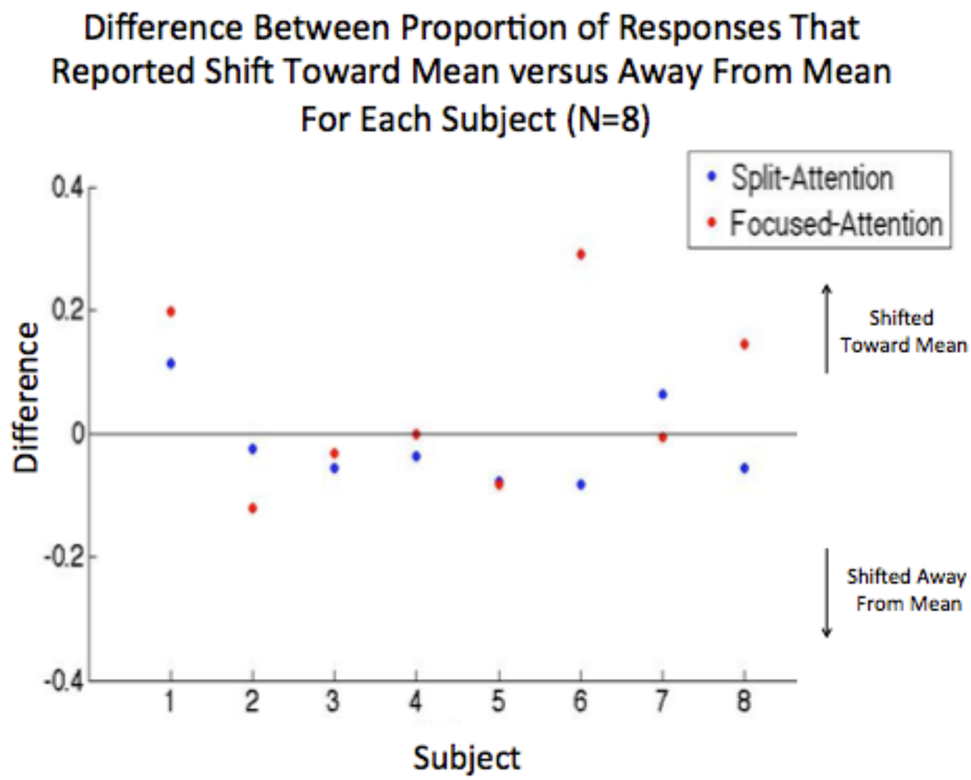


Figure 3: Difference between proportion of responses reporting toward mean versus away from mean for focused-attention and split-attention trials for individual subjects

**Discussion.** Contrary to our hypothesis, we found that subjects were not more likely to choose a color shifted toward the ensemble mean on split-attention trials. It was also shown that subjects were not choosing the color shifted toward or away from the ensemble mean at a greater rate than chance in either trial type. The overall trends suggest that spatial attention does not significantly modulate a subject's perception of a target color within an ensemble, and subjects are not drawn toward the ensemble mean at a significant rate.

However, limitations of the experimental design may have interfered with the experiment's ability to effectively address our research question. We noticed that both focused-attention and split-attention trials had roughly a 50/50 split in the frequency at which each response color was chosen; this null effect could have meant that there was no ensemble bias in either condition, or it could have meant that subjects were simply performing at chance. Because of the format of the response cue, we couldn't tell if subjects were performing the task well with no bias, or randomly guessing. Many subjects expressed that they felt the task was difficult, and that each of the response colors often looked so similar that it was hard to tell the difference between the two. Both of these factors may have caused subjects to guess at which color was correct in the response cue, regardless of whether it was a split-attention or focused-attention trial. Increasing the duration of the color display could cause the task to become more of a memory task than a perceptual task, however, and making the response colors more distinct may also make the task more difficult if neither color seems to match the target color at all. The next experiment addresses the problems associated with the response cue by changing it to a continuous report.

## Experiment 2

Experiment 1 used a forced-choice paradigm in which subjects were asked to choose between two different response colors in order to see if their responses in focused-attention and/or split-attention trials were biased toward the average across all colors. Since subjects often had trouble distinguishing between the two response colors in Experiment 1, we ran a second experiment that followed the same setup, but instead asked subjects to report the target color by clicking along a color wheel (a continuous report). We then used mixture models (Zhang & Luck, 2008) to analyze subject performance. These models can show whether subjects are reporting a color shifted toward the ensemble mean, and can account for random guessing of the target color as well as swapping errors (reporting a distractor color instead of the target color). This allows for differentiation between those who are randomly guessing versus those who are performing the task well but not showing a shift toward the ensemble mean (factors we could not differentiate between in Experiment 1).

**Participants.** Data was collected from 20 subjects for this experiment (mean age 19 years, age range 18-24 years, 11 females, 9 males). All subjects were recruited from the Research Experience Program at The Ohio State University and were compensated with course credit. All subjects reported normal or corrected-to-normal vision.

**Methods.** The overall setup of Experiment 2 was the same as Experiment 1 with the exception of the response cue. Experiment 2 consisted of 8 blocks with 96 trials per block, and colors were distributed and randomized in the same way as Experiment 1. 20 subjects completed Experiment 2, and each subject finished between 5 and 8 blocks. Figure 4 shows a sample trial for Experiment 2 in which subjects report their response by using the mouse to click along a color wheel.

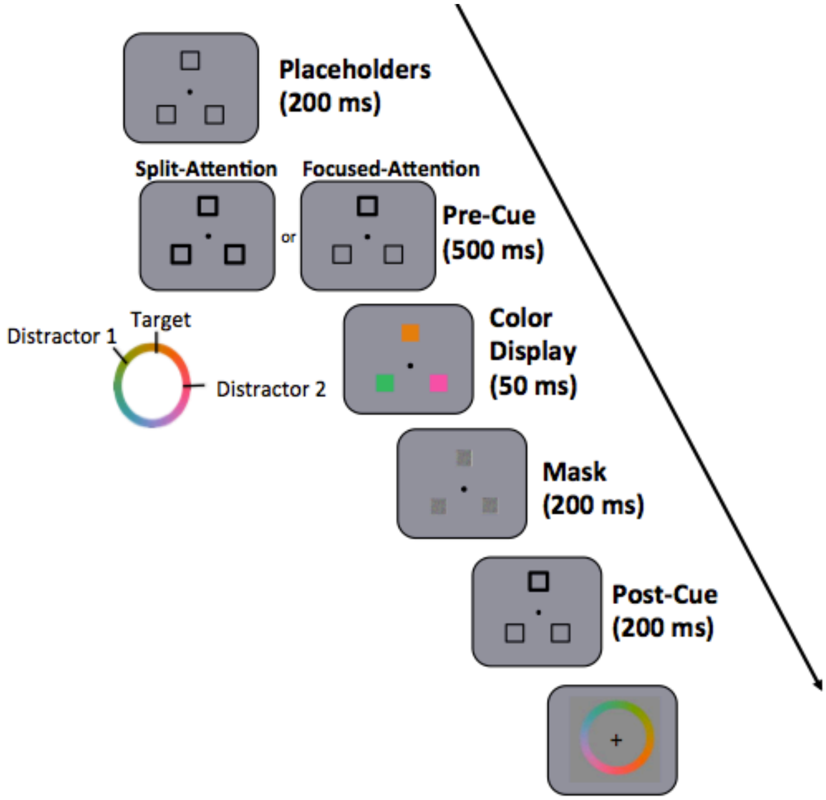


Figure 4: Sample trial for Experiment 2

**Analyses.** Subjects reported the target color by clicking along a color wheel, and their responses were recorded and converted into a difference score in degrees of angle along the color wheel. The target color was set to 0°, and data was aligned so that distractors were always set at -90° and 30°.

We used two probabilistic models to fit the data and account for sources of error. These models were mixture models which assumed that throughout the task, there were some trials in which subjects correctly perceived the stimulus (generating a Gaussian distribution around the correct response), some in which subjects randomly guessed (generating a uniform distribution), and some in which subjects made swapping errors (where they reported one of the distractor colors instead of the target color). The distribution centered on the correct response (the target color) has a $\mu$ parameter that indicates how much the distribution has shifted from the target color. In an ideal case in which subjects perceive and report the target color perfectly on each trial, we would expect μ to have a value of zero. The μ parameter is our primary measure of interest when examining whether subjects are showing a shift toward the ensemble mean.

The first model used was a standard swap model (Brainard, 1997), which accounted for swapping errors (misreports of distractor colors) by allowing for one Gaussian distribution centered on the target color as well as one centered on each distractor color (both distractor colors were assigned the same weight). This model is defined by the equation:

$$p(\theta) = (1-\beta-\gamma)\ \phi_{\mu,\kappa} + 1/2(\beta\phi_{d1,\kappa}) + 1/2(\beta\phi_{d2,\kappa}) + \gamma(1/2\pi)$$

Where $\theta$ is the difference in radians between the reported color and the target color, d1 is the distractor color at -90°, d2 is the distractor color at 30°, β is the probability of misreporting one of the distractor colors (defined by a von Mises distribution with a fixed mean centered on the attended distractor color value), γ is the probability of random guessing, μ is the mean of the primary von Mises distribution, and κ is the concentration of the distributions.

The second was a modified swap model, which fit one Gaussian distribution centered on the target color, as well as one centered on each distractor color. In this model, distributions centered on each distractor color could be assigned different weights depending on whether or not one distractor color had more influence on the overall data than the other. This model is defined by the equation:

$$p(\theta) = (1-\beta-\delta-\gamma)\ \phi_{\mu,\kappa} + \beta\phi_{d1,\kappa} + \delta\phi_{d2,\kappa} + \gamma(1/2\pi)$$

Where γ is the probability of random guessing, d1 is the distractor color at -90°, d2 is the distractor color at 30°, β is the probability of misreporting the distractor color value (defined by a von Mises distribution with a fixed mean centered on the attended distractor color value), δ is the probability of misreporting the control color value (defined by a von Mises distribution with a fixed mean centered on the control distractor color value), μ is the mean of the primary von Mises distribution, and κ is the concentration of the distributions.

Using an AIC model comparison (Akaike information criterion: Akaike, 1974) that took into account the number of parameters in each model, we determined the model of best fit for each individual subject. Additional details about the model comparison results are outlined in the next section.

One-sample t-tests were run for each trial type to determine if subjects were reporting a color shifted toward the ensemble mean at a rate greater than chance (μ compared to 0), and paired t-tests were run to determine if mean distribution shifts (μ), standard deviations ($\sqrt{1/\kappa}$), proportions of swapping errors (β), and guessing rates (γ) were significantly different between conditions.

**Results.** The histograms in Figure 5 show the average response distribution across all subjects in each trial type. Responses are recorded as the difference in degrees from the target color (centered at zero), and data is fit using the standard swap model. Figure 5a shows responses on focused-attention trials and Figure 5b shows responses on split-attention trials. Distractor colors are set at -90° and 30° (the average across all colors lies in the negative direction).

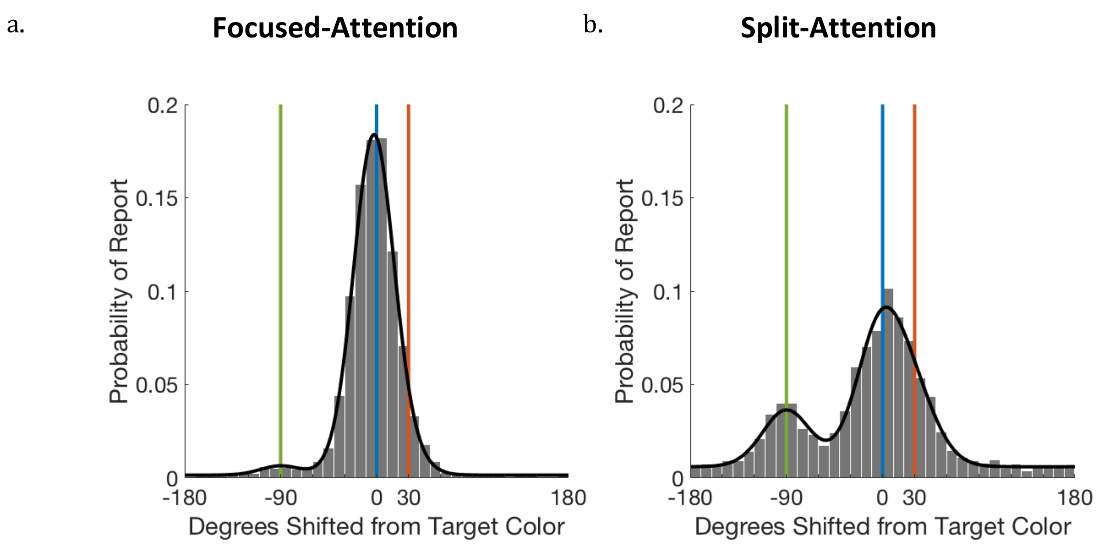


Figure 5: Group data for response distribution in split-attention and focused-attention trials fit with the standard swap model

Figure 6 shows the best-fit parameters from the standard swap model for both conditions across all subjects (n=20). Figure 6a shows the average degree of distribution shift (μ) when reporting the target color. Data is aligned so the average across all colors lies in a negative direction from the target color (negative distribution shifts indicate a tendency to report a color shifted toward the average across all colors). We found a significant tendency for subjects to report a color shifted toward the ensemble mean on both focused-attention trials ($t(19) = -3.013$, $p = 0.0012$) and split-attention trials ($t(19) = -3.73$, $p = 0.0014$), and these two trial types were not significantly different ($t(19) = 0.652$, $p = 0.518$).

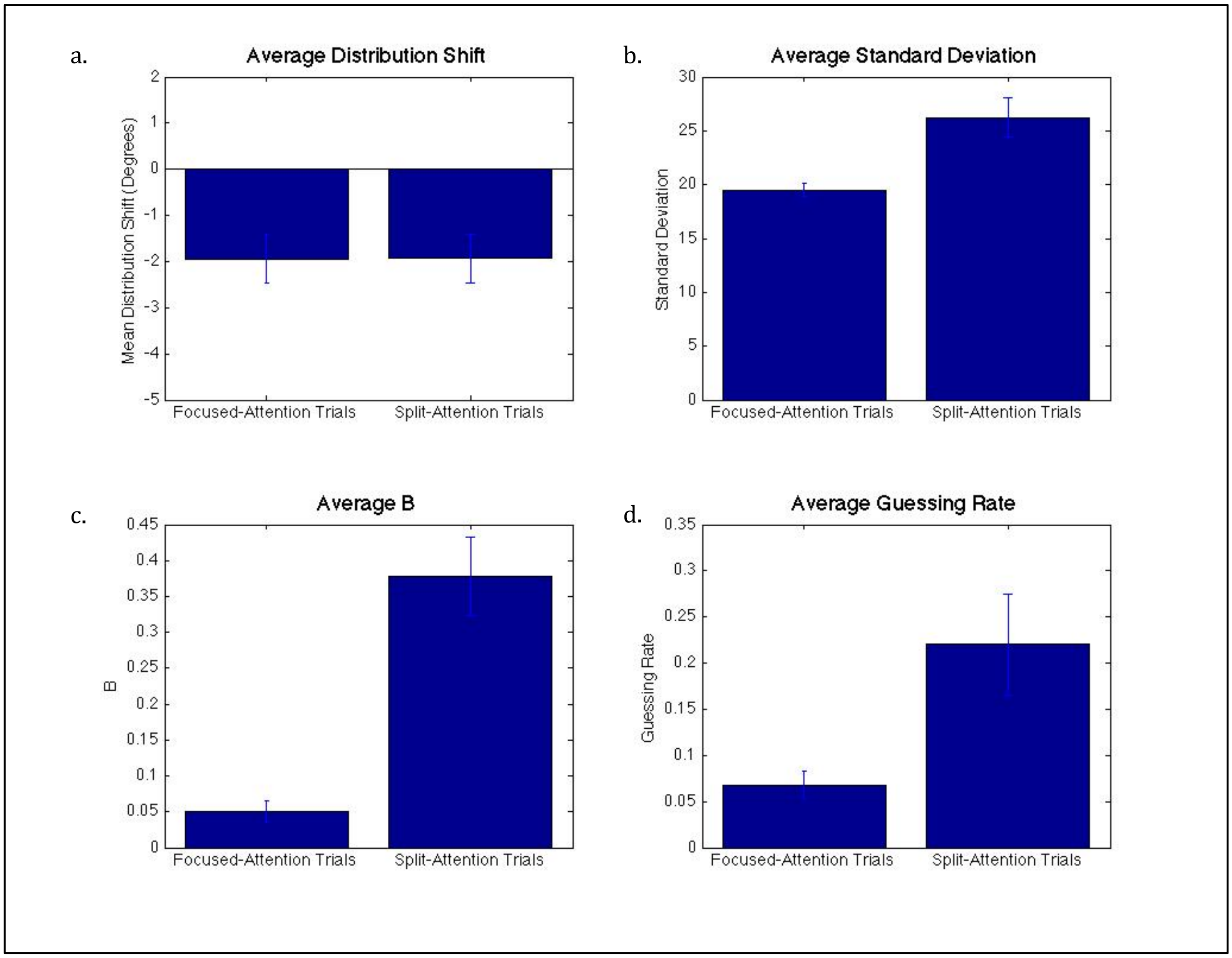


Figure 6: Average distribution shift, average standard deviation, average B, and average guessing rate across all subjects for focused-attention and split-attention trials using the standard swap model

While focused-attention and split attention trials did not significantly differ from each other in measures of the average distribution shift, the average standard deviation based on the standard swap model ($\sqrt{1/\kappa}$), was found to be significantly lower on focused-attention trials than split-attention trials ($t(19) = -4.250$, $p < 0.001$). Smaller standard deviations indicate a higher precision distribution. Figure 6b shows the average standard deviation across all subjects in each trial type.

Subjects also showed significantly fewer swapping errors on focused-attention trials than split-attention trials ($t(19) = -6.792$, $p < 0.001$) . Figure 6c shows the average B value (the probability of misreporting one of the distractor colors as the target color) for each trial type. Higher values for B indicate more swapping errors.

Finally, focused-attention trials showed significantly smaller guessing rates than split-attention trials ($t(19) = -3.930$, $p < 0.001$), indicating that subjects were more precise in their reports of the target color in focused-attention trials (Figure 6d). Guessing rates are expressed as a proportion of total trials in which subjects were guessing the target color (as estimated by the best fit of the proportion accounted for by the uniform distribution).

Using the modified swap model, we found a significant tendency for subjects to report a color shifted toward the average across all colors on focused-attention trials, ($t(19) = -7.338$, $p < 0.001$) but not split-attention trials ($t(19) = -0.582$, $p = 0.568$). However, these two trial types were not significantly different from each other ($t(19) = 2.040$, $p = 0.0555$). Figure 7a shows the average degree of distribution shift when reporting the target color across all subjects (n=20). Data is aligned so the average across all colors lies in a negative direction from the target color (negative distribution shifts indicate a tendency to report a color shifted toward the average across all colors.

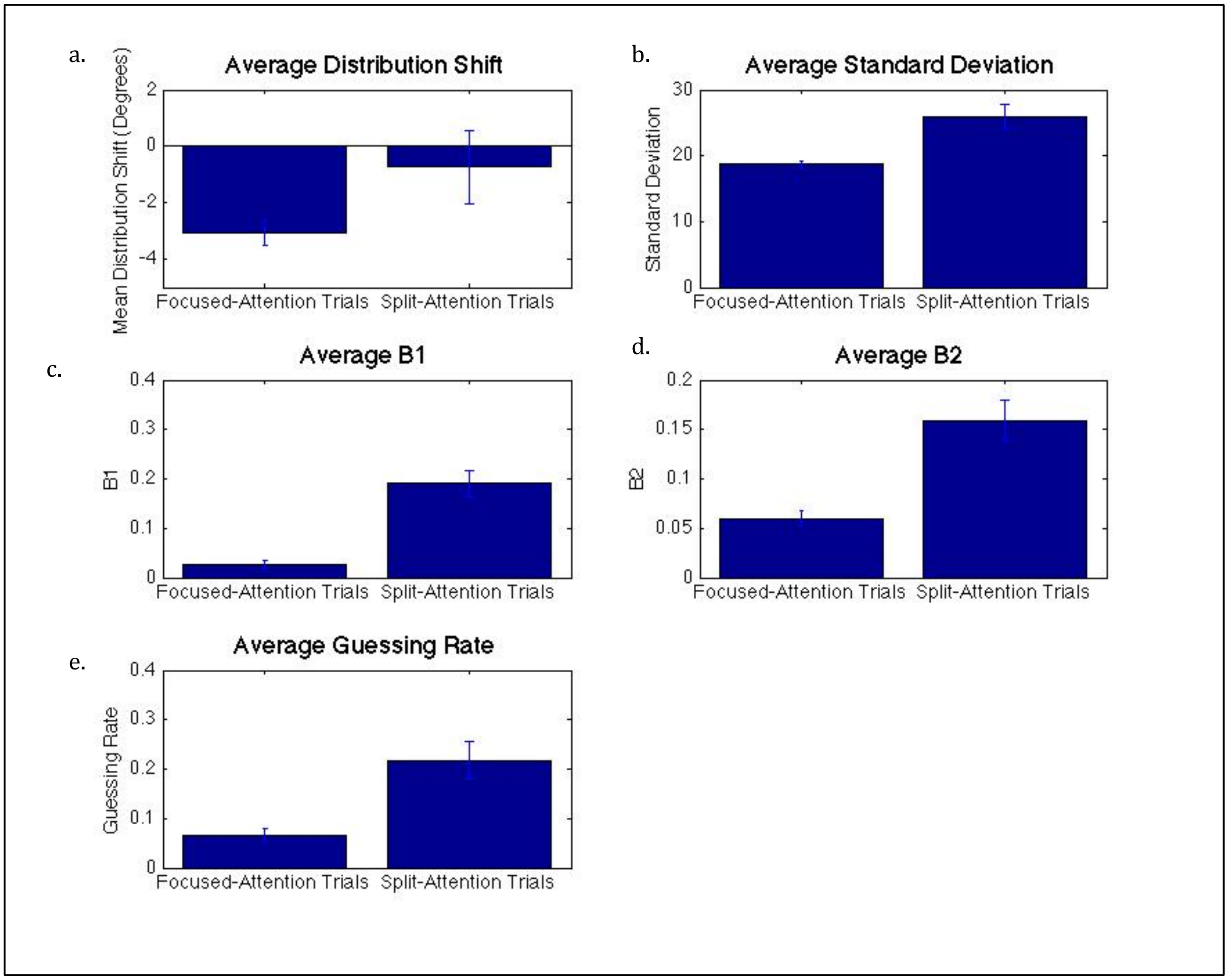


Figure 7: Average distribution shift, average standard deviation, average B1, average B2, and average guessing rate across all subjects for focused-attention and split-attention trials using the modified swap model

The modified swap model showed significantly smaller standard deviations for focused-attention trials as opposed to split-attention trials ($t(19) = 4.411$, $p < 0.001$). These results are outlined in Figure 7b. Focused-attention trials also showed significantly smaller guessing rates than split-attention trials ($t(19) = 4.058$, $p < 0.001$). These results are outlined in Figure 7e.

Figures 7c and 7d show the average B1 and B2 values for each trial type. B1 values correspond to the proportion of trials in which subjects misreported the -90° distractor as the target color, and B2 values correspond to the proportion of trials in which subjects misreported the 30° distractor as the target color. As expected, split-attention trials show a significantly greater proportion of swapping errors than focused-attention trials for both -90° and 30° distractors ($t(19) = -7.05$, $p < 0.001$; $t(19) = -4.719$, $p < 0.001$, respectively). Paired t-tests also showed no significant difference between the proportion of swapping errors for the -90° distractor versus the 30° distractor on split-attention trials ($t(19) = 1.519$, $p = 0.145$). However, focused-attention trials showed significantly more swapping errors for the 30° distractor than for the 90° distractor ($t(19) = -3.174$, $p = 0.0050$).

Figure 8 shows the results of the AIC model comparison for each subject in focused-attention and split-attention trials. Positive values indicate a preference for the standard swap model, and negative values indicate a preference for the modified swap model. Of the 20 total subjects, 14 showed a preference for the standard swap model on focused-attention trials, and 13 showed a preference for the standard swap model on split-attention trials.

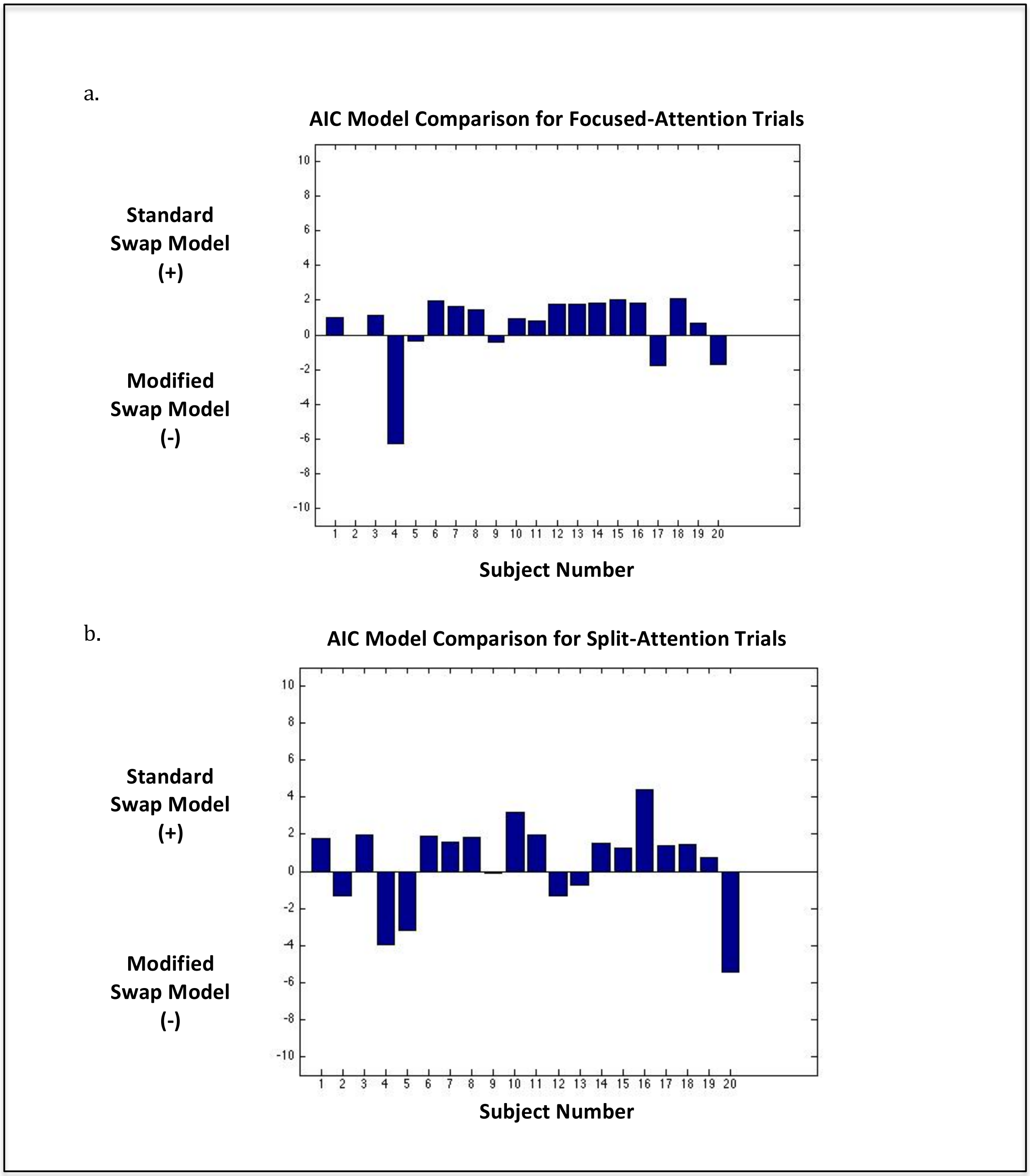


Figure 8: AIC model comparison results across subjects

In addition to the AIC model comparison, we also analyzed what the group results for the mean distribution shift look like when allowing for each subject to be analyzed using their individual model of best fit (standard swap or modified swap). Figure 9 shows the mean distribution shift when selecting for the best model for each individual subject.

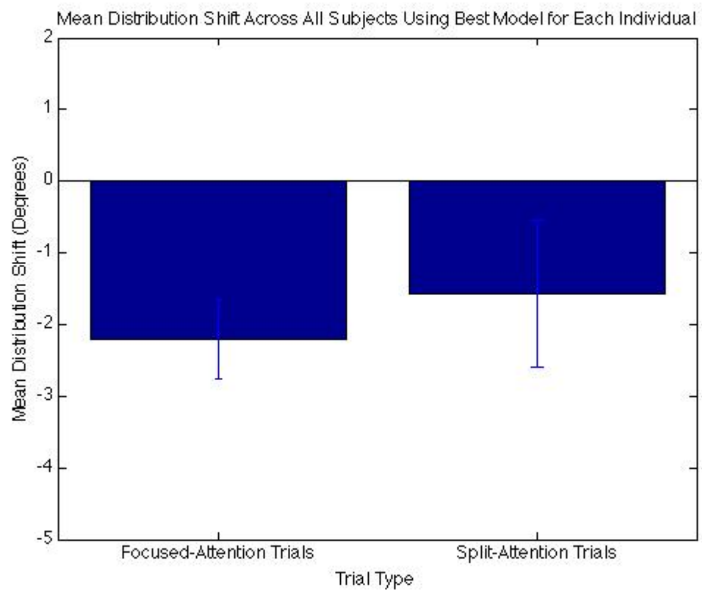


Figure 9: Mean distribution shift across all subjects when selecting for individual models of best fit

When selecting for the model of best fit for each individual subject, statistical results match those of the modified swap model, with focused-attention trials showing a significant shift toward the ensemble mean when compared to chance ($t(19) = -4.063$, $p < 0.001$), but not split attention trials ($t(19) = -1.571$ $p = 0.133$). And when comparing focused-attention trials to split-attention trials, these results are not significantly different ($t(19)= 0.685$, $p = 0.502$).

**Discussion.** Most subjects showed a preference for the standard swap model based on the AIC comparison. Using this model, subjects showed a significant tendency to report a color shifted toward the average across all colors in both focused-attention and split-attention trials, and these conditions were not significantly different. Focused-attention trials, however, showed a decrease in overall noise, with lower standard deviations, guessing rates, and swapping errors. These results suggest that regardless of whether spatial attention was split across objects or focused on a single object, subjects perceived the target color to be drawn toward the ensemble mean. However, the decreases in standard deviation and noise suggest that focused attention made subjects more precise in their responses, as they were only tasked with remembering a single object's feature as opposed to three.

One concern we had when analyzing the data for Experiment 2 was that the distractor set at 30° may be positioned too closely to the target color. This could potentially cause swapping errors centered on the 30° distractor to be counted as a shift in the overall distribution. The fitting showed no clear "bump" in responses around the 30° distractor, and instead fit these responses as part of the overall distribution shift. Since data was aligned so that the 30° distractor lied in the positive direction from the target color, this could cause the average distribution shifts for each trial type to appear at a more positive value than it should.

## Future Directions

Experiments 1 and 2 explored how spatial attention influenced the perception of a target color within an ensemble of three objects. However, an ensemble of three objects is small, and could allow subjects to simply retain each individual color in their memory rather than obtain an ensemble representation of the scene. Previous research has shown that subjects are only able to remember the colors of up to four squares when presented for 100 ms (Luck & Vogel, 1997), but they can perceive the average of a certain feature of more than four objects when presented for just 50 ms (Ariely, 2001; Chong & Treisman, 2003).

We next plan to investigate the effects of spatial attention when subjects are exposed to a larger ensemble in which the ensemble mean lays a value farther away from the target color. To do this, we plan to modify the original experimental design so that subjects are shown an ensemble containing five objects, asked to attend to either all five of the objects or just one, and then told to report their responses by clicking along a color wheel. We will also add one block in which subjects will be asked to report the average across all five colors. This block will allow us to see if subjects can actually perceive and remember the average across all colors. If subjects can perceive and remember the average, this may affect their response strategies.

**Methods.** This proposed experiment will consist of 8 initial blocks containing 96 trials per block as well as a ninth block containing 32 trials in which subjects will be asked to report the average across all colors.

One concern with the setup of this experiment is that displaying five colors may make the task seem too difficult to subjects, since subjects expressed in the previous experiments that it was difficult to see three colors when displayed for such a short period of time. To address this, we will remove the mask that appeared after the color display in previous experiments. This should make it slightly easier for subjects to retain the colors in their memory, but not so easy that it should defeat the purpose of the task.

Similarly to the previous two experiments, the first 8 blocks of this experiment will require subjects to either attend to all five squares (a split-attention trial) or just one square (a focused-attention trial) and then report the color in the post-cued square by clicking along a color wheel. Figure 10 below shows a sample trial for the first 8 blocks of Experiment 3.

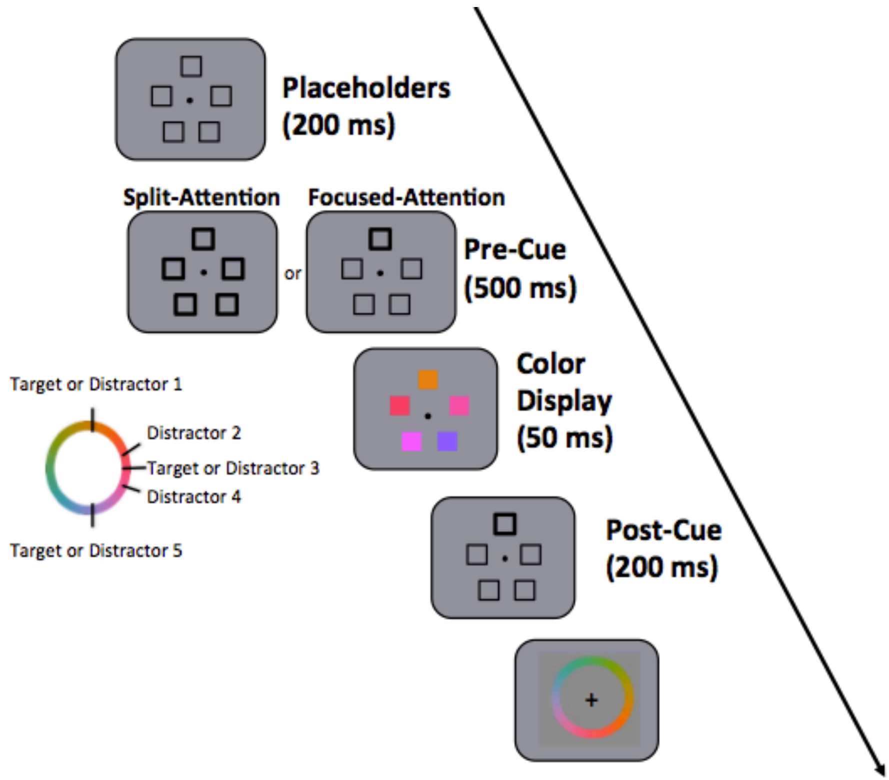


Figure 10: Sample trial for first 8 blocks of Experiment 3

In the ninth block, all squares will become bold in each pre-cue, and instead of making a single square bold in the post-cue like in the first 8 blocks, all squares will become bold again (to keep the pace of each trial the same as in previous blocks). Following the disappearance of the post-cue, subjects will be asked to report the average across the five colors by clicking along a color wheel.

For each block, colors will be distributed so that a random color is chosen and designated to serve as the 0° point on the color wheel, and the remaining four colors will all be shifted in the same direction (positive or negative) from the color at 0°. The second color will be shifted ±80° from this color, the third will be shifted ±90°, the fourth will be shifted ±100°, and the fifth will be shifted 180°.

This experiment has four different conditions for the first 8 blocks. The first two conditions are as follows:

1. Focused-attention trials in which the target color is positioned at 0° or 180°.
2. Split-attention trials in which the target color is positioned at 0° or 180°.

Since distractor colors are positioned at the same relative distances from the target color in trials where the target is at 0° as in trials where the target is at 180°, these two trial types will be counted as the same condition. Colors will be aligned in data analysis so that when the target color is at 0°, the negative value of the distractor colors and responses will be taken. By doing this, all trials where the target is at 0° or 180° will be aligned to have distractor colors set at -80°, -90°, -100°, and 180° from the target color. In these trials, the average across all colors will lay at -90°.

The third and fourth conditions will be as follows:

3. Focused-attention trials in which the target color is positioned at ±90°.
4. Split-attention trials in which the target color is positioned at ±90°.

In trials for conditions 3 and 4, distractor colors will be set at -90°, -10°, +10°, and +90° from the target color, and the average across all colors will lie at the same value as the target color.

The experiment will be balanced so that each of these four conditions receives the same number of trials per block.

**Expected Results.** For each trial, the average across all colors will always lay at ±90° from the color designated at 0°. In focused-attention trials where the target color is positioned at 0° or 180°, we would expect focused attention to mute the effects of the ensemble and allow subjects to report the target color without showing a significant shift toward the ensemble mean.

In focused-attention trials where the target is at ±90°, however, the target is already centered on the average across all colors, and focused-attention should mute any effects of distractor colors. In these trials, we would expect subjects to accurately report the target color without showing any significant shifts in either direction.

In split-attention trials where the target is at 0° or 180°, we would expect to see subjects consistently report a color shifted toward the ensemble mean, since they should be sensitive to the ensemble information when attending to all five colors in the color display.

In split-attention trials where the target is at ±90°, however, subjects should not show a significant shift in their responses since the average across all colors is already centered at the target color. If consistent with the results of Experiment 2, however, these trials may show more noise than focused-attention trials where the target is at ±90°.

## General Discussion

Ensemble statistics are an effective means by which we can retain the gist of a scene without relying on processing of every feature of every object. Ensemble statistics can also bias our perception so that a single feature of a single object is influenced by the features of the objects around it, often drawing the perception of this single feature toward the ensemble mean (Alvarez & Oliva, 2009). Spatial attention is another important factor for feature perception, as it can determine how accurately we bind features to objects (Treisman, 1998).

We examined whether focused attention as opposed to split attention would be able to mute the effects of ensemble statistics and allow for more accurate reports of a target color. To test this, we ran two experiments. In the first experiment, subjects were presented with three differently colored objects, told to attend to either all three or just one of the objects, then asked to report the color of a single object by choosing between two response colors (one slightly shifted toward the ensemble mean and one shifted slightly away from it).

We found that subjects were not more likely to choose the color shifted toward the ensemble mean at a more significant rate than chance for either trial type (focused-attention or split-attention), and these conditions were not significantly different. This would suggest that responses were not being pulled toward the ensemble mean and that spatial attention was not influencing a subject's perception of the target color.

However, we identified several limitations to the experimental design that could have prevented this task from acting as an accurate measure of our research question. The primary limitation was the format of the response cue. Many subjects felt as though they could not tell the difference between the two response colors, as these colors were just 8° apart in color space. In this case, subjects would be forced to guess at which response color best matched the target color regardless of the trial type. This could account for the roughly 50/50 split in the number of responses that reported a shift toward the ensemble mean versus away from the ensemble mean in each trial type.

In the second experiment, we changed the response format to a continuous report in which subjects reported the target color by clicking along a color wheel. Using this paradigm and a standard swap model for data fitting, we found responses for both focused-attention and split-attention trials to be significantly pulled toward the ensemble mean, and these conditions were not significantly different. When using a modified swap model for data fitting, we found responses for focused-attention trials to be significantly pulled toward the ensemble mean, but not split-attention trials. These two conditions were not significantly different, however, indicating that focused-attention cannot be said to modulate the effect differently than split-attention. Both these results suggest that subjects are sensitive to ensemble information when observing a scene, but attention does not significantly modulate this sensitivity.

The standard swap model was found to be the preferred model for most subjects based on an AIC comparison, but the group data when selecting for the model of best fit for each individual subject based on that AIC comparison more closely resembled the results of the modified swap model. The standard swap model includes less parameters, which potentially protects it from "over-fitting" the data, but also assigns the same weight to each distractor color, which could be problematic if one distractor has a greater influence on the overall data than the other. The modified swap model allows for each distractor color to be assigned a different weight, but also includes more parameters, which could decrease overall accuracy. Since this model fits each distractor separately, it may also over-fit or under-fit the swapping errors centered on the 30° distractor (as we saw in Experiment 2 when it looked like responses around 30° were being counted only as part of the overall distribution shift instead of partially as swaps). This could in turn affect the measure of the overall distribution shift ($\mu$) if these errors are not correctly accounted for.

The result that responses on split-attention trials showed a significant tendency to be drawn toward the ensemble mean is consistent with previous studies showing that ensemble statistics can bias feature perceptions when attention is distributed (Brady & Alvarez, 2009).

The result that focused-attention trials also showed significant response shifts toward the ensemble mean, however, was not what we expected. In a study done by Alvarez & Oliva, it was found that subjects were able to accurately report the center of mass of distractors even when asked to focus their attention on a set of targets and ignore the distractors, as they were not relevant to the primary task (Alvarez & Oliva, 2009). This suggests that subjects are still sensitive to ensemble information in a scene, even when attention is focused on a subset of objects instead of split across all objects in the ensemble. This may explain why subjects showed a significant tendency for reporting a color shifted toward the ensemble mean on focused-attention trials in Experiment 2. Even when asked to direct their attention to a certain object, they may still take into account the features of the ensemble, which could in turn affect their responses. However, the Alvarez & Oliva study also found that subjects were still able to accurately report features of individual target objects. This suggests that unlike our experiment, the ensemble mean was not influencing perception of features within the focus of attention.

While we found no significant effect of attention on whether or not responses were being pulled toward the ensemble mean, we did find a significant effect of attention on standard deviations, frequency of swapping errors, and guessing rates (all of which were lower for focused-attention trials). These results suggest that focused attention is effective for increasing precision and decreasing noise, but is not influencing the effects of the ensemble bias. This is consistent with the idea that focused attention can aid in the perception of features within an ensemble, but perhaps it does so through effects on precision rather than accuracy. Ensemble statistics may also be a more relied-upon mechanism for perception than we expected, as they could not be muted by focused attention.

However, it could also be the case that attention does influence the effects of ensemble statistics, but under different conditions. One concern with the design of Experiments 1 and 2 was that three objects is a small number to be considered an ensemble, and that perhaps this could increase the probability of storing the colors in working memory. Displays of up to four colors can be stored in working memory when displayed for 100 ms (Luck & Vogel, 1997), but it has been shown that people can accurately report the average of a certain feature in an ensemble regardless of how many objects are in that ensemble and with presentation times of 50 ms (Ariely, 2001; Chong & Treisman, 2003). So, we next propose an experiment that would display five colors for 50 ms and employ the same attention conditions as Experiment 2 to see if attention significantly modulates feature perceptions within a larger ensemble.

As we take in the world around us, we must collect and prioritize information about the features of surrounding objects to get a coherent depiction of our environment.

Ensemble statistics can help us retain overall accuracy by averaging these features. Spatial attention can also play a role in how we encode features. This study showed that differences in spatial attention (focused versus split) did not significantly modulate whether reports of the color of a single object within an ensemble were pulled toward the ensemble mean. However, spatial attention was found to significantly reduce standard deviations, swapping errors, and guessing rates, indicating that it made subjects more confident and precise in their responses. Future studies will test whether this trend persists when subjects are exposed to a larger ensemble in which the ensemble mean lays at a value further away from the target color.

**References**

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, AC-19, 716–723.

Alvarez, G. and Oliva, A. (2008) The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19(4), 392-398.*

Alvarez, G. and Oliva, A. (2009) Spatial ensemble statistics: efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, USA, 106*, 7345–7350.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. Psychological Science, 12(2), 157-162.

Brady T. F. Alvarez G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22, 384–392.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

Chong S. C., Treisman A. (2003). Representation of statistical properties. *Vision Research*, 43 (4), 393–404.

De Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences*, 108(32), 13341-13346

Golomb, J.D., L'Heureux, Z.E., and Kanwisher, N. (2014). Feature-binding errors after eye movements and shifts of attention. *Psychological Science.* 25(5), 1067-1078.

Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281 (1997).

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437-442.

Palmer J. (1990). Attentional limits on the perception and memory of visual information. *J Exp Psychol Hum Percept Perform*, 16(2), 332-50.

Simon, J. R.,Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. Ergonomics, 6(1), 99-105.

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *353* (1373), 1295–1306.

Triesman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233– 235.

## Appendix

**Response Bias in Experiment 1.** In the initial version of Experiment 1, the two response colors were positioned at the center of the computer screen. As we ran the experiment, we analyzed the data and noticed a bias in the results. Specifically, when the bottom left square was post-cued, subjects tended to choose the left-hand response color at a much higher rate than the right-hand response color. The same was found for the right-hand response color when the bottom right square was post-cued (See Figure A1 for example data from one subject). This effect is similar to a previously reported bias called the Simon Effect, which states that people often respond more quickly and more accurately if the response is shown in the same relative location as the stimulus (Simon & Wolf, 1963).

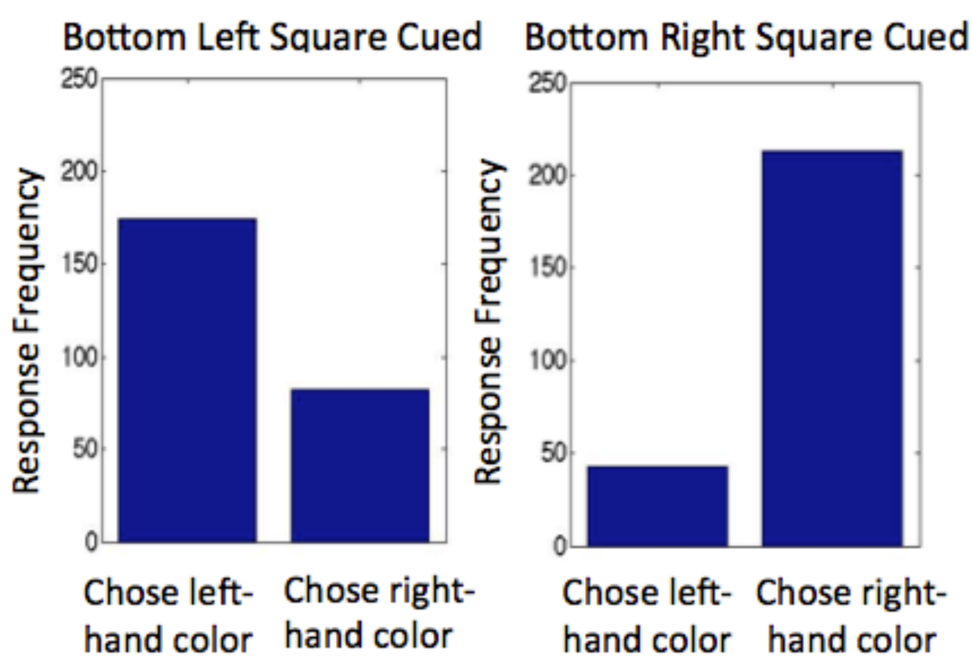


Figure A1: Response frequency for each response color based on post-cued square for one subject

In order to address this bias, the experiment was modified so that the response cue placement was centered on the location of the post-cued square. This reduced the bias, and 8 additional subjects were run with this new version of the experiment. Data from these 8 subjects are presented in the paper.