**Supplemental Material**

**Analysis including all participants above chance**

As explained in the methods section of each experiment, our preregistered analysis plans called for the exclusion of all participants with d’<0.5, which resulted in a large number of exclusions in some cases. Below, we report the results of a post-hoc reanalysis including all participants above chance for each experiment.

*Experiment 1:*

Data from three participants were excluded due to their overall d’ value across all conditions being lower than 0. Twelve participants were excluded due to more than 10% of the trials not meeting the response time exclusion criterion. Using the remaining 71 participants’ data, we replicated the main effect of higher memory performance for the intact object condition (mean d’=0.55) relative to the scrambled object condition (mean d’=0.39; *t*(70)=6.14, *p*<0.001; Cohen’s dz=0.60).

*Experiment 2A:*

Data from one participant was excluded due to their overall d’ value across all conditions being lower than 0. Six participants were excluded due to more than 10% of the trials not meeting the response time exclusion criterion. Using the remaining 71 participants’ data, we replicated the effect of higher memory performance for the intact object condition (mean d’=0.54) relative to the scrambled object condition (mean d’=0.35; *t*(70)=6.42, *p*<0.001; Cohen’s dz=0.68).

*Experiment 2B:*

Data from one participant was excluded due to their overall d’ value across all conditions being lower than 0. Three participants were excluded due to more than 10% of the trials not meeting the response time exclusion criterion. Using the remaining 43 participants’ data, we replicated the effect of higher memory performance for the intact object condition (mean d’=0.92) relative to the scrambled object condition (mean d’=0.82; *t*(42)=2.06, *p*=0.045; Cohen’s dz=0.18).

*Experiment 4:*

No participants were excluded due to their overall d’ value across all conditions being lower than 0. Two participants were excluded due to more than 10% of the trials not meeting the response time exclusion criterion. Using the remaining 46 participants’ data, we replicated the effect of higher memory performance for the intact object condition (mean d’=0.66) relative to the scrambled object condition (mean d’=0.47; *t*(45)=4.64, *p*<0.001; Cohen’s dz=0.58).

*Experiment 5:*

No participants were excluded due to their overall d’ value across all conditions being lower than 0. Five participants were excluded due to more than 10% of the trials not meeting the response time exclusion criterion. Using the remaining 37 participants’ data, we did not find a significant effect but a trend of higher memory performance for the intact object condition (mean d’=1.14) relative to the scrambled object condition (mean d’=1.06; *t*(36)=1.98, *p*=0.055; Cohen’s dz=0.14).

**Binding failures**

As explained in the General Discussion, some previous work has found that making objects less distinct tends to result primarily in binding failures, sometimes termed swap errors (e.g., Oberauer & Lin, 2017; Brown et al., 2021). Thus, we re-analyzed the results from the key experiments looking at trials where the foil color was similar to, vs. dissimilar to, other non-target items on the display to look at whether such binding failures were the main cause of reduced performance for scrambled or inverted stimuli.

In particular, because all four stimulus colors during display were randomly picked with the only constraint of being at least 30 degrees apart, it is possible that on some trials the foil color during the 2-AFC closely matched one of the (untested) colors of the objects in the initial memory display, potentially causing ‘swap’ errors. This could especially be true if there were binding errors between object identities and colors among the memoranda, which potentially could be more likely for scrambled objects relative to the intact object condition.

*Experiment 1 Potential Swap Error Trials Analysis:*

To test whether the d’ benefit for intact objects was driven by a higher swap rate in the scrambled object condition relative to the intact object condition, we compared performance for trials without potential swaps (foil colors distinct from non-targets) to performance for trials with potential swaps (foil colors similar to one of the non-targets). To define these potential swap trials, we computed the difference between the foil color and three non-target colors in each trial. We then identified potential swap trials in three different similarity ranges (differences between foil color and non-target colors are within 5, 10, and 15 degrees on the color wheel) and conducted statistical analyses comparing performance in the intact vs. scrambled condition while including potential swap trial as a factor (potential swap vs. no-swap trial). In Experiment 1, 8.5% of the trials were flagged as potential swap trials within the 5 degrees range, 20% for 10 degrees, and 30% for 15 degrees. For the 5 degrees range, a 2x2 ANOVA on d’ with stimulus type (intact vs. scrambled) and potential swap trial (potential swap vs. no-swap) as factors yielded a main effect of stimulus type (F(1,29)=6.28, *p*=0.02) but no significant main effect of swap trial (F(1,29)=0.12, *p*=0.73) and no interaction (F(1,29)=0.004, *p*=0.95). The same exact pattern was found for 10 degrees range (stimulus type main effect: F(1,29)=6.77, *p*=0.01; swap trial main effect: F(1,29)=0.19, *p*=0.66; interaction: F(1,29)=0.23, *p*=0.63), and 15 degrees range (stimulus type main effect: F(1,29)=16.58, *p*<0.001; swap trial main effect: F(1,29)=0.38, *p*=0.55; interaction: F(1,29)=0.37, *p*=0.55). Thus, binding errors (differences in swap errors across the intact and scrambled conditions) do not explain the performance difference between the conditions.

*Experiment 2A Potential Swap Error Trials Analysis:*

The same analyses were performed for Experiment 2A. 8.1% of the trials were flagged as potential swap trials within the 5 degrees range, 19.5% for 10 degrees, and 29.3% for 15 degrees. For the 5 degrees range, 2x2 ANOVA on d’ with stimuli type and potential swap trial as factors yielded no significant main effects or interaction (stimulus type main effect: F(1,29)=3.00, *p*=0.09; swap trial main effect: F(1,29)=0.10, *p*=0.76; interaction: F(1,29)=2.91, *p*=0.10). For the 10 degrees range, a 2x2 ANOVA yielded a main effect of stimulus type: F(1,29)=19.68, *p*<0.001, but no significant main effect of swap trial: F(1,29)=0.10, *p*=0.75, or interaction: F(1,29)=0.15, *p*=0.71. The same pattern was found for 15 degree range (stimulus type main effect: F(1,29)=24.87, *p*<0.001; swap trial main effect: F(1,29)=0.94, *p*=0.34; interaction: F(1,29)=0.19, *p*=0.67).