## Supplemental Material

**Figure S1. Mixture model parameters for free response order experiments (1a and 1b).** Mixture model parameter estimates are shown for all responses and set-sizes in Experiments 1a and 1b. Note, error bars represent standard error of the mean across subjects and do not accurately reflect goodness of fit for individual subjects' fits. Fits for responses 5 and 6 were particularly unreliable, and a uniform distribution provided a better fit for these responses.

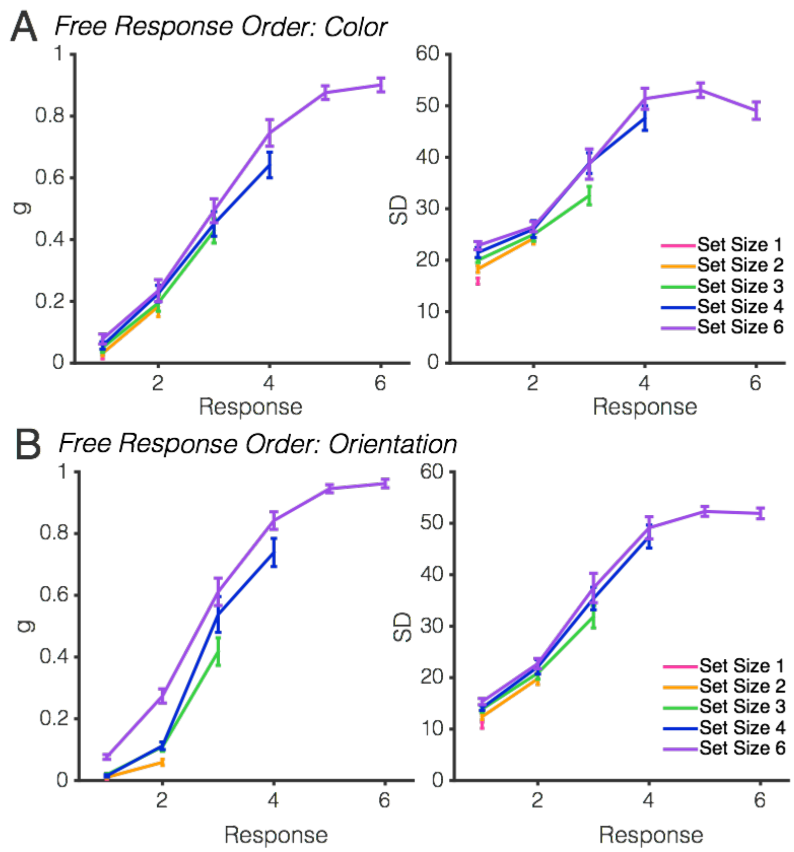

**Figure S2. Mixture model parameters for random response order experiments (2a and 2b).** Mixture model parameter estimates are shown for all responses and set-sizes in Experiments 2a and 2b. Error bars represent standard error of the mean across subjects.

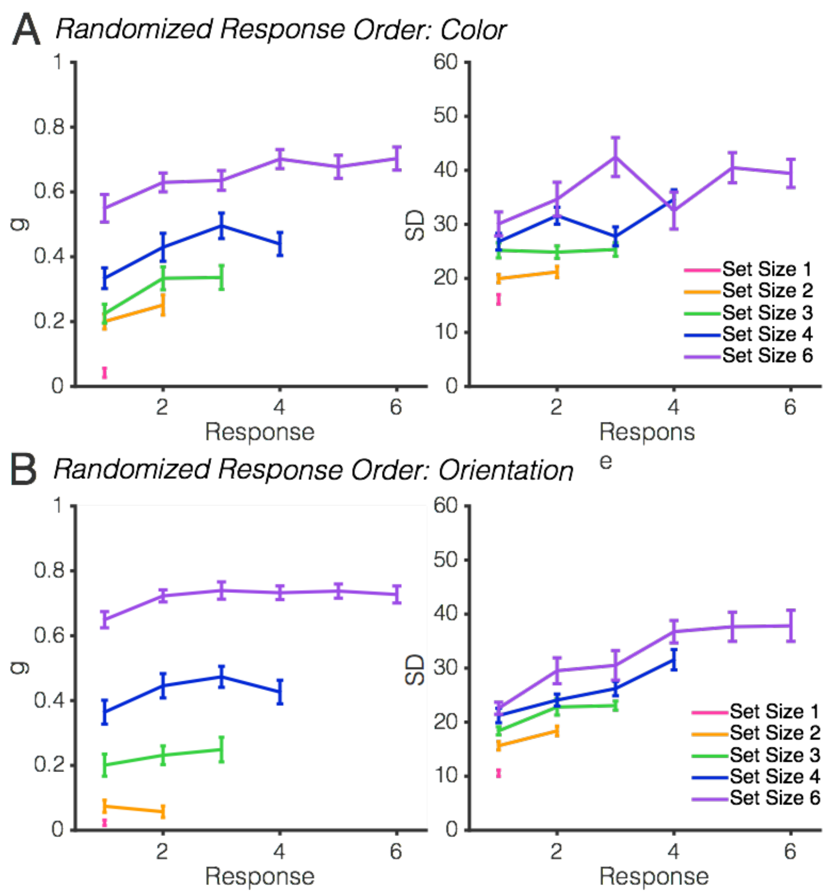

**Table S1. BIC values for within-subject model competition in Experiment 1a (Fig. 5b).** Values are a subtraction between the BIC values for the uniform model and the standard mixture model. Positive values represent evidence for a uniform model, negative values represent evidence for the mixture model.

| *Subject* | *Response 1* | *Response 2* | *Response 3* | *Response 4* | *Response 5* | *Response 6* |
|---|---|---|---|---|---|---|
| *1* | -192.12 | -58.86 | **0.27** | **3.23** | **7.59** | **7.58** |
| *2* | -286.34 | -128.25 | -26.82 | **5.07** | **7.57** | **6.41** |
| *3* | -223.05 | -161.57 | -33.57 | -8.12 | **1.83** | **9.19** |
| *4* | -248.51 | -193.43 | -63.92 | **9.00** | **8.95** | **8.68** |
| *5* | -187.50 | -37.51 | **8.90** | **8.98** | **8.74** | **9.19** |
| *6* | -107.42 | -39.33 | -7.44 | **9.13** | **3.38** | **8.36** |
| *7* | -189.06 | -56.93 | -24.90 | **2.98** | **8.93** | **8.94** |
| *8* | -165.30 | -105.96 | -16.74 | **7.24** | **3.59** | **7.77** |
| *9* | -252.67 | -141.33 | -49.91 | -6.40 | **4.42** | **1.90** |
| *10* | -146.65 | -91.60 | -13.94 | **6.89** | **4.09** | **8.98** |
| *11* | -186.46 | -171.42 | -79.86 | -12.34 | **8.30** | -12.01 |
| *12* | -237.33 | -136.57 | -8.21 | **8.81** | **8.28** | **6.63** |
| *13* | -141.82 | -59.48 | -13.82 | **7.48** | **9.11** | **5.00** |
| *14* | -264.87 | -208.29 | -90.62 | -2.82 | **6.89** | **8.61** |
| *15* | -167.08 | -115.96 | -31.37 | **5.50** | **8.50** | **9.15** |
| *16* | -207.49 | -124.19 | -60.39 | -8.27 | **6.78** | **2.39** |
| *17* | -217.39 | -129.32 | -14.29 | **3.94** | **5.98** | **7.06** |
| *18* | -183.91 | -100.47 | -44.01 | -26.83 | **3.21** | **4.05** |
| *19* | -272.45 | -135.36 | -29.69 | -1.40 | **8.82** | **2.89** |
| *20* | -206.87 | -113.08 | -28.75 | **0.89** | **5.13** | -9.44 |
| *21* | -254.71 | -191.71 | -42.12 | **6.19** | **8.60** | **7.92** |
| *22* | -192.12 | -58.86 | **0.27** | **3.23** | **7.59** | **7.58** |

**Table S2. BIC values for within-subject model competition in Experiment 1b (Fig. 5b).** Values are a subtraction between the BIC values for the uniform model and the standard mixture model. Positive values represent evidence for a uniform model, negative values represent evidence for the mixture model.

| *Subject* | *Response 1* | *Response 2* | *Response 3* | *Response 4* | *Response 5* | *Response 6* |
|---|---|---|---|---|---|---|
| *1* | -447.98 | -155.63 | -61.25 | -12.93 | **8.43** | **9.11** |
| *2* | -710.76 | -250.96 | **1.13** | **10.56** | **10.60** | -13.33 |
| *3* | -788.71 | -277.89 | -3.59 | **6.58** | **6.87** | **10.60** |
| *4* | -593.99 | -253.28 | -45.42 | **2.51** | **10.29** | **10.43** |
| *5* | -672.01 | -409.71 | -105.72 | -1.43 | **9.59** | **7.50** |
| *6* | -508.77 | -284.16 | -24.75 | **10.58** | **10.60** | **9.89** |
| *7* | -555.06 | -69.98 | **7.42** | **8.39** | **9.69** | **9.71** |
| *8* | -516.83 | -309.68 | -115.96 | **3.89** | **9.81** | **5.99** |
| *9* | -556.70 | -385.04 | -86.80 | **2.48** | **8.78** | **10.60** |
| *10* | -587.12 | -321.41 | -80.63 | -0.60 | **10.60** | **10.60** |
| *11* | -585.34 | -266.81 | -29.74 | **4.90** | **10.03** | **10.57** |
| *12* | -412.77 | -158.82 | -44.50 | -9.95 | **8.02** | **10.60** |
| *13* | -604.93 | -276.93 | -79.86 | -0.20 | **6.92** | **10.60** |
| *14* | -556.60 | -304.80 | **10.14** | **7.70** | **8.42** | **3.89** |
| *15* | -543.98 | -189.28 | -62.66 | **4.87** | **10.60** | **10.43** |
| *16* | -666.89 | -353.54 | -194.41 | -24.41 | -94.52 | **10.60** |
| *17* | -417.49 | -179.06 | -28.86 | **8.90** | **5.94** | **5.94** |
| *18* | -615.83 | -425.59 | -125.24 | **3.84** | **5.51** | **10.60** |
| *19* | -497.71 | -293.22 | -17.98 | **9.94** | **9.45** | **9.67** |
| *20* | -597.94 | -255.96 | **5.50** | **3.77** | **10.60** | **9.42** |

**ORIGINAL – CONTAINS ERROR**

~~**Analysis S1. Swap analysis for Experiment 1.**~~

~~We considered what proportion of the uniform distributions observed for the later Set Size 6 responses could be explained by an increase in swap errors. First, we fit data from Experiment 1 using a Mixture Model with Swapping in Memtoolbox (Suchow, Brady, Fougnie, & Alvarez, 2013). Swapping represented a small proportion of Set Size 6 responses (Table S3). On average, swapping occurred for 4% of all Set Size 6 responses in Experiment 1a and 5% of all Set Size 6 responses in Experiment 1b.~~

~~Next, we ran a repeated-measures ANOVA on the Set Size 6 swap rates with Response Number as a within-subjects factor. Planned contrasts compared each earlier response (1-5) to the final response (6). In Experiment 1a, there was no main effect of response order on swaps, $F(3.5,73.4) = 1.2$, $p = .327$. That is, swaps were no more likely to occur for the later, uniform responses than they were to occur for early responses. In Experiment 1b, there was a significant main effect of response order on swaps, $F(3.2,60.3) = 4.41$, $p = .006$. However, planned contrasts revealed that only responses 1 and 2 were different from response 6 ($p < .02$), whereas responses 3 through 5 were indistinguishable from the final response ($p > .10$).~~

~~In sum, the uniformity observed for responses 5 and 6 cannot be explained by a sudden increase in swap errors. We saw an increase in swap errors across responses in only one of the two experiments. When a swap error increase was observed (Exp. 1b), the increase was relatively modest (swap rate increased by on average 1% per response).~~

~~**Table S3.** *Average swap rate for Set Size 6 responses in Experiments 1a and 1b.* Numbers in parentheses represent one standard deviation.~~

| | ~~*Response 1*~~ | ~~*Response 2*~~ | ~~*Response 3*~~ | ~~*Response 4*~~ | ~~*Response 5*~~ | ~~*Response 6*~~ | ~~***Average***~~ |
|---|---|---|---|---|---|---|---|
| ~~***Exp. 1a***~~ | ~~.02 (.03)~~ | ~~.05 (.10)~~ | ~~.03 (.05)~~ | ~~.07 (.12)~~ | ~~.03 (.06)~~ | ~~.06 (.12)~~ | ~~**.04 (.04)**~~ |
| ~~***Exp. 1b***~~ | ~~.01 (.02)~~ | ~~.02 (.03)~~ | ~~.05 (.06)~~ | ~~.08 (.10)~~ | ~~.06 (.07)~~ | ~~.09 (.11)~~ | ~~**.05 (.04)**~~ |

**CORRECTED ANALYSIS**

## Analysis S1. Swap analysis for Experiment 1.

We considered what proportion of the uniform distributions observed for the later Set Size 6 responses could be explained by an increase in swap errors. First, we fit data from Experiment 1 using a Mixture Model with Swapping in Memtoolbox (Suchow, Brady, Fougnie, & Alvarez, 2013). Swapping represented a small proportion of Set Size 6 responses (Table S3). On average, swapping occurred for 10% of all Set Size 6 responses in Experiment 1a and 5% of all Set Size 6 responses in Experiment 1b.

Next, we ran a repeated-measures ANOVA on the Set Size 6 swap rates with Response Number as a within-subjects factor. In Experiment 1a, there was a main effect of response order on swap rate, $F(5,105) = 6.40$, $p < .001$. Critically, however, this significant effect was not due to a monotonic increase in swap rate as a function of response number. Instead, the effect was non-monotonic, peaking at the third response. In fact, the swap rate for responses 5 and 6 was no different than for response 1, $p > .45$. In Experiment 1b, there was no significant main effect of response order on swaps, $F(5,95) = 1.50$, $p = .197$. That is, swaps were no more likely to occur for the later, uniform responses than they were to occur for early responses. In sum, the uniformity observed for responses 5 and 6 cannot be explained by a sudden increase in swap errors.

**Table S3.** *Average swap rate for Set Size 6 responses in Experiments 1a and 1b.* Numbers in parentheses represent one standard deviation.

| | *Response 1* | *Response 2* | *Response 3* | *Response 4* | *Response 5* | *Response 6* | ***Average*** |
|---|---|---|---|---|---|---|---|
| ***Exp. 1a*** | .03(.03) | .13(.11) | .21(.19) | .13(.18) | .05(.12) | .06(.11) | **.10(.08)** |
| ***Exp. 1b*** | .03(.03) | .08(.05) | .07(.09) | .04(.06) | .06(.09) | .04(.07) | **.05(.03)** |

**Analysis S2. Circular statistics approach to testing for uniformity in Experiments 1 & 3.**

**Experiment 1.** As an alternative to the BIC model comparison between uniform and mixture model distributions, we used a modified Rayleigh Test for uniformity ('circ_vtest.m'; Zar (2010) pp. 626). The modified Rayleigh test makes a priori hypotheses that the data is uni-modal and has a specific mean direction (in this case, a response error value of 0 degrees). We ran a modified Rayleigh test on each individual's six response distributions for set size 6, corrected for six within-subject multiple comparisons. This revealed an average of 2.36 uniform distributions per subject in Experiment 1a (SD = .79, ranging from1 to 4) and 2.85 uniform distributions per subject in Experiment 1b (SD = .59, ranging from 2 to 4). With a set size of 6 items, the prevalence of uniform distributions was aligned with past estimates of putative item limits in the literature (i.e., subjects had non-zero information about 2-5 items). A large majority of subjects (82% in Experiment 1a, 90% in Experiment 1b) showed evidence of two or three uniform items, in line with a capacity of 3-4 items.

**Experiment 3.** The modified Rayleigh's test for uniformity estimated that participants had a mean of 2.4 uniform distributions (SD = .97, range from 1-4).

**Table S4. BIC values for model competition in Experiment 2a (Figure 9a).** Values are a subtraction between the BIC values for the uniform model and the standard mixture model. Positive values represent evidence for a uniform model, negative values represent evidence for the mixture model.

| | *Response 1* | *Response 2* | *Response 3* | *Response 4* | *Response 5* | *Response 6* |
|---|---|---|---|---|---|---|
| *First 3 Guesses* | -4.2 | **3.3** | **3.1** | -50.1 | -19.4 | -27.9 |
| *Last 3 Guesses* | -171.0 | -69.7 | -45.7 | **5.8** | **5.0** | **5.6** |

**Table S5. BIC values for model competition in Experiment 2b (Figure 9b).** Values are a subtraction between the BIC values for the uniform model and the standard mixture model. Positive values represent evidence for a uniform model, negative values represent evidence for the mixture model.

| | *Response 1* | *Response 2* | *Response 3* | *Response 4* | *Response 5* | *Response 6* |
|---|---|---|---|---|---|---|
| *First 3 Guesses* | **5.3** | **11.7** | **11.2** | -63.2 | -60.3 | -73.2 |
| *Last 3 Guesses* | -181.7 | -103.2 | -101.5 | **3.0** | **11.7** | **9.2** |