Supplementary Accuracy Results for Experiments 1-3

Experiment 1: In Experiment 1a (75% valid cue), a two-way repeated-measures ANOVA with object size (large, small) and cue validity (valid, invalid) as within-subjects factors and accuracy as a dependent measure revealed a main effect of cue validity ($F(1, 19) = 7.865$, $p < .05$, $\eta_p^2 = .293$), such that attention was allocated less efficiently in invalid locations (94.8%) compared to valid locations (96.2%). No effect of inferred object size or interaction between inferred size and cue validity was observed ($F < 1$). In Experiment 1b (50% valid cue), a two-way repeated-measures ANOVA with object size (large, small) and cue validity (valid, invalid) as within-subjects factors and accuracy as a dependent measure revealed main effect of cue validity ($F(1, 19) = 5.435$, $p < .05$, $\eta_p^2 = .222$), such that attention was allocated less efficiently in invalid compared to valid locations. No main effect of size or significant interaction was observed. These findings are consistent with our response time results, and no speed-accuracy tradeoff was present.

A between-experiment three-way ANOVA with validity type as a between-subject variable and accuracy as a dependent measure revealed a main effect of validity ($F(1, 38) = 13.14$, $p < .001$, $\eta_p^2 = .257$), such that, across both experiments, attention was less efficiently allocated in invalid locations. No other main effects or interactions were significant.

Experiment 2: In Experiment 2a, in which retinotopic scrambling was employed, a two-way repeated-measures ANOVA for accuracy revealed a marginal effect of validity ($F(1, 19) = 4.261$, $p < .06$, $\eta_p^2 = .183$), with higher accuracies in valid locations (95.9%) than in invalid locations (95.0%). No other effects were significant. In Experiment 2b, which employed the full scrambling technique, a main effect of validity ($F(1, 19) = 12.140$, $p < .01$, $\eta_p^2 = .390$) was observed, with higher accuracies in valid locations (96.9%) than in invalid locations (96.0%). No main effect of size or interaction was observed. These findings are consistent with our response time results, and no speed-accuracy tradeoff was observed.

Experiment 3: For Experiment 3, a three-way repeated-measures ANOVA for accuracy was conducted with object size, cue validity, and target spatial frequency (low, high) as within-subject factors. ANOVA revealed marginal effects of cue validity ($F(1, 19) = 3.890$, $p < .07$, $\eta_p^2 = .170$), with higher accuracies in valid (96.7%) than invalid locations (96.1%), and target spatial frequency ($F(1, 19) = 3.203$, $p < .09$, $\eta_p^2 = .144$), with higher accuracies in low (96.6%) compared to high spatial frequency (96.1%) targets. A three-way interaction between object size, cue validity, and spatial frequency was marginally significant ($F(1, 19) = 4.269$, $p < .06$, $\eta_p^2 = .183$), suggesting that target spatial frequency (low or high) modulates the degree to which attentional shifts are constrained by object size, as shown with response times. These findings are consistent with our response time results, and no speed-accuracy tradeoff was observed.

Supplementary Orientation Analyses

Experiment 3: A two-way interaction between size and orientation was observed ($F(1, 19) = 21.37$, $p < .001$, $\eta_p^2 = .529$), with larger size effect in the vertical orientation than the horizontal orientation (vertical orientation, $\Delta = 28.5$ ms; horizontal orientation, $\Delta = 4.1$ ms). An additional two-way interaction between target frequency and object orientation was significant ($F(1, 19) = 8.02$, $p < .05$, $\eta_p^2 = .297$), with greater orientation effect at higher target frequency. A three-way interaction between size, validity, and orientation was also observed ($F(1, 19) = 11.7$, $p < .05$, $\eta_p^2 = .381$), showing that attentional shifts are more strongly modulated by object size in the vertical condition (large objects, $\Delta = 66.5$ ms; small objects, $\Delta = 46.9$ ms) than the horizontal condition (large objects, $\Delta = 34.2$ ms; small objects, $\Delta = 58.3$ ms).

Additional orientation analysis was conducted in order to address a possible confound that some of the objects in the 'large' object group have very different canonical orientations than the rest of the objects (basketball hoop, bulldozer, car, sofa, and tub).To address the possible confound we re-analyzed the data with problematic large objects removed. Upon removal of these objects, main effects of inferred object size ($F(1, 19) = 14.61$, $p < .001$, $\eta_p^2 = .435$), cue validity ($F(1, 19) = 24.46$, $p < .001$ $\eta_p^2 = .573$), and target frequency ($F(1, 19) = 51.75$, $p < .001$, $\eta_p^2 = .731$) were significant, but the main effect of object orientation ($F(1, 19) = 2.62$, $p = .122$) was not. Notably, the two-way interaction between size and orientation was still significant ($F(1, 19) = 10.24$, $p < .05$, $\eta_p^2 = .350$), suggesting that when the remaining objects were in canonical orientation (vertical), the size effect was larger than when they were not. This effect lends further evidence to size-based attentional scaling, as inferred object size had a greater effect on attention when objects were in the preferred orientation and high-level object information was more readily parsed. Additionally, the target frequency by orientation analysis was significant ($F(1, 19) = 4.50$, $p < .05$, $\eta_p^2 = .191$), suggesting the effect of orientation was stronger when attentional demand was higher. Finally, the three-way interactions between size, validity, and frequency ($F(1, 19) = 11.79$, $p < .05$, $\eta_p^2 = .383$) and size, validity, and orientation ($F(1, 19) = 11.6$, $p < .05$, $\eta_p^2 = .379$) were also significant, but the four-way interaction between size, validity, frequency, and orientation was not ($F(1, 19) = .20$, $p = .657$). Together, these effects suggest that object size affects attentional shifting when attentional demand is high, and when objects are in their preferred orientation.

Clutter Control:

Experiment 1a. Following removal of the 5 most cluttered large objects and 5 least cluttered small objects, main effects of validity ($F(1, 19) = 19.37$, $p < .001$, $\eta_p^2 = .505$) and object orientation ($F(1, 19) = 89.00$, $p < .001$, $\eta_p^2 = .824$) were significant. No interactions were significant, suggesting that object orientation did not modulate the size effect.

Experiment 1b. Following removal of the 5 most cluttered large objects and 5 least cluttered small objects, main effects of validity ($F(1, 19) = 6.20$, $p < .05$, $\eta_p^2 = .246$) and object orientation ($F(1, 19) = 48.93$, $p < .001$, $\eta_p^2 = .720$) were significant. Also, a two-way interaction between validity and object orientation ($F(1, 19) = 14.10$, $p < .05$, $\eta_p^2 = .426$) was significant, suggesting that the effect of cue validity was larger in horizontal objects. However, the size effect remained uninfluenced by object orientation ($F(1, 19) = 0.36$, $p = .559$, $\eta_p^2 = .018$).

Experiment 3. Following removal of the 5 most cluttered large objects and 5 least cluttered small objects, the main effect of orientation ($F(1, 19) = 3.60$, $p = .073$, $\eta_p^2 = .159$) and three way interaction between size, validity, and object orientation ($F(1, 19) = 3.45$, $p = .079$, $\eta_p^2 = .154$) were both marginally significant. The two-way interaction between size and orientation was significant ($F(1, 19) = 12.02$, $p < .05$, $\eta_p^2 = .388$), again suggesting the size effect was stronger in the vertical orientation. An additional two-way interaction between target frequency and object orientation was significant ($F(1, 19) = 4.43$, $p < .05$, $\eta_p^2 = .189$), suggesting that orientation effect was larger with increased attentional demand.

Supplementary Experiments

Experiment S1: In the three experiments reported in the main article, inferred object size has been demonstrated to constrain attentional selection. Namely, canonically smaller objects elicit more efficient attentional deployment. This effect is robust and is not a product of low-level differences between differently-sized objects. In Experiment 1, while inferred size influences the focus of attention in Experiment 1, it did not modulate shifts of attention within objects of different sizes. These results suggest that the effect of size influences focusing of attention, which may occur independent of spatial cueing affecting, which is a fast and automatic process that affects attentional shifts. In Supplemental Experiments S1a and S1b, the time-course of spatial cueing was manipulated to test whether size-based attentional modulation lasts beyond the time-course of the cueing effect. To address the time-course of the effect of inferred object size on attention, cue validity was varied in Experiments S1a and S1b (75% valid cue, Experiment S1a; 50% valid cue, Experiment S1b) and importantly, the SOA was raised to 500 ms, as this is beyond the time-course of the spatial cueing effect. Two-way repeated-measures ANOVA for RT and accuracy were conducted for both experiments. For RTs, in Experiment S1a ($F(1, 15) = 12.72$, $p < .01$, $\eta_p^2 = .459$), and Experiment S1b ($F(1, 16) = 30.75$, $p < .001$, $\eta_p^2 = .658$), size remained influential to attentional allocation, though validity effects were no longer observed (Figs. S1a, b). While the lack of spatial validity effects was expected in these experiments, in accordance with the time-course of spatial attention (Drummond & Shomstein, 2013), the effect of size persisted. These results indicate that the influence of object size is long-lasting and robust, in that it has been replicated across numerous independent manipulations. These robust replications are particularly notable in that they demonstrate strong support for the original hypothesis of size-based attentional scaling. For both Experiments S1a and S1b, two-way repeated-measures ANOVA was also conducted for accuracy results. No effects were found in accuracy (all *F*s < 1).

Experiment S2: In Supplemental Experiment S2, the robustness of the effect of real-world object size was tested using a target detection task, rather than a target identification task. This detection task was conducted to further elucidate the role of object size in attention. As the detection task is less attentionally demanding than identification task, the contribution of object size should be reduced, if inferred object size constrains attentional allocation. In this experiment, participants were tasked with detecting a T or L target in either the cue-valid or cue-invalid location on an object, with no distractor. Cue validity was set to 75% and SOA was 250 ms, as in Experiment 1a. In this experiment, targets appeared on 80% of trials, while no target was present on 20%. Under these conditions, and in participants exceeding 75% accuracy on catch trials, no main effects of size ($F(1,17) = .151$, $p = .703$) or validity ($F(1, 17) = 1.220$, $p = .285$) were observed (Fig. S2a).

However, given that this task was exceedingly easy and did not successfully elicit any attentional effects, Supplemental Experiment 2b was conducted, in which the same design was used. However, all possible target locations were masked following a 50 ms target display and a 50 ms ISI. Under these conditions, and in participants exceeding 75% accuracy on catch trials, main effects of size ($F(1, 21) = 7.461$, $p < .05$, $\eta_p^2 = .262$) and validity ($F(1,21) = 9.498$, $p < .01$, $\eta_p^2 = .311$) were again observed (Fig. S2b), demonstrating that an effect of size only arises when attention is necessary for the task, and when objects successfully contribute to attentional deployment. For both Experiments S2a and S2b, two-way repeated-measures ANOVA was also conducted for accuracy results. No significant effects were observed in accuracy.

Experiment S3: Critically, we suggest here that inferred real-world object size modulates attention to and within objects. However, it remains possible that potential mid- and high-level differences between large and small objects, including the degree of object curvature and object familiarity, could serve as additional explanations for the attentional cost for large objects. For these reasons, it is crucial to equate the objects across these measures. Previous work has suggested that small objects may be curvier than large objects, as they tend have a more ergonomic structure (Long, Konkle, Cohen, & Alvarez, 2016). Additionally, it may be possible that participants were more familiar with small objects compared to large objects, thus processing the object more quickly. Further experiments and analyses were conducted to demonstrate the absence of such mid- and high-level level differences among the employed objects.

We first explored the degree to which object line drawings across both large and small object sets varied in boxy/curvy features. Ratings were collected using Amazon Mechanical Turk (N = 20, limited to master workers within the US) for each object using a previously established boxy/curvy scale in which participants rated each object according to the following scale: 1 (very curvy), 2 (somewhat curvy), 3 (equally boxy and curvy), 4 (somewhat boxy), 5(very boxy). Two independent two-tailed t-test analyses were used to assess differences in boxy/curvy feature judgements. First, a subject-centered paired-samples t-test revealed no difference in boxy/curvy ratings across object size groups (large objects, $M = 3.23$; small objects, $M = 3.35$; $t(19) = -1.792$, $p = .089$, $d = -.401$). A second, object-centered independent-samples t-test also revealed no difference in object features between large and small objects (large objects, $M = 3.23$; small objects, $M = 3.35$; $t(28) = -.295$, $p = .770$, $d = -.108$). While these results clearly suggest that the object sets used in this study do not differ in rectilinear/curvilinear features, a replication was conducted using an additional set of 20 master workers on mTurk in order to validate these findings. Again, neither a paired-samples t-test (large objects, *M* = 3.24; small objects, *M* = 3.30; *t*(19) = -1.792, *p* = .487, *d* = -.158) nor an independent-samples t-test (large objects, *M* = 3.24; small objects, *M* = 3.30; *t*(28) = -.145, *p* = .886, *d* = -.053) revealed any difference between large and small object groups. Furthermore, an additional between-subjects repeated-measures ANOVA conducted across all 40 participants revealed no difference in boxy/curvy object features between large and small groups ($F(1,38) = 2.779$, $p = .104$, $\eta_p^2 = .068$). Crucially, no difference was observed across participant groups ($F(1,38) = .309$, $p = .582$, $\eta_p^2 = .008$), suggesting the two participant groups rated the objects consistently. These results indicate that the two object groups do not differ in object curvature, suggesting that curvature does not contribute to attentional scaling within our paradigm.

Within the same survey, we also investigated whether object size groups differed in participant familiarity with the objects. Ratings were collected using the following scale: 1 (not at all familiar), 2 (slightly familiar), 3 (moderately familiar), 4 (very familiar), 5 (extremely familiar). The first survey (N = 20, limited to US master workers), revealed no familiarity difference between object groups across both paired-samples t-test (large objects, *M* = 4.25; small objects, *M* = 4.37; *t*(19) = -1.363, *p* = .189, *d* = -.305) and independent-samples t-test (large objects, *M* = 4.25; small objects, *M* = 4.37; *t*(28) = -.585, *p* = .563, *d* = -.214). The additional replication of this survey (N = 20 US master workers), following removal of two participants with overall ratings greater than two standard deviations below the group average, revealed no difference in the paired-samples t-test for familiarity (large objects, *M* = 4.40; small objects, *M* = 4.53; *t*(17) = -1.472, *p* = .159, *d* = -.347). An additional between-subjects repeated-measures ANOVA conducted across 37 participants (3 eliminated for ratings 2 SD below group mean) revealed only a marginal difference in familiarity between large and small groups ($F(1,35) = 3.964$, $p = .054$, $\eta_p^2 = .102$), and no difference was observed across participant groups ($F(1,35) = .001$, $p = .974$, $\eta_p^2 = .005$), suggesting the two participant groups rated the objects consistently. Taken together, these results suggest that familiarity is not a viable alternative explanation for the observed RT differences between small and large object groups.

In addition to measuring the perceived curviness of each object, we also tested for differences in object-boundary curvature by using a steerable pyramid model (Freeman & Adelson, 1991; Simoncelli & Freeman, 1995). Steerable pyramid is an image decomposition technique in which images are passed through a set of wavelet filters for a range of scales and orientations, and it can be used for texture and orientation analysis (Freeman & Adelson, 1991; Simoncelli & Freeman, 1995). Here, each image is decomposed across a single subband for eight orientations, resulting in eight unique image orientation filters. The resulting overcomplete representation for each image contained both location and frequency information for each orientation in the image. Orientation maps for large and small object exemplars are shown in Fig. S6b. Here, the orientation frequency information was used to assess the overall curvilinearity of each object.

Using the data from the Steerable pyramid image decomposition, orientation histograms were generated for each image, for which the calculated dominant orientations for each pixel were grouped into 10º bins (ranging from -90º – 90º). The standard deviation of each image's angular distribution was calculated, as we hypothesized that objects with curvier features would have a more uniform angular distribution with less variance. According to this hypothesis, if small objects have curvier features, then they will also have less variance in angular distribution than large objects. An object-centered independent-samples t-test revealed no difference in the standard deviation of angular distribution between large and small objects (large objects, *M* = 27878.11; small objects, *M* = 28345.57; *t*(28) = -.9, *p* = .376, *d* = -.329). This result suggests that small objects are not curvier than larger objects, in accordance with the Amazon Mechanical Turk ratings.

The same analysis was also conducted on an independent set of images in order to provide validation for this technique. The images used for validation were exemplars of 'cuby' and 'smoothy' object classes derived from a classic study in perceived shape (see Fig. S6c; Op de Beeck, Torfs, & Wagemans, 2008). Orientation maps for these object classes are shown in Fig. S6d. If standard deviation of angular distribution is sufficient to distinguish boxy from curvy objects, 'smoothy' exemplars should have a significantly lower angular standard deviation than 'cuby' exemplars due to a due to a smoother, more uniform gradient of angles. An object-centered independent-samples t-test conducted on these shape exemplars revealed a significant difference in the standard deviation of angular distribution between 'cuby' and 'smoothy' exemplars ('cuby' objects, *M* = 10298.85; 'smoothy' objects, *M* = 10004.02; *t*(20) = -2.08, *p* = .05, *d* = -.888). These results suggest that our analysis is sensitive to shape feature differences, and validate the finding that small and large object stimuli are similarly curvilinear.

Supplementary Experiment Methods

Participants. For Experiment S1a, 20 participants were recruited. Three participants were excluded for failing to meet the 90% accuracy criteria. For Experiment S1b, 23 participants were recruited. Seven participants were excluded for failing to meet the 90% accuracy criteria. For Experiment S2a, 20 participants were recruited. Two were eliminated for failing to meet the 75% catch trial accuracy criteria. For Experiment S2b, 40 participants were recruited. 18 were eliminated for failing to meet the 75% catch trial accuracy criteria. All participants were recruited from The George Washington University subject pool, reported normal or corrected-to-normal visual acuity, and were naïve to the purpose of the experiment.

Design and Procedure. In Experiment S1, cue validity (75%, 50%) was manipulated between subjects (Experiments S1a and S1b) and SOA was held constant at 500 ms. For each experiment, a 2 x 2 within-subjects factorial design was employed, with inferred object size (large vs. small) and cue-target relation (valid, invalid) as within-subject factors. Response times (RT) were the primary measure of interest. The letters T and L were used as target stimuli, while a T/L hybrid, which contains components of both a T and an L, was used as a distractor (see Fig 1a). The "T" target was mapped onto the "c" keyboard response, while the "L" target was mapped on the "m" keyboard response. Targets were considered valid if they appeared in the same location as the cue, and invalid if they appeared in the opposite location. In Experiments S1a and S1b, both the cue and ISI lasted 250 ms. Following the ISI, targets and distractors remained on the screen for 2 seconds, or until response. Participants were encouraged to response as quickly and as accurately as possible. For Experiments S2a and S2b, following the practice block, each participant completed ten experimental blocks, each consisting of 120 trials. Participants responded by pressing the "space" bar to indicate the presence of a target letter. Targets appeared on 80% of trials, while no target was present in the remaining 20% of trials, which served as catch trials. Also, in Experiment S2b, targets were displayed for 50 ms, followed by a 50 ms ISI, followed by masks at both possible target locations for 2 seconds or until response.

Supplementary Figure 1

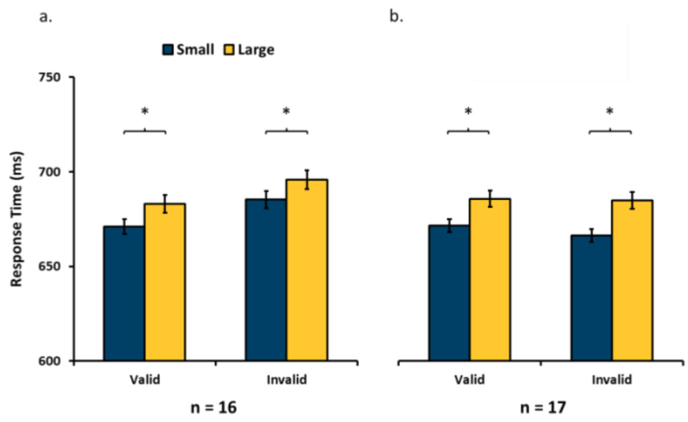

Figure S1: Data for Experiment S1
a.) Results for Experiment S1a – 75% Valid, Long SOA. A significant main effect of size is observed. b.) Results for Experiment S1b – 50 % Valid, Long SOA. Significant main effect of size is observed. All error bars represent standard error of the mean (SEM) corrected for within-subjects variance.

Supplementary Figure 2

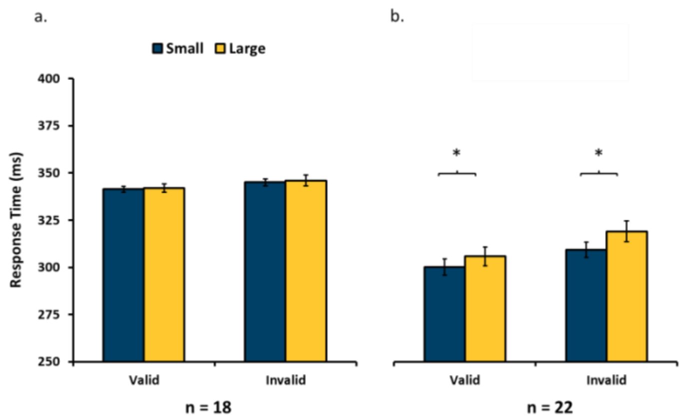

Figure S2: Data for Experiment S2
a.) Results for Experiment S2a – target detection task. No effects of size or validity are observed
b.) Results for Experiment S2b – target detection task with mask. Main effects of both size and validity are observed. All error bars represent standard error of the mean (SEM) corrected for within-subjects variance.

Supplementary Figure 3

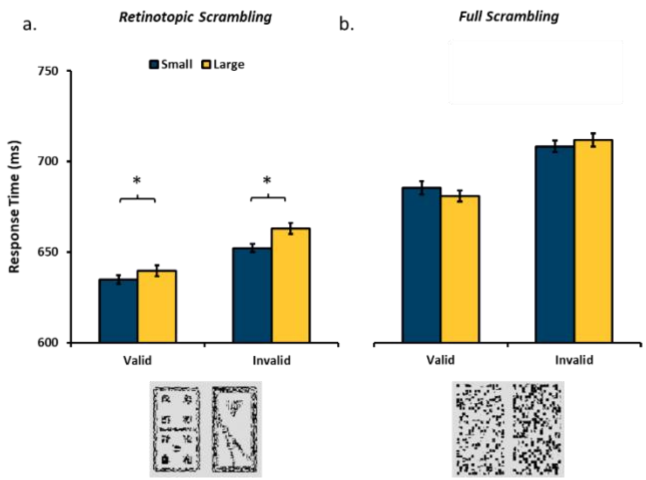

Figure S3: Supplemental Data for Experiment 2 broken down by scrambling type
a.) Results for Experiment 2a (retinotopic scrambling) b.) Results for Experiment 2b (full scrambling); Both experiments were conducted using a 75% valid cue and a 250 ms short SOA. All error bars represent standard error of the mean (SEM) corrected for within-subjects variance.

Supplementary Figure 4

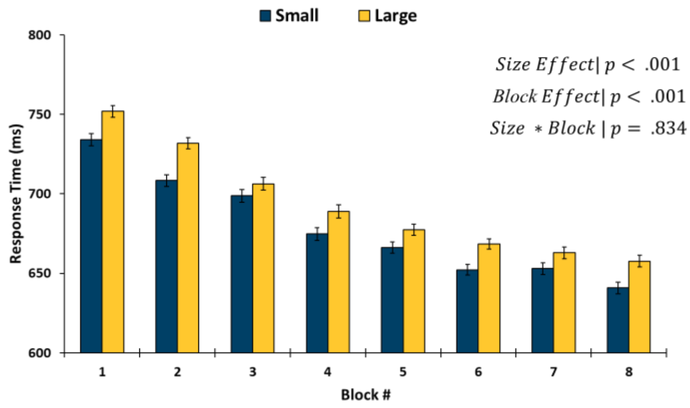

Figure S4: Size Effect across blocks for Experiment 1a
Results from block-by-block analysis for Experiment 1a. No decrease in the size effect was observed across all 8 blocks. All error bars represent standard error of the mean (SEM) corrected for within-subjects variance.

Supplementary Figure 5

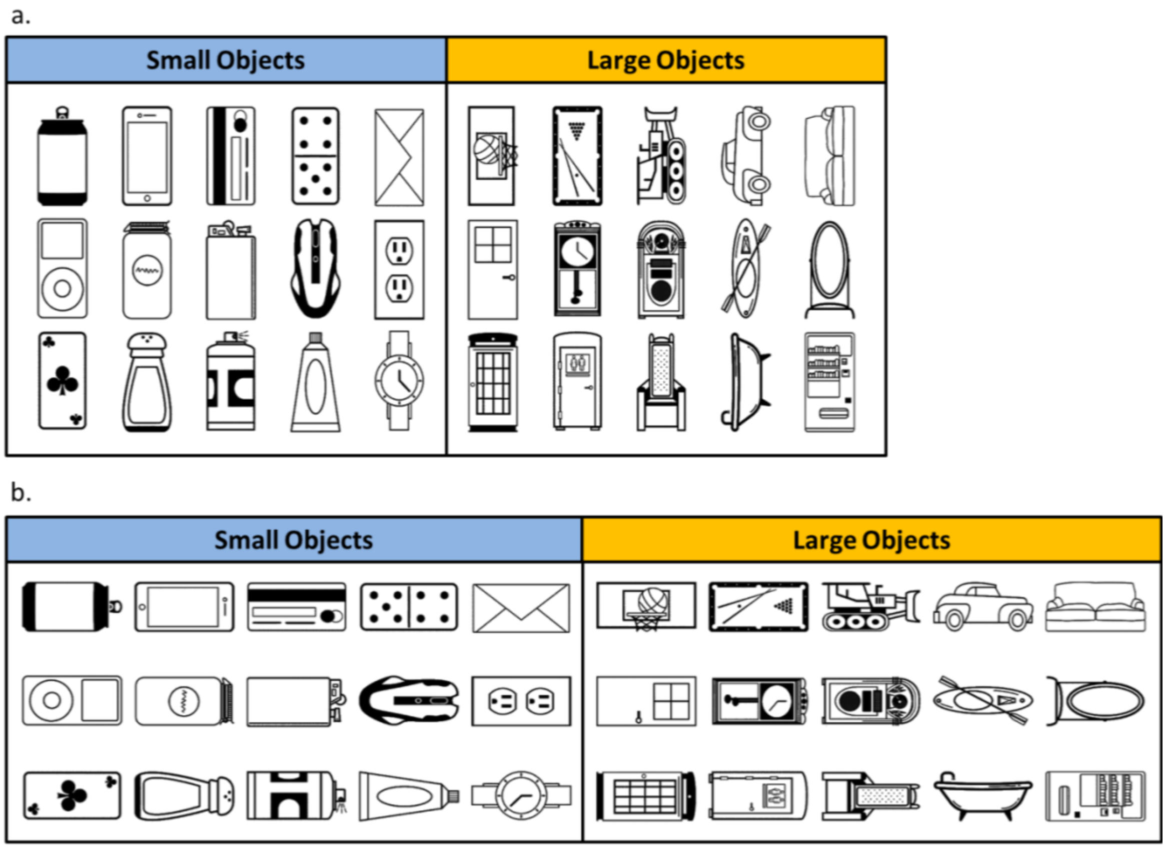

Figure S5: Object line-drawing stimuli
All object line drawings used in experiments, shown grouped according to size. There were 15 small items and 15 large items. All objects were presented in both a) *vertical* orientation and b) *horizontal* orientation, counterbalanced within subject in all experiments.

Supplementary Figure 6

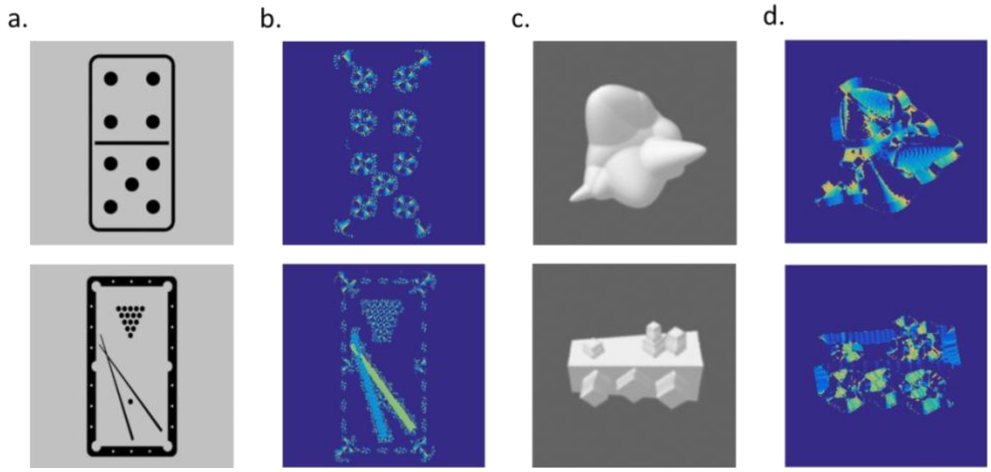

Figure S6: Steerable Pyramid Orientation Analysis to address object curvilinearity
a.) Examples of small (domino) and large (billiards) object stimuli used for experiments and steerable pyramid orientation analysis; b.) Examples of steerable pyramid-generated orientation maps for the corresponding small and large object stimuli; c.) Examples of 'smoothy' (top) and 'cuby' (bottom) objects used for validation of the steerable pyramid orientation analysis; d.) Examples of steerable pyramid-generated orientation maps for corresponding 'smooth' and 'cuby' object examples.