LETTERS
https://doi.org/10.1038/s41562-018-0485-2

nature human behaviour

Attention scales according to inferred real-world object size

Andrew J. Collegio[1], Joseph C. Nah[1], Paul S. Scotti[1,2] and Sarah Shomstein[1]*

Natural scenes consist of objects of varying shapes and sizes. The impact of object size on visual perception has been well-demonstrated, from classic mental imagery experiments[1], to recent studies of object representations reporting topographic organization of object size in the occipito-temporal cortex[2]. While the role of real-world physical size in perception is clear, the effect of inferred size on attentional selection is ill-defined. Here, we investigate whether inferred real-world object size influences attentional allocation. Across five experiments, attentional allocation was measured in objects of equal retinal size, but varied in inferred real-world size (for example, domino, bulldozer). Following each experiment, participants rated the real-world size of each object. We hypothesized that, if inferred real-world size influences attention, selection in retinal size-matched objects should be less efficient in larger objects. This effect should increase with greater attentional demand. Predictions were supported by faster identified targets in objects inferred to be small than large, with costlier attentional shifting in large than small objects when attentional demand was high. Critically, there was a direct correlation between the rated size of individual objects and response times (and shifting costs). Finally, systematic degradation of size inference proportionally reduced object size effect. It is concluded that, along with retinal size, inferred real-world object size parametrically modulates attention. These findings have important implications for models of attentional control and invite sensitivity to object size for future studies that use real-world images in psychological research.

Visual attention is a spatially deployed mechanism enabling efficient processing of a vast array of sensory information[3–6]. Whereas spatial selection is a fundamental aspect of attentional deployment, there is a rich history of attention research suggesting that objects can also be the unit of attentional selection (that is, above and beyond spatial contribution). In our environment, objects vary across multiple dimensions, including colour, shape, boundaries, size and semantic meaning. Such object properties enable an observer to distinguish between different objects, and also contribute to attentional deployment[7–14].

While the contributions of some object properties (for example, object boundaries, strength of object representations) to attention have been well studied, the contribution of object size to attentional selection remains ill-defined. Principally, the size of an object on the retina and the inferred size of an object are independent[15–19]. For example, the retinal size of a cup on your desk is larger than that of a car that can be seen from your window a block away. Changes in retinal object size that can occur due to a variety of reasons (for example, changes in depth)[20] do not correspondingly reflect changes in inferred object size, demonstrating the importance of object context[17,21] and the principle of size constancy. Images of real-world objects, such as those employed in our study, are rich in high-level information, including information about each object's real-world size[15,22–24]. Therefore, we hypothesize that spatial attentional selection might in fact be constrained by top-down inferences of object size.

Whether inferred object size affects perception has been at the centre of much object processing and mental imagery research since 1975. There is strong evidence suggesting that size of real-world object representations is inferred to be consistent with the true size of those objects in the environment, such that when participants are asked to imagine, draw or adjust objects to a preferred viewing size, large real-world objects are consistently represented as larger than small real-world objects in each task[1,15,25]. This consistent visual size information has been termed canonical size[15] and is dependent on inferences of real-world object size. Further demonstrating the impact of object size category on object perception, it was observed that canonical object size representations are automatic, that mid-level object properties such as texture and overall shape are sufficient to distinguish objects of different real-world sizes and that the topography of occipito-temporal cortex is organized to reflect canonical object size[2,22,23,26] in addition to an object's perceived size[18]. Interestingly, recent evidence also shows that observers are more likely to miss an object while performing a visual search of a natural scene if the object size is proportionally inconsistent with the scene in which it would typically appear (for example, a giant toothbrush placed on sink)[27,28] further supporting evidence suggesting that not only is an object's canonical size computed automatically, but that it has consequence for perception.

Whereas cortical organization seems to reflect differentiation by an object's inferred size in addition to the object's retinal size, whether such cortical segmentation has direct impact on attentional selection, rather than general perception, remains unclear. Work by Castiello and Umiltà[29] delineated the effect of retinal object size on attentional allocation, showing that, in a spatial cueing task, attention was less efficiently allocated within larger compared with smaller squares. Findings from this line of work suggest that attentional allocation within large objects is more diffuse while attention in smaller objects is more focused[29]. That attention is influenced by the physical size of the object is then well established, but does the same influence extend to inferred, rather than physical, size of objects? The potential mechanism is as follows: following computation of the retinal size of the object, this representation is then influenced, augmented, by the inferred knowledge of an object's real-world size either from occipito-temporal canonical size topography[2] or potentially through the contribution from parietal cortex object representations that are action-/size-sensitive[30,31]. These more complex object representations then serve as the basis for attentional selection.

[1]Department of Psychology, The George Washington University, Washington DC, USA. [2]Department of Psychology, The Ohio State University, Columbus, OH, USA. *e-mail: shom@gwu.edu

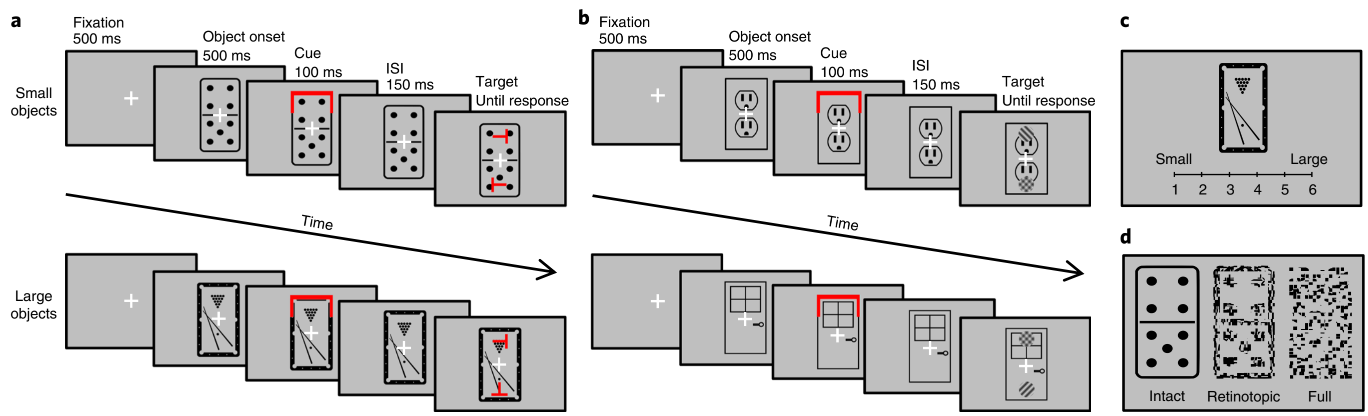

Fig. 1 | Experimental design. a, A modified Posner cueing task with task-irrelevant real-world objects was employed in experiments 1 and 2. Retinal size was fixed, while canonical size varied and participants identified target letters. **b**, Modified Posner cueing task employed for experiment 3, in which participants identified Gabor patch orientation. **c**, Post-experiment survey. Subjects ranked size of each object on a scale from 1–6. **d**, Examples of object stimuli (domino) used for scrambling, from left: intact, retinotopically scrambled, and fully scrambled. Credit: Domino (**a**,**d**), Agarunov Oktay-Abraham; pool table (**a**,**c**), Juan Pablo Bravo; outlet (**b**), Blaise Sewell (all icons reproduced from https://thenounproject.com/)

Here, we directly investigate whether inferred real-world object size, rather than simply retinal object size, influences the focus and shifting of attention within objects. If inferred object size has direct impact on attentional allocation, the efficiency of allocation should be modulated by whether an object is inferred to be canonically large (for example, a billiards table) or small (for example, a domino), as well as by the degree of attentional engagement. We hypothesized that if attention is influenced by, and thus constrained by, the functional localization of inferred object size in the occipito-temporal cortex[2,15], then attention will be less efficiently deployed in canonically large objects, in a manner referred to henceforth as attentional scaling. Conversely, if attention operates solely according to retinal object size, then attention should be allocated at the same rate, without attentional scaling, across all objects of equal retinal size (independent of inferred size).

Inferred object size was directly manipulated by creating two experimental sets of objects. Objects were classified, a priori, into two groups (canonically small and large) according to a pre-determined metric: objects classified as 'large' were body-sized or larger (for example, billiards table, door), while 'small' objects were those that could be held in hand (for example, domino, cell phone). A modified Posner spatial cueing paradigm was employed[5], with cues highlighting one end of a single line drawing of a real-world object presented in the centre of the screen (Fig. 1a). Line drawings were used to control for viewpoint and low-level information. According to several controls and metrics, there were no low-level differences between small and large objects (see Methods). Participants performed a letter T or L target discrimination task with targets presented either at the cued (valid) or uncued (invalid) location within the object (see Methods). Importantly, while objects subtended the same physical size, they varied in inferred real-world size (domino versus billiards table).

We hypothesized that the standard attentional cueing facilitation should be observed independent of an object's inferred size, such that targets will be detected faster and more accurately in validly cued locations compared to invalidly cued locations, termed the spatial validity effect and reflecting spatial selection[5]. Furthermore, and the contribution here, if attentional allocation scales with the inferred size of the object, target identification response time will be slower (and responses less accurate) in large real-world objects compared to smaller objects, in a diffuse/focused manner analogous to the effect of retinal object size delineated in prior research. Notably, the influence of object size may be more nuanced than a categorical difference between small and large objects, as observers have prior knowledge of real-world object size beyond category-level[15]. If allocation scales with increasing inferred real-world size, a direct relationship should be observed between real-world object size and response time, such that as objects increase in size, attentional allocation should slow accordingly. To test this, on completion of each experiment, participants rated the real-world size of each object, in random order, on a scale of one to six (one, very small; six, very large, see Fig. 1c). This served two purposes: (1) validating our a priori classification, and (2) elucidating a direct relationship between an individual's perception of object size and corresponding response times.

In the first set of experiments (experiments 1a and 1b), validity of the spatial cue (75 or 50% valid) was manipulated to test whether the top-down spatial predictability of the cue modulates the degree to which object size constrains attention. Given that inference of object's real-world size is a top-down process, it was important to dissociate top-down knowledge from top-down attention. If top-down inference is only triggered by top-down attentional allocation, object size should only constrain attention in the 75% valid condition. However, if object size has a general effect on attentional allocation, size-based attentional modulation should be observed independent of spatial predictability of the cue. To anticipate our findings, we observed direct evidence that attentional selection is modulated by inferred object size, independent of cue validity, in a manner consistent with the attentional scaling hypothesis proposed here (Fig. 2a,b). These results suggest more efficient deployment of attention in inferred smaller objects compared to larger objects, analogous to the impact of retinal object size on attention. Importantly, a strong correlation between participants' ratings of real-world object size and response times for individual objects strongly suggest that attentional allocation scales as a function of individual real-world object size, rather than just size category (Fig. 2).

In experiment 1a (75% valid cue, Fig. 2a), a three-way repeated-measures analysis of variance (ANOVA, $\alpha = 0.05$) with object size (large, small), cue validity (valid, invalid) and object orientation (vertical, horizontal) as within-subjects factors and response time as a dependent measure revealed main effects of inferred object size ($F(1, 19) = 14.50$, $P < 0.001$, partial eta squared, $\eta_p^2 = 0.433$, 90% confidence interval $(\mathrm{CI})\eta_p^2 = (0.138, 0.607)$, see Lakens for use of 90% CI for F-tests with $\alpha = 0.05$ (ref. [32])) and cue validity ($F(1, 19) = 22.94$, $P < 0.001$, $\eta_p^2 = 0.547$, 90% $\mathrm{CI}\eta_p^2 = (0.252, 0.689)$), such that attention was allocated less efficiently in canonically large than

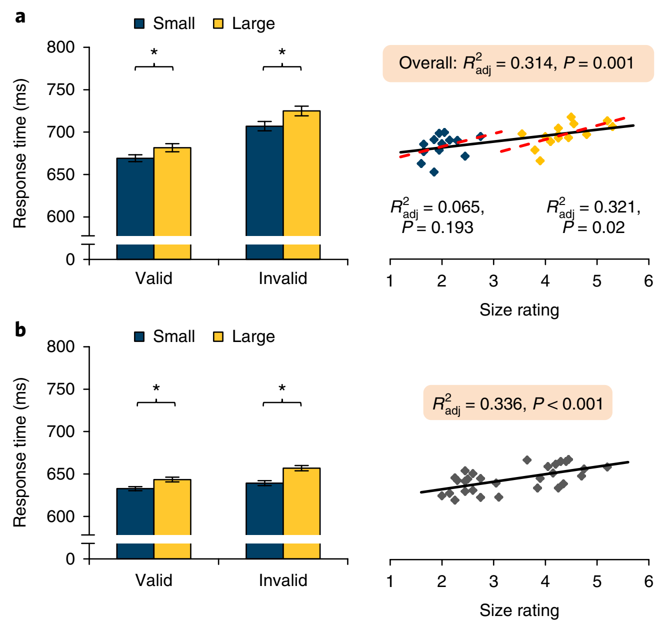

Fig. 2 | Data for experiment 1. a, Results for experiment 1a ($n=20$), 75% validity and short stimulus-onset asynchrony. Significant main effects of both size ($F(1, 19)=14.50$, $P<0.001$, $\eta_p^2=0.433$, 90% CI$\eta_p^2=(0.138, 0.607)$) and validity ($F(1, 19)=22.94$, $P<0.001$, $\eta_p^2=0.547$, 90% CI$\eta_p^2=(0.252, 0.689)$) were observed. No interaction between inferred size and cue validity was observed ($F(1, 19)=0.52$, $P=0.479$, $\eta_p^2=0.027$, 90% CI$\eta_p^2=(0.000, 0.209)$). In linear regression, ratings of object size ($n=28$, following removal of two outliers) were predictive of overall target response time ($R^2_{adj}=0.314$, $F(1, 26)=13.38$, $P=0.001$, $\eta_p^2=0.340$, 90% CI$\eta_p^2=(0.102, 0.517)$). **b**, Results for experiment 1b ($n=20$), with 50% validity and short stimulus-onset asynchrony. Significant main effects of both size ($F(1, 19)=49.01$, $P<0.001$, $\eta_p^2=0.721$, 90% CI$\eta_p^2=(0.486, 0.809)$) and validity ($F(1, 19)=5.60$, $P=0.029$, $\eta_p^2=0.227$, 90% CI$\eta_p^2=(0.014, 0.444)$) were observed. No interaction between inferred size and cue validity was observed ($F(1, 19)=1.21$, $P=0.285$, $\eta_p^2=0.060$, 90% CI$\eta_p^2=(0.000, 0.266)$). In linear regression, ratings of object size ($n=29$, following removal of one outlier) were predictive of overall target response time ($R^2_{adj}=0.336$, $F(1, 27)=15.19$, $P<0.001$, $\eta_p^2=0.360$, 90% CI$\eta_p^2=(0.122, 0.531)$). All error bars represent s.e.m. corrected for within-subjects variance. *$P<0.05$.

canonically small objects. No interaction between inferred size and cue validity was observed ($F(1, 19)=0.52$, $P=0.479$, $\eta_p^2=0.027$, 90% CI$\eta_p^2=(0.000, 0.209)$), indicating that while attention was allocated generally more slowly in large objects than in small objects, no evidence was observed that object size directly influenced attentional shifts away from the valid location. A main effect of object orientation ($F(1, 19)=51.65$, $P<0.001$, $\eta_p^2=0.731$, 90% CI$\eta_p^2=(0.503, 0.816)$) was observed, suggesting that attention was allocated more efficiently across horizontally oriented objects than vertically oriented objects. Imperatively, neither the two-way interaction between inferred size and object orientation ($F(1, 19)=2.40$, $P=0.138$, $\eta_p^2=0.112$, 90% CI$\eta_p^2=(0.000, 0.331)$) nor the three-way interaction between inferred size, cue validity and object orientation ($F(1, 19)=0.06$, $P=0.810$, $\eta_p^2=0.003$, 90% CI$\eta_p^2=(0.000, 0.115)$), were significant, providing no evidence that rotation away from an object's canonical orientation modulated size-based attentional scaling. The two-way interaction between cue validity and object orientation was also non-significant ($F(1, 19)=0.89$, $P=0.358$, $\eta_p^2=0.045$, 90% CI $\eta_p^2=(0.000, 0.242)$). It should also be noted that should there have been any orientation bias influencing size effects (whether canonical orientation or general orientation peculiarities) the orientation factor would have interacted with the size effect.

In a follow-up analysis, participants' size ratings were linearly regressed against overall target response times to measure the degree to which participants' perception of object size predicted attentional orienting. Crucially, individuals' ratings of object size were predictive of overall target response time (overall significant regression: $R^2_{adj}=0.314$, $F(1, 26)=13.38$, $P=0.001$, $\eta_p^2=0.340$, 90% CI$\eta_p^2=(0.102, 0.517)$, Fig. 2a) following removal of two outliers (residuals greater than two standard deviations from the mean). To assess whether this significant relationship was driven solely by a bimodal distribution (large and small objects), we performed a median split on response times according to participants' size ratings for each object and conducted linear regression on each half of the data. This split yielded two results: (1) there was a perfect correspondence between the pre-experiment size classification and participant ratings (median split on the data perfectly split objects into small and large (Fig. 2a; $t(26)=-14.45$, $P<0.001$ (two-tailed), Cohen's $d=-5.462$, 95% CI$_d=(-7.102, -3.799)$) and (2) size ratings were strongly predictive of response times within the large object group, ($R^2_{adj}=0.321$, $F(1, 12)=7.13$, $P=0.020$, $\eta_p^2=0.373$, 90% CI$\eta_p^2=(0.037, 0.590)$). Additionally, a tertile split across all objects, conducted to reduce the effect of moderately sized objects on regression, revealed a significant relationship between size rating and response time within the smallest object group ($R^2_{adj}=0.385$, $F(1, 7)=6.01$, $P=0.044$, $\eta_p^2=0.462$, 90% CI$\eta_p^2=(0.008, 0.676)$) following removal of one outlier (residual greater than two standard deviations from the mean). Taken together, these results indicate that inferred knowledge of real-world object size directly predicts the efficiency of attentional allocation within objects of varying sizes. These findings are consistent with our prediction of attentional scaling according to object size and, imperatively, dovetail with the finding that attention is more diffuse and allocated less efficiently to locations within retinally large objects[29]. In addition to response time, performance accuracy was also analysed for all experiments and is reported in the Supplementary Materials (see Supplementary Materials, experiments 1–3). Accuracy results were consistent with those reported in response time. Namely, no effect of size was observed, suggesting that all significant effects were absorbed by the response time measure and no speed-accuracy tradeoffs were observed, indicating consistency with response time results.

In experiment 1a, more efficient attentional deployment was observed in small objects compared to larger objects, consistent with our original hypothesis that attention scales according to inferred real-world object size. The observed top-down attentional scaling effects could be partly driven by top-down attentional signal from the spatial predictability of the cue. To test this possibility, cue validity was decreased to 50%, eliminating top-down spatial bias. While spatial validity effects should be reduced with the decrease in spatial predictability, the effect of size should remain. In experiment 1b (50% valid cue, Fig. 2b), a three-way repeated-measures ANOVA for response time again revealed significant main effects of inferred size ($F(1, 19)=49.01$, $P<0.001$, $\eta_p^2=0.721$, 90% CI$\eta_p^2=(0.486, 0.809)$) and validity ($F(1, 19)=5.60$, $P=0.029$, $\eta_p^2=0.227$, 90% CI $=(0.014, 0.444)$), with no observed interaction between size and validity ($F(1, 19)=1.21$, $P=0.285$, $\eta_p^2=0.060$, 90% CI $\eta_p^2=(0.000, 0.266)$). Additionally, a main effect of object orientation was observed ($F(1, 19)=69.78$, $P<0.001$, $\eta_p^2=0.786$, 90% CI $\eta_p^2=(0.593, 0.853)$), but the two-way interaction between size and orientation ($F(1, 19)=0.03$, $P=0.855$, $\eta_p^2=0.002$, 90% CI$\eta_p^2=(0.000, 0.088)$), as well as the three-way interaction between size, validity and orientation ($F(1, 19)=0.08$, $P=0.787$, $\eta_p^2=0.004$, 90% CI$\eta_p^2=(0.000, 0.127)$) did not reach significance, again providing no evidence that object orientation modulated the effect of object size on attention.

These results replicated findings from experiment 1a, indicating an effect of size-based attentional scaling: again, individuals' ratings of object size were predictive of response time (overall significant linear regression: $R^2_{adj} = 0.336$, $F(1, 27) = 15.19$, $P < 0.001$, $\eta_p^2 = 0.360$, 90% CI$\eta_p^2 = (0.122, 0.531)$) following removal of one outlier (residual greater than two standard deviations from the mean). Additionally, a between-experiment four-way ANOVA with experiment (75% valid, 50% valid) as a between-subject variable, inferred object size, cue validity and object orientation as within-subject factors and response time as a dependent measure revealed main effects of size ($F(1, 38) = 42.82$, $P < 0.001$, $\eta_p^2 = 0.530$, 90% CI$\eta_p^2 = (0.332, 0.647)$), validity ($F(1, 38) = 28.53$, $P < 0.001$, $\eta_p^2 = 0.429$, 90% CI$\eta_p^2 = (0.223, 0.567)$) and orientation ($F(1, 38) = 115.67$, $P < 0.001$, $\eta_p^2 = 0.753$, 90% CI$\eta_p^2 = (0.621, 0.812)$) such that, across both experiments, attention was less efficiently allocated in larger objects, in invalid locations and in the vertical orientation. A two-way interaction between validity and experiment ($F(1, 38) = 10.43$, $P = 0.003$, $\eta_p^2 = 0.215$, 90% CI$\eta_p^2 = (0.050, 0.381)$) was significant, consistent with the prediction that validity effect should be larger in the 75% valid condition (experiment 1a). A two-way interaction between validity and object orientation was also significant ($F(1, 38) = 7.06$, $P = 0.011$, $\eta_p^2 = 0.157$, 90% CI$\eta_p^2 = (0.021, 0.322)$), suggesting that validity effects are stronger in the horizontal orientation. Importantly, the influence of size on attention remained unaffected by experiment (size by experiment interaction, $F(1, 38) = 0.07$, $P = 0.793$, $\eta_p^2 = 0.002$, 90% CI$\eta_p^2 = (0.000, 0.068)$) or orientation (size by orientation interaction, $F(1, 38) = 1.85$, $P = 0.181$, $\eta_p^2 = 0.046$, 90% CI$\eta_p^2 = (0.000, 0.185)$). In experiments 1a and 1b, as well as two additional supplemental experiments, in which the time-course and robustness of these effects are tested (see Supplementary Materials, Experiments 1 and 2), we observe that attention is deployed less efficiently in canonically larger compared to canonically smaller objects. It could also be argued that the observed size effect is driven by the surprise or novelty associated with seeing small objects that share the same retinal size as the large objects. If this were the case, it would be anticipated that the effect of size would diminish with time spent on the experiment (that is, blocks). It was observed, however, that the effect of size remained stable over time (see Supplementary Fig. 4)

Imperative to the original hypothesis of attentional scaling, inferred knowledge of object size has been shown to modulate deployment of attention within objects. If inferred object size is driving this difference, then a systematic reduction in object identifiability should result in a parametric decrease in the contribution of inferred object size to attentional allocation and thus lead to diminished differences in response time between large and small objects. In an additional set of two experiments (experiments 2a and 2b), objects were scrambled to parametrically reduce knowledge of object size (Fig. 1d). Two forms of image scrambling were employed: (1) retinotopic scrambling, in experiment 2a, preserved the gestalt of the object, such that the object might still be recognizable, despite a lack of low-level fidelity due to scrambled pixels and (2) full scrambling, in experiment 2b, in which pixels were scrambled within an image irrespective of the gestalt of the original object. In this set of experiments, with 75% validity and a cue-to-target stimulus-onset asynchrony of 250 ms (as in experiment 1a), it was predicted that the magnitude of the size-driven effect should decrease parametrically with an increase in scrambling. Specifically, retinotopic scrambling will moderately reduce the effect of size on attentional deployment, while full scrambling will reduce the effect above and beyond retinotopic scrambling, due to the greater degree of object information loss in full scrambling compared to retinotopic scrambling.

A between-experiment three-way repeated-measures ANOVA was conducted with scrambling level (no scrambling, experiment 1a; retinotopic scrambling, experiment 2a and full scrambling, experiment 2b) as a between-subject variable and object size (large, small) as a within-subject factor (Fig. 3). ANOVA revealed a significant effect of object size ($F(1, 57) = 17.18$, $P < 0.001$, $\eta_p^2 = 0.232$, 90% CI$\eta_p^2 = (0.085, 0.367)$) and importantly, a significant two-way interaction between object size and scrambling type ($F(2, 57) = 6.35$, $P = 0.003$, $\eta_p^2 = 0.182$, 90% CI$\eta_p^2 = (0.041, 0.308)$) was observed, such that size effect was only significant in the no-scramble condition ($F(1, 19) = 14.40$, $P = 0.001$, $\eta_p^2 = 0.431$, 90% CI$\eta_p^2 = (0.137, 0.606)$) and the retinotopic scrambling experiment ($F(1, 19) = 8.58$, $P = 0.009$, $\eta_p^2 = 0.311$, 90% CI$\eta_p^2 = (0.052, 0.514)$; Fig. 3 and Supplementary Fig. 3a), but not in the full scramble experiment ($F(1,19) = 0.05$, $P = 0.826$, $\eta_p^2 = 0.003$, 90% CI$\eta_p^2 = (0.000, 0.108)$; Fig. 3 and Supplementary Fig. 3b). Interestingly, dovetailing with the findings, the retinotopically scrambled objects were rated consistently with our pre-defined large/small groups with 93% accuracy, suggesting a somewhat preserved size inference. On the other hand, size rating consistency was at chance (53.33%) for the fully scrambled objects. Taken together, these results indicate that the size effect, as predicted, was reduced (Fig. 3) in the full scramble condition (no canonical size inference, experiment 2b, with difference in response times between large and small objects of $\Delta = -0.67$ ms) relative to the retinotopic condition (somewhat preserved size inference, experiment 2a, $\Delta = 7.73$ ms) and the no-scramble condition (intact size inference, experiment 1a, $\Delta = 13.79$ ms).

Experiment 3 was designed with three aims: (1) internal replication with a different target discrimination task; (2) testing the effect of increased attentional demand; and (3) addressing a possible confound of target spatial frequency. One of the important aspects of object size is a consideration that objects of varying size also vary in terms of the spatial frequency of their features (that is, large objects, low spatial frequency; small objects, high spatial frequency)[33,34] and that the results demonstrated in experiment 1 and 2 could partly be confounded by a mismatch between target spatial frequency and object spatial frequency. High spatial frequency targets may be processed less efficiently in low spatial frequency, large objects, leading to the attentional scaling effect. To address this potential confound, and to provide additional internal replication of our original results, in experiment 3 participants performed a target identification task as in experiments 1 and 2, but the nature of the target was different. In this experiment, instead of performing a T or L discrimination task, participants identified the orientation of a target Gabor patch (45° or −45°) (Fig. 1b). Additionally, the spatial frequency of the target Gabor patches varied between four cycles per degree (low frequency) and eight cycles per degree of visual angle (high spatial frequency) in a within-subjects design. If attention is affected by the spatial frequency of targets as they relate to object spatial frequency, a congruency effect should be observed, such that attention will be allocated more efficiently when target spatial frequency matches object spatial frequency (for example, high spatial frequency, small object; low spatial frequency, large object). Alternatively, if attention is only constrained by inferred object size, as we have argued up to this point, we expect to observe two findings: (1) a main effect of object size, same as in prior experiments and (2) an interaction with frequency such that effect of size should be large in a more difficult, high-frequency (less discriminability), task.

For experiment 3, a repeated-measures ANOVA on response times was conducted with object size (large, small), cue validity (valid, invalid), object orientation (vertical, horizontal) and target spatial frequency (low, high) as within-subject factors (Fig. 4a). ANOVA revealed main effects of inferred object size ($F(1, 19) = 25.11$, $P < 0.001$, $\eta_p^2 = 0.569$, 90% CI$\eta_p^2 = (0.278, 0.704)$) with faster response times to small objects; cue validity ($F(1, 19) = 29.35$, $P < 0.001$, $\eta_p^2 = 0.607$, 90% CI$\eta_p^2 = (0.325, 0.731)$) with faster response times to validly cued targets; object orientation ($F(1, 19) = 6.20$, $P = 0.022$, $\eta_p^2 = 0.246$, 90% CI$\eta_p^2 = (0.021, 0.460)$) with faster response times in horizontally oriented objects and target spatial frequency ($F(1, 19) = 43.99$, $P < 0.001$, $\eta_p^2 = 0.698$, 90% CI$\eta_p^2 = (0.452, 0.793)$), with faster response times for the low frequency targets. In addition to replicating main effects of object size and validity, the main effect

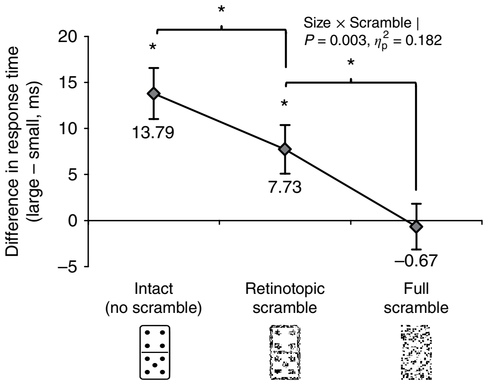

Fig. 3 | Data for experiment 2, comparison across all scrambling conditions. Scrambling of objects reduced contribution of object size to attention. As objects become more difficult to identify, the contribution of size is reduced. A between-experiment repeated-measures ANOVA conducted across experiment 1a ($n = 20$), experiment 2a ($n = 20$) and experiment 2b ($n = 21$) revealed a significant effect of object size ($F(1, 57) = 17.18$, $P < 0.001$, $\eta_p^2 = 0.232$, 90% CI$\eta_p^2 = (0.085, 0.367)$) and a significant two-way interaction between object size and scrambling type ($F(2, 57) = 6.35$, $P = 0.003$, $\eta_p^2 = 0.182$, 90% CI$\eta_p^2 = (0.041, 0.308)$). The size effect was only significant in the no-scramble condition ($F(1, 19) = 14.40$, $P = 0.001$, $\eta_p^2 = 0.431$, 90% CI$\eta_p^2 = (0.137, 0.606)$) and the retinotopic scrambling experiment ($F(1, 19) = 8.58$, $P = 0.009$, $\eta_p^2 = 0.311$, 90% CI$\eta_p^2 = (0.052, 0.514)$), but not in the full scramble experiment ($F(1,19) = 0.05$, $P = 0.826$, $\eta_p^2 = 0.003$, 90% CI$\eta_p^2 = (0.000, 0.108)$). All error bars represent s.e.m. corrected for within-subjects variance. *$P < 0.05$. Credit: Domino, Agarunov Oktay-Abraham (reproduced from https://thenounproject.com/)

of target spatial frequency indicates that participants are slower for high spatial frequency targets compared to low frequency targets, regardless of object size, indicating this is a result of task difficulty and not spatial frequency mismatch.

A two-way interaction between inferred object size and cue validity was marginally significant ($F(1, 19) = 4.09$, $P = 0.057$, $\eta_p^2 = 0.177$, 90% CI$\eta_p^2 = (0.000, 0.398)$). Crucially, a significant three-way interaction between inferred object size, cue validity and target spatial frequency was also significant ($F(1, 19) = 7.99$, $P = 0.011$, $\eta_p^2 = 0.296$, 90% CI$\eta_p^2 = (0.044, 0.502)$), indicating that target spatial frequency (low or high) modulates the degree to which attentional shifts are constrained by object size (Fig. 4a). These results suggest that, in addition to general attentional scaling resulting from a more diffuse focus of attention in large than small objects (observed in experiments 1 and 2), as predicted, attentional shifts between cue valid and invalid locations are also modulated by inferred object size, such that cue validity effects are larger in large objects ($\Delta = 56.6$ ms) compared to small objects ($\Delta = 45.8$ ms). When target spatial frequency is high, the task is more difficult and attentionally demanding, therefore modulation of attentional shifts by object size is observed (large objects, $\Delta = 70.8$ ms; small objects, $\Delta = 40.3$ ms). This was further supported by a significant relationship observed between object size rating and cost of attention shift for high spatial frequency targets ($R^2_{adj} = 0.210$, $F(1,28) = 8.73$, $P = 0.006$, $\eta_p^2 = 0.238$, 90% CI$\eta_p^2 = (0.043, 0.424)$, Fig. 4c). When target spatial frequency is low, modulation of attentional shifts by object size is reduced (large objects, $\Delta = 43.0$ ms; small objects, $\Delta = 52.5$ ms) and attention is constrained by inferred object size in a more general, scaling fashion (that is, less efficient in large objects). No relationship was observed between object size rating and attentional shifting costs with low spatial frequency targets ($R^2_{adj} = -0.014$, $F(1,28) = 0.59$, $P = 0.450$, $\eta_p^2 = 0.016$, 90% CI$\eta_p^2 = (0.000, 0.162)$, Fig. 4b). No other effects, including any congruency effect of inferred object size and target spatial frequency, were observed (size by frequency, $F(1, 19) = 0.04$, $P = 0.838$, $\eta_p^2 = 0.002$, 90% CI$\eta_p^2 = (0.000, 0.099)$; validity by frequency, $F(1, 19) = 1.98$, $P = 0.176$, $\eta_p^2 = 0.094$, 90% CI$\eta_p^2 = (0.000, 0.311)$; validity by orientation, $F(1, 19) = 0.71$, $P = 0.793$, $\eta_p^2 = 0.004$, 90% CI$\eta_p^2 = (0.000, 0.227)$; size by frequency by orientation, $F(1, 19) = 0.01$, $P = 0.946$, $\eta_p^2 = 0.000$, 90% CI$\eta_p^2 = (0.000, 0.041)$ and validity by frequency by orientation, $F(1, 19) = 0.53$, $P = 0.477$, $\eta_p^2 = 0.027$, 90% CI$\eta_p^2 = (0.000, 0.210)$). The four-way interaction between size, validity, frequency and orientation was also nonsignificant ($F(1, 19) = 0.44$, $P = 0.513$, $\eta_p^2 = 0.023$, 90% CI$\eta_p^2 = (0.000, 0.200)$), providing no evidence for an effect of object orientation on the crucial three-way interaction between size, validity and frequency (see Supplementary Material, orientation analyses section for additional lower-level orientation interactions).

More sluggish processing of spatial locations within canonically larger objects compared to smaller objects suggests that inferred real-world size modulates attentional allocation, in a scaling fashion. Critically, this effect is modulated by the degree of top-down knowledge available about the object, as evidenced by the reduction of size-based attentional modulation according to the degree of information loss about object size due to image scrambling (Fig. 3). Finally, under attentionally demanding conditions, shifts of attention are also constrained by inferred object size, with slower attentional shifts within large objects.

As much progress has been made in the understanding of how objects guide attention, it is becoming imperative to move toward real-world objects to better understand object properties that facilitate attentional selection. Real-world objects contain a rich array of high-level information that may also constrain attentional selection, alongside low-level features of more basic, two-dimensional objects, such as the rectangles in a traditional two-rectangle display[10]. These findings provide evidence of high-level size-based attentional constraint, in the form of attentional scaling within line drawings of real-world objects.

While novel with respect to equal retinal size, the findings dovetail with prior research, suggesting that attention scales according to both the physical and the perceived size of an object[16–18,29,35,36]. One elegant study aimed to disentangle the complex relationship between retinal and perceived object size through the use of the Ponzo illusion and rectangles, concluding that perceived object size constrains attentional selection[16], and that, in concordance with work conducted by Castiello and Umiltà, attention is more tightly focused in objects perceived to be smaller. The study presented here concludes that attention is not influenced by perceived size alone, but by an inferred knowledge of object size in the real-world. This inferred canonical size, an inherent high-level property of real-world objects, constrains topographic organization within object-selective cortical regions, but also influences how attention is deployed. Both the focus and shifting of attention are affected by inferred real-world object size, rather than retinal object size alone.

Mechanistically, our findings point to a possible functional relationship between object-selective cortical regions of occipito-temporal cortex, mapped according to inferred object size and parietal cortex and the attentional priority map, in that real-world object size perception affects attentional deployment[2,16,17]. This has yet to be explored and the time is ripe for pursuits of the neural basis of size-based attention scaling. It should be noted that the real-world size-specific functional response in the inferior temporal cortex lies in good register with the known maps of eccentricity, in many cases extending them. Such retinotopic constraints on the high-level visual representations are widespread in both human and monkey[37] for both animate and inanimate objects. Importantly, this organization is present even in studies where the actual retinal size and position is held constant, as in our study. Thus, our results may be related to the presence of this inferred real-world object size organization in the inferior temporal cortex. The precise mechanism relating these two factors will need to be the subject of further studies.

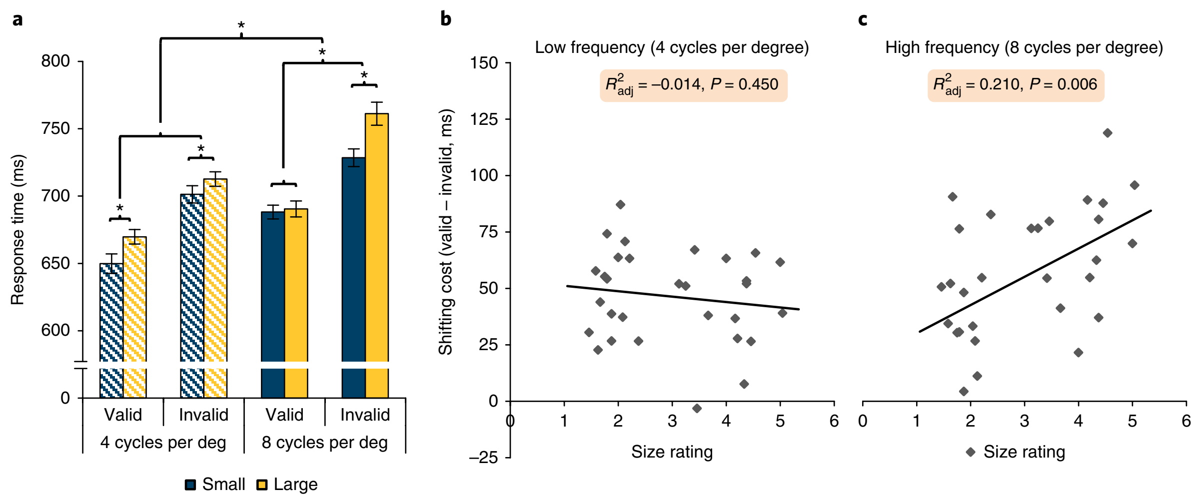

Fig. 4 | Data for experiment 3. **a**, Results for experiment 3 ($n=20$), in which target Gabor spatial frequency varied between four cycles per degree and eight cycles per degree. Low frequency data shown in striped bars; high frequency data shown in solid bars. Repeated-measures ANOVA revealed main effects of inferred object size ($F(1, 19)=25.11$, $P<0.001$, $\eta_p^2=0.569$, 90% CI$\eta_p^2=(0.278, 0.704)$), cue validity ($F(1, 19)=29.35$, $P<0.001$, $\eta_p^2=0.607$, 90% CI$\eta_p^2=(0.325, 0.731)$) and target spatial frequency ($F(1, 19)=43.99$, $P<0.001$, $\eta_p^2=0.698$, 90% CI$\eta_p^2=(0.452, 0.793)$). A two-way interaction between inferred object size and cue validity was marginally significant ($F(1, 19)=4.09$, $P=0.057$, $\eta_p^2=0.177$, 90% CI$\eta_p^2=(0.000, 0.398)$) and a three-way interaction between inferred object size, cue validity and target spatial frequency was significant ($F(1, 19)=7.99$, $P=0.011$, $\eta_p^2=0.296$, 90% CI$\eta_p^2=(0.044, 0.502)$). All error bars represent s.e.m. corrected for within-subjects variance. $^*P<0.05$. **b**, No correlation was observed between object size ratings ($n=30$) and shift costs with low spatial frequency targets ($R^2_{adj}=-0.014$, $F(1,28)=0.59$, $P=0.450$, $\eta_p^2=0.016$, 90% CI$\eta_p^2=(0.000, 0.162)$). **c**, Strong positive correlation between object size ratings ($n=30$) and shift costs with high spatial frequency targets ($R^2_{adj}=0.210$, $F(1,28)=8.73$, $P=0.006$, $\eta_p^2=0.238$, 90% CI$\eta_p^2=(0.043, 0.424)$).

This research also holds strong implications for scene processing literature, given that scene processing requires attentional allocation to objects of varying sizes. According to an observer's prior experience with an object, the context of an object may provide information about an object's size, which will impact the manner in which the object engages attentional selection. High-level object representations and object context modulate attentional allocation, and novel work must be conducted to investigate the role of real-world objects and inferred size on attention, both space- and object-based. Additional high-level properties of objects, such as how large an object is imagined to be via mental imagery or whether an object is known to be animate, may also constrain attentional deployment in a scaling manner. Previous work suggests that if participants are told an object is a toy, the canonical size of the toy object is smaller than the object it represents[22], suggesting a change in the mental representation of the object, while other work suggests that imagined objects are also topographically organized according to real-world size in occipito-temporal cortex[26]. However, animate objects do not show a size-based topographic organization[38]. If this topographic organization is necessary for size-based attentional scaling, animate objects should not elicit a scaled attentional response. Taken together, our findings and these implications suggest that caution should be taken when using object images in psychological research, as inferred object size constrains processing of and attention to objects.

Methods

Participants. For experiment 1a, 25 participants (8 male; average age: 19.4 years) were recruited. Five participants were excluded for failing to meet the 90% accuracy criteria. For experiment 1b, 25 participants (7 male; average age: 19.4 years) were recruited. Five participants were excluded for failing to meet the 90% accuracy criteria. For experiment 2a, 26 participants (6 male; average age: 19.2 years) were recruited. Six participants were excluded for failing to meet the 90% accuracy criteria. For experiment 2b, 26 participants (8 male; average age: 19.7 years) were recruited. Five participants were excluded for failing to meet the 90% accuracy criteria. For experiment 3, 25 participants (9 male; average age: 19.6 years) were recruited. Five participants were excluded for failing to meet the 90% accuracy criteria. All participants were recruited from The George Washington University participant pool, gave written informed consent according to The George Washington University's institutional review board, reported normal or corrected-to-normal visual acuity and were naïve to the purpose of the experiment. All experiments consisted of completely independent sets of participants. No statistical methods were used to pre-determine sample sizes but our sample sizes are larger than those reported in previous publications[5,16,29]. For each experiment, subjects were provided with instructions by the experimenter. Data collection and analysis were not performed blind to the conditions of the experiments.

Apparatus and stimuli. For all experiments, stimuli were presented on a 19" Dell 1908FP colour liquid crystal display monitor with a refresh rate of 60 Hz, positioned at a distance of 62 cm from the viewer. All experiments were generated and presented using PsychoPy v1.82[39,40]. Stimuli consisted of line drawings of real-world, everyday objects drawn both from The Noun Project, an online repository of object icons and clip art (https://thenounproject.com), as well as from a set of images rendered by the authors. The object stimuli consisted of 15 line drawings of small objects and 15 line drawings of large objects. The small objects were as follows: soda can, cell phone, cigarette lighter, credit card, domino, envelope, iPod, mason jar, computer mouse, electric outlet, playing card, salt shaker, aerosol canister, toothpaste tube and watch. The large objects were: basketball hoop, billiards table, bulldozer, car, sofa, door, grandfather clock, jukebox, kayak, mirror, phone booth, porta-potty, throne, bathtub and vending machine. All objects are displayed in the Supplementary Materials (Supplementary Fig. 5).

For each experiment, the display consisted of a single central fixation point, which subtended a visual angle of 1° × 1°, and a single central line drawing of an object, subtending a visual angle of 4° × 8° oriented either vertically (as shown in Supplementary Fig. 5a) or horizontally (rotated 90° clockwise, as shown in Supplementary Fig. 5b). Object orientation was fully counterbalanced within-subjects for each experiment and all 30 objects were displayed with equal probability. Following the presentation of an object, a cue was presented at either end of the object, with equal probability. For experiments 1 and 2, following the offset of the cue, a single target letter (T or L, equiprobable) always appeared at either of the two ends of the object, while a distractor letter (T/L hybrid) always appeared at the opposite end of the object. For experiment 3, a single Gabor patch oriented 45° or −45° equiprobably appeared at either of the two ends of the object, while a distractor chequerboard appeared at the opposite end. Gabor spatial frequency equiprobably varied between four cycles per degree of visual angle and eight cycles per degree of visual angle. Note that we used a factor of two rather than factor of around ten (more reflective of real-world differences) for spatial frequencies with the goal of increasing spatial attention demands and not necessarily matching real-world frequency differences. Targets and distractors each subtended a visual angle of 1° × 1° and were presented with a centre-to-centre distance of 5°. Line drawings were rendered in black (red, green, blue: RGB = 0, 0, 0) against a light grey background (RGB = 192, 192, 192). Cues, targets and distractors were rendered in red (RGB = 255, 0, 0) in experiments 1 and 2, while targets and distractors were rendered in black (RGB = 0, 0, 0) and white (RGB = 255, 255, 255) in experiment 3. Additionally, all stimuli were controlled for overall luminance and tested to ensure that small and large objects did not differ.

To demonstrate that size effects in the two main and two supplementary experiments are truly driven by inferred object size knowledge, rather than any low-level group or individual differences between canonically large and small objects, low-level features (specifically, luminance and retinal size) of each object were directly compared[14]. A 100-bin histogram based on luminance values was created for each object. Intensities greater than or equal to 128 were discounted from analysis, as these pixels belonged to the grey background of each image. Histograms were normalized and all objects were compared to each other and to the group mean. The pixel counts for each object were compared across conditions to account for potential differences in retinal object size. Two-tailed, two-sample t-tests for both relative pixel intensity ($t(28) = -0.80$, $P = 0.428$, $d = -0.293$, 95% $\mathrm{CI}_d = (-1.011, 0.429)$) and retinal object size ($t(28) = 1.40$, $P = 0.174$, $d = 0.509$, 95% $\mathrm{CI}_d = (-0.223, 1.233)$) revealed no significant differences between groups. Two-tailed one-sample t-tests comparing relative pixel intensity (small objects: $t(14) = 0.56$, $P = 0.584$, $d = 0.145$, 95% $\mathrm{CI}_d = (-0.367, 0.651)$; large objects: $t(14) = -0.58$, $P = 0.574$, $d = -0.149$, 95% $\mathrm{CI}_d = (-0.655, 0.363)$) and retinal object size (small objects: $t(14) = -0.90$, $p = 0.384$, $d = -0.232$, 95% $\mathrm{CI}_d = (-0.741, 0.285)$; large objects: $t(14) = 1.11$, $P = 0.288$, $d = 0.285$, 95% $\mathrm{CI}_d = (-0.236, 0.797)$) to the overall group average (average pixel intensity: 115.72; average retinal object size: 10,250.73 pixels) revealed no individual object differed from the group mean, failing to provide evidence that less efficient allocation of attention within larger objects was driven by systematic differences in low-level features between large and small objects.

To control for any further possible conflation between inferred object size and low-level features, individual ratings of inferred real-world object size were also regressed against average luminance value for each object and luminance values were regressed with response time. Linear regression was done to test whether average object luminance was predictive of performance (that is, worse performance in darker objects), despite a lack of group differences in low-level features. For experiment 1a, neither size by luminance ($R^2_{\mathrm{adj}} = -0.014$, $F(1,28) = 0.60$, $P = 0.447$, $\eta_p^2 = 0.021$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.163)$), nor luminance by response time ($R^2_{\mathrm{adj}} = -0.005$, $F(1,28) = 0.86$, $P = 0.361$, $\eta_p^2 = 0.030$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.181)$) relationships were observed, indicating a lack of evidence that the average luminance of each object, and not the inferred object size, was predictive of response time. For experiment 1b, neither size by luminance ($R^2_{\mathrm{adj}} = -0.026$, $F(1,28) = 0.25$, $P = 0.619$, $\eta_p^2 = 0.009$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.128)$), nor luminance by response time ($R^2_{\mathrm{adj}} = -0.020$, $F(1,28) = 0.43$, $P = 0.515$, $\eta_p^2 = 0.015$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.149)$) relationships were observed.

While low-level features (luminance, retinal object size) were controlled across objects, it remains possible that clutter, another low-level feature, is responsible for differences between canonically large and small objects. Due to remaining concerns about possible clutter differences between object groups, a clutter analysis assessing contrast was conducted across all objects and then analysed for differences between groups[41]. Per a two-tailed, two-samples t-test of clutter values for large and small real-world objects, large objects were on average more cluttered than small objects ($t(28) = 3.43$, $P = 0.002$, $d = 1.221$, 95% $\mathrm{CI}_d = (0.428, 1.995)$). However, when eliminating the top five most-cluttered large items (that is, grandfather clock, jukebox, phone booth, porta-potty, vending machine) and top five least-cluttered small items (that is, computer mouse, lighter, playing card, soda can, toothpaste tube) from analysis, leaving only the middle 20 objects, the groups no longer differed in average clutter ($t(18) = 0.80$, $P = 0.44$ (two-tailed), $d = 0.357$, 95% $\mathrm{CI}_d = (-0.532, 1.236)$), yet the size-based modulation of attentional allocation remained when accounting for only the middle 20 objects in experiment 1a ($F(1, 19) = 16.66$, $P < 0.001$, $\eta_p^2 = 0.467$, 90% $\mathrm{CI}\eta_p^2 = (0.169, 0.632)$) and experiment 1b ($F(1, 19) = 25.98$, $P < 0.001$, $\eta_p^2 = 0.578$, 90% $\mathrm{CI}\eta_p^2 = (0.288, 0.710)$). Furthermore, the degree of contrast clutter in the middle 20 objects did not significantly predict response time in a regression analysis in either experiment (experiment 1a: $R^2_{\mathrm{adj}} = -0.015$, $F(1,28) = 0.56$, $P = 0.459$, $\eta_p^2 = 0.020$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.159)$; experiment 1b: $R^2_{\mathrm{adj}} = -0.001$, $F(1,28) = 0.97$, $P = 0.333$, $\eta_p^2 = 0.033$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.188)$), despite the significant prediction of response time by object size rating. Finally, in experiment 3, size-based modulation of attentional shifting also remained when accounting for only the middle 20 objects, as the analysis revealed a significant main effect of size ($F(1, 19) = 31.61$, $P < 0.001$, $\eta_p^2 = 0.625$, 90% $\mathrm{CI}\eta_p^2 = (0.348, 0.743)$), a marginally significant interaction between size and validity ($F(1, 19) = 3.52$, $P = 0.076$, $\eta_p^2 = 0.156$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.378)$) and a significant three-way interaction between size, validity and target spatial frequency ($F(1, 19) = 5.19$, $P = 0.034$, $\eta_p^2 = 0.215$, 90% $\mathrm{CI}\eta_p^2 = (0.009, 0.433)$), indicating that both size and spatial frequency modulated attentional shifts. Crucially, the four-way interaction between size, validity, object orientation and frequency was not significant ($F(1, 19) = 0.10$, $P = 0.757$, $\eta_p^2 = 0.005$, 90% $\mathrm{CI}\eta_p^2 = (0.000, 0.135)$), failing to indicate that object orientation affects the critical three-way interaction between object size, cue validity and target frequency. Additional lower-level orientation interactions are included in the Clutter Control section of Supplementary Materials. Taken together, these results further suggest that the effect of size is not driven by low-level features or object orientation.

Possible mid- and high-level differences between small and large objects were investigated using a two-pronged approach (see Supplementary Experiment 3). First, in a separate Mechanical Turk survey, participants were asked to rate the overall curvature and familiarity of each object in the object set, as previous work has suggested that large and small objects may differ in both curvature (mid-level) and familiarity (high-level) due to the ergonomic features of small objects. Second, a steerable pyramid image decomposition analysis was performed to assess the angular features and relative curvilinearity of small and large objects. To preface the findings of Supplementary Experiment 3, no differences were observed across object groups in either the Mechanical Turk survey or the steerable pyramid analysis. These results suggest that size-based attentional scaling is not due to differences in shape features or familiarity between large and small objects.

Design and procedure. In experiment 1, cue validity (75%, 50%) was manipulated between subjects (Experiments 1a and 1b). For each experiment, a 2 × 2 × 2 within-subjects factorial design was employed, with inferred object size (large, small), cue-target relation (valid, invalid) and object orientation (vertical, horizontal) as within-subject factors. Response times 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 experiment 2, the procedure was identical to that used in experiment 1a. However, line drawing images were scrambled using MATLAB (Mathworks) to reduce the contribution of inferred object knowledge to attention. Two forms of scrambling were used. In Experiment 2a, images were scrambled using a retinotopic technique, in which each object was divided into 2,500 6 × 6 pixel grains (36 pixel^2 region). The 6 × 6 grains were then scrambled across all objects in a manner that preserved the average luminance of each 6 × 6 pixel grain location. This conservative method roughly preserved the boundaries of each object, but reduced the quality of each image. In experiment 2b, a more general scrambling technique was used, in which no pixel information was preserved, conserving no identifiable information about each object.

In Experiment 3, the procedure was also identical to that used in experiment 1a, with the exception that targets were Gabor patches rather than T/L renderings. The left-leaning Gabor target (−45°) was assigned to the 'c' keyboard response, while the right-leaning Gabor target (45°) was assigned to the 'm' keyboard response.

Each trial began with a 500 ms presentation of one object line drawing. Following the presentation of the line drawing, a red cue was presented on one of the two ends of the object, followed by a brief interval prior to the presentation of target stimuli. In experiments 1, 2 and 3, the cue lasted 100 ms, while the interstimulus interval (ISI) was 150 ms. Following the ISI, targets and distractors remained on the screen for 2 s or until response. Participants were encouraged to respond as quickly and as accurately as possible. Each participant in each experiment completed a practice block consisting of 30 trials, followed by eight experimental blocks, each consisting of 120 trials. Trial order was fully randomized across each experiment.

In the post-experiment survey, a single line drawing was presented centrally, with a six-point rating scale presented underneath (Fig. 1c). Participants were instructed to rate the size of each of the objects on a scale of one to six, one being very small and six being very large. The size ratings were used as an independent measure of each object's real-world size and participants were agnostic to the size classifications of the objects made prior to the experiments.

Data analyses. Data for each experiment were analysed using repeated-measures analysis of variance with an α set to 0.05. Two-sample and paired t-tests were also employed where indicated in the text and were two-tailed unless otherwise specified. Anticipatory response times faster than 100 ms as well as any response time longer than 1,500 ms were removed from analysis (1.9, 2.9, 1.1, 5.0 and 6.6% data removal for experiments 1a, 1b, 2a, 2b and 3, respectively. Mean response times for each condition (small, large object; valid, invalid location) were calculated for each subject. For each analysis, all assumptions, including normality and equality of variances, were formally tested and confirmed. Confidence intervals are reported for all effect sizes, including 90% CI of η_p^2 for F-tests (see Lakens for commentary on the use of 90% CI for F-tests with $\alpha = 0.05$) and 95% CI of Cohen's d for t-tests[32].

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. The custom PsychoPy and MATLAB programs generated for this study are available from the corresponding authors on reasonable request.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 20 March 2018; Accepted: 30 October 2018;
Published online: 7 January 2019

References

1. Kosslyn, S. M. Information representation in visual images. *Cogn. Psychol.* **7**, 341–370 (1975).
2. Konkle, T. & Oliva, A. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* **74**, 1114–1124 (2012).
3. Carrasco, M. Visual attention: the past 25 years. *Vision Res.* **51**, 1484–1525 (2011).
4. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
5. Posner, M. I., Snyder, C. R. R. & Davidson, B. J. Attention and the detection of signals. *J. Exp. Psychol. Gen.* **109**, 160–174 (1980).
6. Eriksen, C. W. & St. James, J. D. Visual attention within and around the field of focal attention: a zoom lens model. *Percept. Psychophys.* **40**, 225–240 (1986).
7. Duncan, J. Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.* **113**, 501–517 (1984).
8. Valdes-Sosa, M., Bobes, M. A., Rodriguez, V. & Pinilla, T. Switching attention without shifting the spotlight object-based attentional modulation of brain potentials. *J. Cogn. Neurosci.* **10**, 137–151 (1998).
9. O'Craven, K. M., Downing, P. E. & Kanwisher, N. fMRI evidence for objects as the units of attentional selection. *Nature* **401**, 584–587 (1999).
10. Egly, R., Driver, J. & Rafal, R. D. Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Gen.* **123**, 161–177 (1994).
11. Drummond, L. & Shomstein, S. The timecourse of space- and object-based attentional prioritization with varying degrees of certainty. *Front. Integr. Neurosci.* **7**, 1–10 (2013).
12. Muller, N. G. & Kleinschmidt, A. Dynamic interaction of object- and space-based attention in retinotopic visual areas. *J. Neurosci.* **23**, 9812–9816 (2003).
13. Shomstein, S. & Behrmann, M. Cortical systems mediating visual attention to both objects and spatial locations. *Proc. Natl Acad. Sci. USA* **103**, 11387–11392 (2006).
14. Malcolm, G. L., Rattinger, M. & Shomstein, S. Intrusive effects of semantic information on visual selective attention. *Attention, Perception, Psychophys.* **78**, 2066–2078 (2016).
15. Konkle, T. & Oliva, A. Canonical visual size for real-world objects. *J. Exp. Psychol. Hum. Percept. Perform.* **37**, 23–37 (2011).
16. Jefferies, L. N., Gmeindl, L. & Yantis, S. Attending to illusory differences in object size. *Attention, Perception, Psychophys.* **76**, 1393–1402 (2014).
17. Robertson, L. C. & Kim, M.-S. Effects of perceived space on spatial attention. *Psychol. Sci.* **10**, 76–79 (1999).
18. Murray, S. O., Boyaci, H. & Kersten, D. The representation of perceived angular size in human primary visual cortex. *Nat. Neurosci.* **9**, 429–434 (2006).
19. Konkle, T. & Oliva, A. Normative representation of objects: evidence for an ecological bias in perception and memory. In *Proc. 29th Annual Cognitive Science Society* (eds McNamara, D. S. & Trafton, J. G.) 407–412 (Cognitive Science Society, 2007).
20. Epstein, W. & Baratz, S. S. Relative size in isolation as a stimulus for relative perceived distance. *J. Exp. Psychol.* **67**, 507–513 (1964).
21. Bar, M. Visual objects in context. *Nat. Rev. Neurosci.* **5**, 617–629 (2004).
22. Long, B., Konkle, T., Cohen, M. A. & Alvarez, G. A. Mid-level perceptual features distinguish objects of different real-world sizes. *J. Exp. Psychol. Gen.* **145**, 95–109 (2016).
23. Konkle, T. & Oliva, A. A familiar-size Stroop effect: real-world size is an automatic property of object representation. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 561–569 (2012).
24. Malcolm, G. L. & Shomstein, S. Object-based attention in real-world scenes. *J. Exp. Psychol. Gen.* **144**, 257–263 (2015).
25. Biederman, I. & Cooper, E. E. Size invariance in visual object priming. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 121–133 (1992).
26. Kosslyn, S. M., Thompson, W. L., Kim, I. J. & Alpert, N. M. Topographical representations of mental images in primary visual cortex. *Nature* **378**, 783–785 (1995).
27. Eckstein M. P., Koehler K., Welbourne L. E. & Akbas E. Humans, but not deep neural networks, often miss giant targets in scenes. *Curr. Biol.* **27**, P2827–P2832 (2017).
28. Wolfe, J. M. Visual attention: size matters. *Curr. Biol.* **27**, R1002–R1003 (2017).
29. Castiello, U. & Umiltá, C. Size of the attentional focus and efficiency of processing. *Acta Psychol. (Amst.)* **73**, 195–209 (1990).
30. Martin, A. The representation of object concepts in the brain. *Annu. Rev. Psychol.* **58**, 25–45 (2007).
31. Chao, L. L. & Martin, A. Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* **12**, 478–484 (2000).
32. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Front. Psychol.* **4**, 1–12 (2013).
33. Sekuler, R. & Blake, R. *Perception* 3rd edn (McGraw-Hill Education, New York, 1994).
34. Peterson, M. A. in *Pioneer Visual Neuroscience: A Festschrift for Naomi Weisstein* (ed. Brown, J.) 151–166 (Taylor Francis, New York, 2018).
35. Galera, C., Von Grünau, M. & Panagopoulos, A. Automatic focusing of attention on object size and shape. *Psicologica* **26**, 147–160 (2005).
36. Herrmann, K., Montaser-Kouhsari, L., Carrasco, M. & Heeger, D. J. When size matters: attention affects performance by contrast or response gain. *Nat. Neurosci.* **13**, 1554–1559 (2010).
37. Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G. & Mishkin, M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends. Cogn. Sci.* **17**, 26–49 (2013).
38. Konkle, T. & Caramazza, A. Tripartite organization of the ventral stream by animacy and object size. *J. Neurosci.* **33**, 10235–10242 (2013).
39. Peirce, J. W. PsychoPy-Psychophysics software in Python. *J. Neurosci. Methods* **162**, 8–13 (2007).
40. Peirce, J. W. Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform.* **2**, 1–8 (2009).
41. Rosenholtz, R., Li, Y. & Nakano, L. Measuring visual clutter. *J. Vis.* **7**, 1–22 (2007).

Acknowledgements

This work was supported by a National Science Foundation grant no. BCS-1534823 to S.S. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. We thank G. Malcolm, D. Kravitz and M. Behrmann for insightful comments on an earlier version of this manuscript. Special thanks go to M. Peterson for suggesting a spatial frequency control experiment reported in experiment 3.

Author contributions

A.J.C. and S.S. contributed to theoretical motivation, developed the study design and wrote the paper. A.J.C. programmed and conducted the experiments, performed data collection and analysis. J.C.N. and P.S.S. assisted with theoretical motivation, programming experiments and data collection, and provided revisions on manuscript drafts.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-018-0485-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to S.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

natureresearch

Corresponding author(s): Andrew Collegio

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	All data collection was performed using custom code in PsychoPy 1.82 and MATLAB R2013b
Data analysis	All data contained in this manuscript were analyzed using Microsoft Excel 2010 and JASP version 0.8.5 or higher

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study was conducted using a quantitative experimental design.
Research sample	Research sample was powered on previous studies examining influence of task-irrelevant objects on attentional allocation. Sample sizes for these studies vary between 16-25 participants. 25 participants (8 male; average age: 19.4 years) were recruited for Experiment 1a. 25 participants (7 male; average age: 19.4 years) were recruited for Experiment 1b. 26 participants (6 male; average age: 19.2 years) were recruited for Experiment 2a. 26 participants (8 male; average age: 19.7 years) were recruited for Experiment 2b. 25 participants (9 male; average age 19.6 years) were recruited for Experiment 3.
Sampling strategy	Participants were randomly sampled from within the George Washington University undergraduate population. All participants were undergraduate students at George Washington University, gave written informed consent according to The George Washington University's institutional review board (IRB) and received research credit for participation. Sample was representative of the GWU student population.
Data collection	All data were collected on a 19" Dell 1908FP Color LCD monitor with a refresh rate of 60 Hz, positioned at a distance of 62 cm from the viewer. All experiments were generated and presented using PsychoPy v1.82. The researcher was present with the participant during the practice block but not during experimental blocks. The researcher was not blind to the experimental condition or study hypothesis.
Timing	Data collection began on September 14, 2016 and ended on December 6, 2017
Data exclusions	Participants were excluded from analysis based on an established exclusion criteria of 90% accuracy. Any participant performing below 90% accuracy was excluded. For experiment 1a, 5 participants were excluded. From Experiment 1b, 5 participants were excluded. From experiment 2a, 6 participants were excluded. From experiment 2b, 5 participants were excluded. For experiment 3, 5 participants were excluded.
Non-participation	No participants dropped out or declined participation in the study.
Randomization	Within subject experimental method was used. Each individual trail was randomized.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study	
☐	☐	Unique biological materials
☐	☐	Antibodies
☐	☐	Eukaryotic cell lines
☐	☐	Palaeontology
☐	☐	Animals and other organisms
☐	☒	Human research participants

Methods

n/a	Involved in the study	
☐	☐	ChIP-seq
☐	☐	Flow cytometry
☐	☐	MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	*Describe any restrictions on the availability of unique materials OR confirm that all unique materials used are readily available from the authors or from standard commercial sources (and specify these sources).*

Antibodies

Antibodies used

Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

State the source of each cell line used.

Authentication

Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

Mycoplasma contamination

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.

Wild animals

Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Human research participants

Policy information about studies involving human research participants

Population characteristics

See above

Recruitment

All participants were recruited from the George Washington University subject pool. Participants received research credit for partaking in the study. Participants were told they could leave or refuse participation during the study and still receive credit, and each gave written informed consent according to The George Washington University's institutional review board (IRB).

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session
(e.g. UCSC)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- [ ] The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- [ ] The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- [ ] All plots are contour plots with outliers or pseudocolor plots.
- [ ] A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument

Identify the instrument used for data collection, specifying make and model number.

Software

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

- [ ] Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)

Specify: functional, structural, diffusion, perfusion.

Field strength

Specify in Tesla

Sequence & imaging parameters

Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, Sequence & imaging parameters | *slice thickness, orientation and TE/TR/flip angle.*

Area of acquisition | *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software | *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization | *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template | *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal | *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring | *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings | *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested | *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016) | *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction | *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity | *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis | *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis | *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*