Where Is Size in the Brain of the Beholder?

Dietrich Samuel Schwarzkopf*
Experimental Psychology & Institute of Cognitive Neuroscience, UCL, London, UK
Received 29 October 2014; accepted 26 December 2014

Abstract
Despite advances in our understanding of how the brain represents visual space, it remains unresolved how the subjective experience of an object’s size arises. While responses in retinotopic cortex correlate with perceived size, this does not imply that those brain regions mediate perceived size differences. Here I describe how the percept of an object’s size could be generated in the brain and outline unanswered questions that future research should seek to address.

Keywords
Size perception, retinotopy, primary visual cortex (V1), illusion, contextual interaction

1. Introduction
Size matters. Leaving jokes and innuendo aside, an accurate judgment of the size of objects is a critical skill for survival. Sizing up another organism can make the difference between identifying prey and potential predators. Estimates of size and spatial distance are also crucial for seemingly simple actions, such as grasping and picking up a coffee cup or safely placing one’s feet in difficult terrain. In light of this, it is surprising that the neural processes underlying size perception remain only poorly understood.

Are there neurons tuned to the size of visual stimuli, as there are neurons selective for orientation and disparity (Hubel and Wiesel, 1962, 1968)? Neurons in higher visual areas with large receptive fields could be sensitive to the size or position of complex objects. There is some evidence for this (Afraz et al., 2010; Kravitz et al., 2010) but it seems unlikely that this neural substrate can mediate the excellent acuity for fine spatial judgments of the human visual system. Considering that these brain areas are presumably also involved with
resolving the challenge of recognizing objects, invariance to spatial characteristics like size and position would be economical because it disentangles the computation of an object’s identity from its spatial properties.

Instead an object’s size could be read out directly from its representation in the retinotopically-organized maps of visual space in early visual areas, especially V1, an idea related to the local sign hypothesis (Fischer et al., 2011; Whitney and Bressler, 2007). This refers to the idea that the response of neurons organized into a visual field map is used directly to encode the position of a stimulus in visual space. The tiny receptive fields of neurons in early visual cortex afford these maps with high spatial resolution, which makes them suitable for supporting even very subtle position judgments. In the same way the size of an object could also be determined by judging the position of its edges.

This mechanism however does not seem to account intuitively for several aspects of size perception. The apparent size of an object is subject to distortions by contextual illusions. A powerful example is the Moon, which can appear gigantic when it is low above the horizon compared to when it is high in the sky. And even the basic geometry of an image can induce such effects, as in the Ebbinghaus, the Müller-Lyer, or the Ponzo illusions as well as depth cues from perspective or retinal disparity (Fig. 1). Thus, if the stimulus representation in primary visual cortex (V1) is used to determine the size of an object, it should also reflect such illusory size differences.

2. Evidence That Perceived Size Is Encoded in V1

Interestingly, this view is now supported by a host of experimental findings. Neuroimaging experiments demonstrated that the activation produced by a stimulus in V1 reflect how large it appears to the observer (Fang et al., 2008; Murray et al., 2006). These experiments exploited the fact that the apparent size of a stimulus of constant retinal size depends on its assumed distance from the observer (e.g., Fig. 1D). Physically identical images were placed in the context of a corridor either at its near or far end. Shadows appropriate for the objects’ real-world size further facilitated the interpretation that the distant stimulus dwarfed the near one. When the stimulus was at the far end of the corridor and it was thus perceived as larger, its cortical representation in V1 was also larger in terms of which parts of the retinotopic map it activated. Further, this effect was related to the magnitude of the perceptual size difference measured behaviorally.

A subsequent study confirmed this effect also for Emmert’s Law, the finding that the perceived size of a retinal afterimage is larger when it is viewed on a distant surface than a near one (Sperandio et al., 2012a). Contrary to Richard Gregory’s tentative prediction (Gregory, 2008), the activation caused
by the afterimage in V1 is also modulated according to the perceived size. The finding that the perceived size of a retinal afterimage is subject to contextual size illusions further supports this hypothesis (Sperandio et al., 2012b). Similarly, V1 responses also accord with differences in perceived size induced by adaptation (Pooresmaeili et al., 2013). Taken together, such experiments suggest that distortions of visual space due to distance interpretations must occur in the brain as early as V1. Recently, electrophysiological recordings demonstrated that neuronal receptive fields in the V1 map are shifted relative to the physical location of a stimulus depending on whether it is interpreted as near or far (Ni et al., 2014). This provides the clearest evidence thus far that the representation of visual space itself is warped based on our interpretation of distance from the observer.

3. Relationship of Size Illusions with V1 Surface Area

Additional evidence for this idea also comes from experiments which compared the strength of common size illusions, the tunnel illusion similar to that employed in aforementioned studies and the Ebbinghaus illusion, to the surface area of V1 delineated by retinotopic mapping experiments (Schwarzkopf and Rees, 2013; Schwarzkopf et al., 2011). The rationale of these experiments was that if the cortical extent of the neural effects related to these illusions were spatially constant, as would be expected if lateral, horizontal axons or interneurons in V1 mediate these effects, the illusion strength should be reduced in individuals with larger V1 areas. Such an inverse relationship is precisely what these experiments found. In the Ebbinghaus illusion a larger V1 area increases the cortical distance between the inducing context and the target stimulus. This in turn changes the sign of the modulatory effect the context exerts on the perceived size of the target, which may relate to facilitatory and inhibitory receptive field effects (Schwarzkopf and Rees, 2013). Interestingly, the cortical distance where this change in the effect occurs was estimated to be close to the size of the cortical point image (Harvey and Dumoulin, 2011; Palmer et al., 2012). This indicates that this illusion might indeed relate to a constant factor in neuronal circuits operating along the cortical sheet in V1. Therefore, effects that largely depend on such V1 circuits should be weaker for individuals with a large V1 (Fig. 2). For instance, the effect of adaptation on perceived size (Pooresmaeili et al., 2013) is presumably also mediated by neurons in early visual cortex, which theoretically predicts weaker adaptation in individuals with a larger V1.

However, a correlation between V1 area and illusion strength does not necessarily imply that V1 neurons actually create the illusion. In particular for the tunnel illusion and the Ponzo illusion (which may be similar to the tunnel illusion, i.e., compare Fig. 1C and 1D) it seems improbable that the computations
necessary to infer the three-dimensional distance of the stimulus from perspective cues are performed in V1. Rather it seems more likely that feedback connections from higher visual areas modulate the stimulus representation there, perhaps by distorting the entire map of visual space through shifting of receptive fields as suggested by electrophysiological recordings. And even for simple geometric illusions like the Ebbinghaus it is quite possible that higher-level processes are at least one of the factors involved in creating the perceived size differences. Size contrast, in which the size of objects in the vicinity influences the perceived size of the target object, may be a higher-level process involved in that illusion, even if the illusion partly arises through low-level circuits in V1. Support for such higher-level processes comes from the fact that illusion strength depends on the shape similarity between inducers.
Figure 2. Interactions between the profile of contextual interactions between visual stimuli and the macroscopic surface area of V1. Each surface represents the V1 of an individual. The color code denotes the retinotopic organization (e.g., eccentricity). In an individual with a large V1 (top), the same amount of visual space is represented with a greater cortical territory than in an individual with a small V1 (bottom). The sombrero-shaped profile illustrates a putative center-surround profile of contextual interactions [e.g., the peak region may denote increases in perceived size while the troughs denote perceived shrinkage as suggested by Schwarzkopf and Rees (2013)]. If this effect is constant in terms of cortical space, for instance because it is mediated by intrinsic lateral connections in V1, in an individual with a small V1 the perceptual effect will extend farther in visual space (bottom left). Conversely, if the effect is calibrated to the retinotopic map, possibly because it is mediated by feedback connections from higher brain regions, the effect in visual space will be constant (bottom right) and therefore not correlated with V1 surface area. Yet for both kinds of effect the magnitude of the percept will be reflected by responses in V1 neurons.

and target and other presumably higher cognitive factors (Choplin and Medin, 1999; Deni and Brigner, 1997; Hughes and Fernandez-Duque, 2010; Jaeger and Guenzel, 2001; Jaeger et al., 2014; Muise et al., 1997). Moreover, even if V1 circuits mediate these illusions this would not preclude that additional factors, e.g., the neurotransmitter balance or synaptic plasticity, modulate their strength. Thus reports that illusion strengths differ in different cultures (Caparos et al., 2012; De Fockert et al., 2007, 2011; Doherty et al., 2008), with age (Doherty et al., 2010; Thelen and Watt, 2010), gender (Phillips et al., 2004), or in conditions like autism and schizophrenia (Chouinard et al., 2013;
Dakin and Frith, 2005; Yoon et al., 2010) might be explained not by morphological but functional differences between individuals. However, not all such findings may truly reflect the perceptual differences and should be replicated with modern psychophysical methods that are more robust to non-perceptual factors like decision bias (Jogan and Stocker, 2014; Morgan et al., 2012, 2013).

In general, even for low-level effects that are spatially constant in V1, as appears to be the case for our Ebbinghaus experiments, V1 circuits might not necessarily be involved directly. It is possible that feedback from higher areas induces the expansion or contraction of the target stimuli due to the target-inducer distance on the cortical sheet. As such it is possible that effects that depend on feedback connections fanning out over a region within V1 are calibrated to the individual’s cortical magnification factor in that region. This would imply that while perceived size is represented in V1, illusion magnitude does not correlate with V1 surface area (Fig. 2).

Several other pieces of evidence in fact support the notion that different processes mediate the Ebbinghaus and the Ponzo/tunnel illusion apart from the conjecture that the latter must involve processing of perspective cues. First, the strengths of the two illusions are not correlated (Schwarzkopf et al., 2011). Second, the Ebbinghaus is greatly reduced in strength when inducers and target are presented dichoptically to different eyes (Song et al., 2011), a hallmark test of a process mediated by early stages of the visual system where input from the two eyes is still segregated. In contrast, the Ponzo illusion manifests with similar strength regardless of whether it is presented dichoptically or with inducers and targets in the same eye. Third, only one experiment showed the correlation in the strength of the tunnel illusion with V1 area. In these data the correlation is strongly affected by one influential outlier (Schwarzkopf et al., 2011). Future experiments should therefore test if this effect replicates and extend the study to other size illusions (Fig. 1) to ascertain which effects are likely to be low-level processes.

Naturally, the presence or absence of a correlation between illusion strength and V1 surface area is not alone sufficient to confirm the involvement of feedback mechanisms. Functional or effective connectivity analyses could provide converging evidence of the circuits involved in generating these illusions, as in a recent study that employed dynamic causal modeling to implicate intrinsic connections in V1 as mediating the tilt illusion (Song et al., 2013). Behavioral estimates of the velocity of a perceptual effect may further help reveal the underlying circuits (Cass and Spehar, 2005; Genç et al., 2013).

4. The Role of Fronto-Parietal Areas

Object size could also be at least partially encoded by higher visual areas and prefrontal decision making regions even if the retinotopic representation in
early visual cortex is used to read out size. Recent mapping studies identified a region in parietal cortex containing what the authors suggest is a topographic map of numerosity (Harvey et al., 2013). While the authors used a range of stimuli to control for potential low-level confounds to this interpretation, such as that this region may be selective for stimulus size, it remains possible that tuning to size explains at least some of these effects. Spatially overlapping neuronal populations in the same area could also encode different stimulus attributes. Alternatively, there may be other brain regions that explicitly encode object size, possibly by incorporating prior expectations on the real-world size of objects (Konen and Kastner, 2008; Konkle and Oliva, 2012). In this context, it is also remains unclear in how far size perception differs between stimuli that are encoded by specific neuronal populations, such as faces, and novel objects. In order to localize a potential brain area involved in size judgments, future neuroimaging experiments should disentangle judgments of relative size from physical differences. Is there any brain region in parietal or frontal cortex whose activation tracks the magnitude of perceived size difference between two stimuli? In addition, learning studies could test how familiarity with a stimulus affects size judgments.

The causal role of feedback in size illusions could also be tested using transcranial magnetic stimulation (TMS). Following an experimental design used previously to infer the role of feedback from lateral occipital cortex to V1 in the perception of illusory contours (Wokke et al., 2013), TMS could be applied to V1 either at an early or a late phase after stimulus onset. If the illusory percept is induced by feedback from higher regions only late TMS should interfere with it. Conversely if the effect arises in V1 itself early TMS should interfere with it. If it is possible to identify a higher-level region involved in size judgments, this experiment could further be extended to include TMS over that region. The prediction of a feedback account would then be that only early TMS over this higher-level region but only late TMS over V1 disrupts the illusion, while for assumed low-level illusions the pattern should be reversed.

5. Going Beyond Illusory Size

If size judgments indeed rely on reading out retinotopic representations from V1, this could in fact explain other misperceptions of stimulus size beyond such contextual illusions (Anstis, 1998). Helmholtz already reported perceptual distortions that are consistent with the cortical magnification profile of the early visual system (Helmholtz, 1867). Thus even subtle errors in compensating for this warped representation could result in misperception of the size and shape of stimuli in the periphery. In this context it is interesting that variability in basic size judgments occurs across the visual field (Afraz et al.,
2010; Schwarzkopf and Rees, 2013). Could such spatial heterogeneity in size perception and apparent geometric distortions be related to idiosyncrasies in cortical magnification in V1? While previous experiments observed no correlation between the perceptual biases in size judgments in each visual hemifield and the contralateral V1 surface areas (Schwarzkopf and Rees, 2013), in those experiments biases were measured by asking observers to compare stimuli between the two hemifields so the estimated bias would have always contained distortions potentially related to those locations. Moreover, V1 surface area was measured through standard phase-encoded retinotopic mapping (Sereno et al., 1995). Modern advances in methodology permit the measurement of population receptive field (pRF) parameters, which not only control biases in traditional methods but also provide estimates of the spatial selectivity, local cortical magnification factor, and possibly even of contextual inhibitory interactions in visual cortex (Dumoulin and Wandell, 2008; Harvey and Dumoulin, 2011; Zuiderbaan et al., 2012). Such measures may thus be more suitable to address questions about the neural representation of object size.

6. Implications for Shape Perception and Grasping Action

However, the idea that size and shape are read out from retinotopic representations seems inconsistent with claims that while perceived size is susceptible to contextual illusions, grasping behavior accurately reflects physical size. Such findings support the dual-stream hypothesis by which perceived stimulus size is processed in the ventral visual cortex, while dorsal regions controlling motor actions are unaffected by perceptual bias (Goodale, 2011; Goodale and Milner, 1992; Milner and Goodale, 2008). This idea however remains controversial (Franz and Gegenfurtner, 2008; Franz et al., 2005). Future research must seek to establish in how far motor behavior is truly unaffected by perceived size. It is of course also possible that there are two overlapping neuronal populations in V1, only one of which reflects perceived size but the other encodes the physical size. The dorsal stream might selectively receive input from the latter population. If parietal areas indeed track perceived size but read it out from position signals early visual cortex, this would suggest that size represents a hybrid attribute that does not truly fit the dorsal vs. ventral stream dichotomy. However, if there is indeed a representation of perceived size in ventral and veridical size in dorsal regions, it could also be the case that there are distinct areas in either stream that encode object size.

The V1-read-out hypothesis is also at odds with psychophysical experiments suggesting that the position of stimulus elements is actually of surprisingly little importance for perceiving the shape or curvature of a contour. Figure 3 shows images of small grating patches that have been positioned on a perfectly circular path (Fig. 3A). However, due to the orientation of each
Figure 3. Orientation exerts a strong influence on perceived location. The grating patches in all three panels are positioned on perfect circles. However, the orientation of each patch affects whether one perceives a circle (A), a square (B), or a diamond (C).

local patch we perceive these stimuli as either square (Fig. 3B) or diamond shaped (Day and Loffler, 2009; Loffler et al., 2003). The position of individual elements dominates shape perception to a lesser degree than orientation (Wang and Hess, 2005). If size is indeed read out based on which locations in retinotopic cortex are activated by a stimulus, geometric shape and size could be closely related. This poses the question whether the retinotopic activation by these stimuli in V1 also reflects their perceived position in the square, diamond or circle configuration. Such a finding could also be consistent with the observation that the V1 response to a drifting grating stimulus is shifted along the cortical sheet in the direction of motion (Jancke et al., 2004; Onat et al., 2011), which seems to accord with perceived displacements of the moving stimulus (De Valois and De Valois, 1991; Nishida and Johnston, 1999; Ramachandran and Anstis, 1990; Snowden, 1998). It however remains unclear how to reconcile this with human neuroimaging experiments that suggested responses in visual cortex are displaced along the trailing, not the leading, edge of a drifting grating stimulus (Whitney and Bressler, 2007). Naturally, it is also possible that inferences about shape are simply computed by neurons in higher visual areas that rely mostly on orientation information rather than position.

7. Conclusion

Evidence is mounting that the perceived size of visual stimuli is reflected by neural responses in V1 and early visual cortex. Future experiments should seek to localize which higher brain regions track relative size judgments, resolve the role of feedback to V1 in modulating perceived size, and explore the link between perceived position, geometric shape, and size.
References


