The mechanisms of size constancy

Irene Sperandio¹, Philippe A. Chouinard²

1) School of Psychology, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom.

2) School of Psychological Science and Public Health, La Trobe University, Bendigo Campus, Bendigo, Victoria, 3550, Australia.

Corresponding Author:

Correspondence should be addressed to Irene Sperandio or Philippe A. Chouinard at the above address. E-mail: i.sperandio@uea.ac.uk or p.chouinard@latrobe.edu.au
Abstract

Size constancy is the result of cognitive scaling operations that enable us to perceive an object as having the same size when presented at different viewing distances. In this article, we review the literature on size and distance perception to form an overarching synthesis of how the brain might combine retinal images and distance cues of retinal and extra-retinal origin to produce a perceptual visual experience of a world where objects have a constant size. A convergence of evidence from visual psychophysics, neurophysiology, neuropsychology, electrophysiology and neuroimaging highlight the primary visual cortex (V1) as an important node in mediating size-distance scaling. It is now evident that this brain area is involved in the integration of multiple signals for the purposes of size perception and does much more than fulfil the role of an entry position in a series of hierarchical cortical events. We also discuss how information from other sensory modalities can also contribute to size-distance scaling and shape our perceptual visual experience.

Keywords

Size perception, size-distance scaling, visual illusions, Emmert’s law, extra-retinal signals, primary visual cortex
Our visual experience is not a mere Cartesian replication of the external world. Rather, it is likely the product of a combination of bottom-up processing of retinal input, top-down processing driven by our expectations and biases, and horizontal processing about contextual information from various sensory inputs. These principles set the basis for perceptual constancy, which is the ability to perceive objects as having the same physical properties despite changes in retinal input. For example, we perceive an apple hanging from a tree as having the same red colour irrespective of variations in the spectral content of sunlight throughout the day or changes in weather conditions. And, we perceive an elephant as being physically large irrespective of whether it is close by and creates a large image on the retina or far away and creates a small image on the retina. In this review, we focus on how the brain might achieve the latter example of perceptual stability, known as size constancy. We start by describing the basic principles of how different parts of the retina are processed by different parts of the thalamus and cortex. We then present evidence from behavioural studies demonstrating how information about distance from retinal as well as extra-retinal origins can shape our perceptual visual experience of object size. We then review how certain types of brain lesions in the monkey and brain damage in patients can affect size perception, depth perception, and consequently size-distance scaling. We then summarise electrophysiological findings and recent functional magnetic resonance imaging (fMRI) results showing how the primary visual cortex (V1) can respond in a manner that reflects the perceived output and how these responses are modulated by retinal as well as extra-retinal signals about distance.

**Retinal representations in the brain**

Retinotopy is the mapping of visual input from the retina to neurons in the brain. The basic principles of retinal organisation will be reviewed briefly (for a more detailed account on cortical structures, see Engel et al. 1997; Wandell and Winawer 2011; for a more detailed account on subcortical structures, see Kastner et al. 2006). To begin, it is worthwhile to consider the similarities between the eye and camera. In the case of the camera, light passes through an aperture and gets refracted by a lens such that an inverted image (i.e. an image that is rotated 180 degrees) is projected and focused on a roll of film with colour and perspective preserved. In the case of the eye, light passes through the pupil and gets refracted through the cornea, enters the pupil, and gets refracted once again by the crystalline lens such that a focused and inverted image is projected in a similar manner on the retina. From there, the similarities between the eye and camera end.

There is a division of labour between the fovea and peripheral parts of the retina. The former is specialised for spatial resolution (i.e. the fovea is specialised to discriminate between two points that are close together in space) and the latter is specialised for sensitivity (i.e. the peripheral parts of the retina are specialised to detect small quanta of light). The transduced signal from the retina leaves the eye through the optic nerve and projects mostly to the V1 via the lateral geniculate nucleus (LGN). Given that the computational demands for discriminating between two points in space are high, a greater amount of brain tissue is devoted to processing information from the fovea. This differential representation in the brain between the fovea and peripheral parts of retina is known as magnification. The fovea is represented by large amounts of tissue in the posterior regions of the visual thalamus and visual cortex, while the more peripheral parts of the retina are progressively represented in more anterior regions of these same brain areas by smaller amounts of tissue. Another important feature of retinal organisation is that the topographical representation of the inverted image on the retina is maintained throughout the visual system. The upper visual field is represented by inferior brain regions, while the lower visual field is
represented by superior brain regions. In addition, the left visual field is represented by the right hemisphere while the right visual field is represented by the left hemisphere.

Currently, retinotopic mapping of the human brain is typically done with fMRI using a variant of the phase-encoding approach developed by Sereno et al. (1995). The approach is based on the principle that when a stimulus is presented in a cyclical manner, the fMRI blood-oxygen-level dependent (BOLD) signal will follow a similar cyclical profile (Fig. 1a). Typically, a movie of a stimulus moving across different parts of the visual field is played to the participant in a repeating loop. The BOLD signal is then analysed using a Fourier analysis to determine where in the visual field the response is the highest. Two types of visual stimulation are usually presented: one for polar angle mapping and the other for eccentricity mapping. For polar angle mapping, many researchers will present a flickering checkerboard wedge that rotates like the arms of a clock in either clockwise or counter-clockwise directions (Fig. 1b). The results from this type of stimulation typically give rise to the retinotopic map in Fig. 1d (see also Engel et al. 1997; Wandell and Winawer 2011). Note how the left and right visual fields are represented contralaterally in the right and left hemisphere respectively. For eccentricity mapping, many researchers present a flickering checkerboard ring that expands outwards or contracts inwards (Fig. 1c). The results from this type of stimulation typically give rise to the retinotopic map in Fig 1e (see also Engel et al. 1997; Wandell and Winawer 2011). Note how in this case the fovea is represented more posteriorly in the brain and occupies a greater extent of cortical tissue than the more peripheral parts of the retina. Both polar angle and eccentricity mapping procedures can be carried out under passive viewing conditions or while the participant performs an attentional task, such as indicating by button press a physical perturbation in the stimulus.

**Size constancy and optical illusions**

The brain recreates a conscious visual experience of the world that is stable. This is a remarkable feat if one considers that the retinal image is in constant flux as we and the objects around us are in movement. For example, as we watch a train depart from a platform at a railway station, its retinal image decreases in size while our perception of its size remains the same. The change in retinal image size obeys Euclid’s law (Fig. 2a). Namely, when the distance between an object and our eyes doubles then the retinal image size for that object decreases by half. Although the size of the retinal image decreases, we do not perceive the object as smaller but farther away. This perceptual re-scaling of size is known as size constancy (Andrews, 1964; Gregory, 1998; Morgan, 1992).

The effects of distance on size perception are well illustrated by afterimages (Fig. 2c). An afterimage is an image impressed on the retina of a stimulus that is no longer present in the environment. It can be induced from staring at a light source, or a picture, for a prolonged period of time. Appearing like a ghost or a shadow, an afterimage can be projected and perceived after a delay, upon a surface such as a wall or a screen. Afterimages are caused by the retina’s photoreceptors adapting to the stimulus. Although the surface area of the bleaching on the retina remains the same following exposure of the inducing stimulus, the size of the afterimage will decrease when we look at a surface that is closer to our eyes and increase when we look at surface that is further away. In fact, the perceived size of the afterimage can be mathematically calculated from Emmert’s law (1881) as:
where \( s \) is the perceived size of the afterimage, \( d \) is the perceived distance between the observer and the surface upon which the afterimage is “projected”, and \( \theta \) is the visual angle that is subtended by the afterimage. According to this law, the perceived size of the afterimage changes in a manner that is proportional to its distance. It then follows that our perception of size and distance are dependent on each other. Therefore, any misperception of distance will involve a misperception of size and vice versa. Studies in which participants were asked to project afterimages on optical illusionary displays illustrate this perceptual co-dependency. For example, it has been shown that the size of an afterimage changes as an observer projects it onto different locations in an Ames room (Dwyer et al., 1990) or different Ebbinghaus backgrounds (Fig. 3a; Sperandio et al., 2012b). In both cases, the contextual cues in the illusionary backgrounds inform the brain about the apparent (but not real) differences in distance, which leads to errors in judgements of size.

Important differences exist between perceiving depth in the real world and a two-dimensional picture, such as a painting. Any two-dimensional picture that produces an impression of depth is an optical illusion. There is no depth in a flat surface. It is only through the interpretation of cues depicted in the image that the brain recreates the impression of a third dimension that does not exist. The cues that painters apply to produce impressions of depth (e.g. shading, aerial perspective, occlusion, relative height, texture gradient) are ultimately in conflict with other cues (e.g. vergence, accommodation, and stereopsis) informing the brain that the eyes are focusing on a flat surface. It is perhaps for this reason why the perceptual scaling of the size of afterimages is typically near-perfect with distance while the apparent depth inferred from a two-dimensional picture has a less pronounced effect on the size perception of objects in the scene (Gregory, 2008).

How does one define an illusion? An illusion is a perceptual departure from reality. The most convincing illusions are those where the discrepancy between what is perceived and real is not recognised until one takes physical measurements and even then the illusion remains cognitively impenetrable (Pylyshyn, 1999). The latter refers to instances when the perceptual deception is not diminished after the observer is aware of the deception. There are many explanations for optical illusions (e.g. Chouinard et al. 2013). Some vision scientists favour a Bayesian account to explain a number of them (Geisler and Kersten 2002). Under this model, the brain is a sensorimotor machine that derives a percept from computing an optimal combination between the existing sensory evidence (e.g. what is captured by the retina) and knowledge acquired from prior experience for the purposes of motor interactions with our surroundings (Kording and Wolpert 2004). If this notion is correct, which we think is a fair assumption, then perception is highly dependent on the rules that the brain has learned about the world (Helmholtz 1867). One of these rules would be that the physical properties of an object are always constant. Namely, their size, shape, colour, and location are always the same regardless of changes in retinal information that arise from changes in viewing distance, angle of perspective, and lighting conditions.

An influential explanation for at least some optical illusions is Gregory's theory of inappropriate constancy scaling (1963, 1968). Gregory's account considers the mechanism responsible for the illusions to be an inappropriate perceptual re-scaling of object size that is triggered by misleading depth information. In other words, some visual illusions are experienced because we tend to keep constant the size of objects. For example, in the Ponzo illusion (Fig. 3b) the converging lines in the background
indicates to the brain that the upper line is farther away in the scene than the bottom line. Ordinarily, according to Euclid’s law, objects of the same size will have a smaller retinal image size when they are further away. Because of the perspective cues, the brain interprets the upper line as being further away and therefore perceptually rescales this line as longer. After all, under normal circumstances, size-distance scaling is applied to counteract variations in retinal image size due to changes in viewing distance, so that we can perceive an object as having a constant size. The end result is that the two lines appear to be different even though they are identical.

Variants of the inappropriate constancy scaling theory have been used to explain what is perhaps the oldest of all known optical illusions: the moon illusion (Fig. 4; Ross and Plug, 2002; McCready, 2004; Kaufman and Kaufman, 2000; Kaufman and Rock, 1962; Rock and Kaufman, 1962; Gregory, 2008). The moon appears progressively smaller as it rises from the horizon. Yet, the moon obviously does not change size and will cast the same retinal image size regardless of its celestial position. So why does the moon appear smallest in its zenith position? One explanation is that our visual system has not evolved to evaluate huge distances such as those separating the earth and the moon, and hence the moon is calculated as being closer than it actually is. Consequently, the moon overhead is interpreted as being closer to the observer than the moon on the horizon for the purposes of size constancy. When the distance of the moon is erroneously interpreted as closer, such as when it is high in the sky, the brain perceptually rescales its size as smaller. But the end result is paradoxically in opposition to size constancy laws: for most people the perceptually larger moon in the horizon actually looks closer than the moon in its zenith position. Hence, Emmert’s law is violated: the moon appears larger and so nearer (Gregory, 2008). If this account of the moon illusion is correct then this would mean that distance can be calculated in multiple ways.

Contributions of extra-retinal cues and other modalities to size constancy

Information about distance does not always arise from the retina. As people focus their gaze on objects at different distances, the crystalline lens changes its curvature such that the image on the retina becomes focused to accommodate for changes in distance. This process is called accommodation. The eyes will also converge when people fixate their gaze on objects close to them and diverge when they fixate their gaze on objects further away. This process is called vergence. Both cases provide extra-retinal signals to the brain about depth. In this section, we review studies showing how these signals contribute to size-distance scaling. These cues are not only powerful but they are also essential when viewing conditions are reduced or restricted. We then review how other types of extra-retinal signals can also provide distance information for the purposes of size constancy.

In a classic study, Holway and Boring (1941) had participants judge the size of circles placed at different distances from their eyes (Fig. 5). The circles always subtended the same visual angle. Under normal viewing conditions, participants perceived the size of the circles based on their physical size. As viewing distance increased, Holway and Boring placed larger circles in order to maintain the 1 degree in visual angle presentation and the participants correctly judged their size as bigger. Holway and Boring then had the participants repeat similar procedures under monocular viewing, which removed binocular cues, and finally had the participants repeat the procedures again under monocular viewing with the use of a peephole, which reduces accommodation. As more sources of depth information were eliminated, the participants began to perceive all the circles as having the same size even though they were physically
different. These experiments illustrate that size perception reflects the actual sizes of objects when there is plenty of depth information available but not when this information is reduced or eliminated. Similar conclusions were obtained by Sperandio et al (2009), who manipulated size and distance of the stimuli as in Holway and Boring's study (1941) but used simple reaction times as an index of perceived object size, the notion being that people are faster to react to a stimulus when it is perceived as bigger. As expected, participants responded faster to larger stimuli as distance increased under normal viewing conditions but not during conditions in which vision was restricted.

The extra-retinal signals that we reviewed so far have intimate relationships with the perceptual system in the brain. The perceptual system needs to be updated quickly whenever there are changes in accommodation or eye positioning in order to recreate a perceptual visual experience that is stable from retinal inputs that are transient and unreliable. Otherwise, the world would appear completely distorted. Whether this updating arises from proprioceptive signals from the eye to the perceptual system (i.e. a feedback account; Helmholtz, 1867) or from the motor centres in the brain relaying a "carbon copy" to the perceptual system as it sends a motor command to the eyes (i.e. a feedforward account; Sherrington, 1918) has been a matter of debate for nearly a century. The discussion we wish to consider next is how less direct forms of extra-retinal signalling from other sensory modalities can influence the visual experience of object size.

Auditory cues provide useful information about distance. For example, the distance of lightning can be calculated by the listener based on the time interval from when it is seen to when it is heard. When seeing and hearing the lightning is almost simultaneous, we know it is close by. In a similar manner, Jaekl and colleagues (2012) manipulated the perceived distance of a visual stimulus presented briefly in complete darkness by varying the delay between its presentation and the presentation of an auditory cue. The perceived size of the visual stimulus increased as the delay between the two stimuli increased, demonstrating that distance information from the auditory domain can influence size-distance scaling and our perceptual visual experience of size.

The Taylor illusion (Fig. 6) provides a demonstration of how the visual perception of size can be influenced by proprioception (Taylor, 1941; Gregory et al., 1959; Carey and Allan, 1996). In this illusion, an afterimage is induced by the presentation of a quick flash of light that bleaches the retina while the participant maintains their gaze on their hand in complete darkness. Several seconds later, an afterimage is conjured where their hand is located in space. The perceived size of the afterimage can then be altered by having the participant move their hand towards their eyes, in which case the afterimage decreases in size, or away from their eyes, in which case the afterimage increases in size. In a series of behavioural and eye-tracking experiments, we recently tested the contribution of both vergence and proprioceptive signals in driving this illusion (Sperandio et al., 2013a). In agreement with earlier work (Taylor, 1941; Mon-Williams et al., 1997), we showed that vergence contributed strongly to this illusion – underscoring its importance in providing distance information. However, in one of our experiments, we also introduced a conflict between vergence and proprioceptive signalling whereby we had participants fixate their gaze on a light that moved along one sagittal direction and move their arm in the opposite direction. Specifically, they converged their eyes while they moved their hand away from their body and diverged their eyes while they moved their hand closer to their body. The illusory effects diminished, indicating that proprioceptive feedback from the arm also affects the visual experience of size.
Likewise, Combe and Wexler (2010) have shown that size-distance scaling under reduced viewing conditions improves when additional distance information is provided by self-motion cues. In this study, participants performed a size constancy task in which they judged the size of stimuli that subtended the same visual angle at different viewing distances. Viewing conditions were reduced by means of presenting the stimuli briefly for only 16 ms. Viewing distance was manipulated either by having participants move their head back and forth from the computer monitor presenting the transient stimuli (i.e. active displacement of the observer) or by displacing the monitor. From this experiment, the authors could show how changes in viewing distance produced by the observer improved size-distance scaling towards near-perfect levels.

In other studies, it has been demonstrated that vestibular and / or proprioceptive signals about unusual body positions can interfere with visual size-distance judgments. For example, Higashiyama and Adachi (2006) had participants judge the size and distance of objects while they sat on a chair, place their head between their legs, and saw their surroundings upside down. Under this viewing condition, size constancy broke down and participants underestimated the size and distance of objects. The results were not attributed to an inversion of the retinal image given that a similar experiment having participants stand upright and see their environment inverted while wearing prism goggles did not affect size constancy. In a different study, Harris and Mander (2014) demonstrated how changes in perceived distance in tumbled and tumbling room facilities can alter size perception. These realistically-decorated rooms can be altered experimentally in ways that allows an investigator to systematically vary body orientation relative to the orientation of the room and consequently alter the participant’s perceived direction of gravity. In particular, the Tumbling Room can pivot around the participant strapped in a chair that can also be rotated, which allows the investigator to produce mismatches between visual, proprioceptive, and vestibular signals. Instances of a mismatch between the signals causes a change in the perceived direction of gravity and also causes the walls of the room to appear further away. When Harris and Mander (2014) made the walls appear further away by having participants seated orthogonally to the perceived direction of gravity, the participants judged the length of lines presented on the wall by means of a projector as being longer than their actual length. Thus, body orientation relative to the perceived direction of gravity interferes with a person’s ability to make accurate size-distance judgments.

Our consciousness of bodily self in relation to the external world can also alter how we perceive the size of objects. Consider an often cited example taken from Alice's Adventures in Wonderland (Carroll, 1865). Alice discovers a bottle on a table labelled *DRINK ME*, the contents of which causes her to shrink and perceive objects in the room as bigger, and a cake labelled *EAT ME*, which causes her to grow and perceive objects in the room as smaller. Recent work has shown that this world of physical distortions can be perceptually recreated by modifying our conscious state of bodily self in a body-swap illusion (van der Hoort *et al.*, 2011; van der Hoort and Ehrsson, 2014). In a body-swap illusion, as the name implies, a person perceptually experiences ownership of another body (Petkova and Ehrsson, 2008; Serino *et al.* 2013). This illusion can be induced by applying tactile stimulation on a person's body while they watch simultaneously by means of virtual reality the same tactile stimulation being applied to a mannequin. After stimulation, the conscious state of ownership transfers to the external body. In one study, van der Hoort and colleagues (2011) had participants experience the ownership of either a toy doll (size = 30 cm) or a giant mannequin (size = 400 cm) and judge the size and distance of objects placed in front of them. Like in the fictional Alice in Wonderland tale, the sense of owning a smaller or bigger body resulted in a change in the perception of object sizes and their viewing distances from the eyes.
Insights about size constancy from lesion studies

Studies that have investigated the effects of brain lesions on size constancy mechanisms in both monkeys and humans have highlighted the importance of the ventral stream of visual processing. Humphrey and Weiskrantz (1969) as well as Ungerleider et al. (1977) independently examined how abilities in size constancy are affected in rhesus monkeys with bilateral lesions to different brain regions. In both studies, monkeys were presented with two stimuli (circles that were projected onto screens in the first study and ‘real’ squares in the second study) placed at different distances such that the larger of the two could have a smaller or a larger retinal image size. The monkey’s task was to indicate which of the two stimuli was physically larger. Prior to surgery, the monkeys in both studies were able to learn the task. In the first study, Humphrey and Weiskrantz (1969) revealed that removing the inferior temporal cortex but not the posterior parietal cortex affected abilities in size constancy; monkeys with the former type of lesion could not perform or relearn the task. An analysis of the types of mistakes that were made suggested that the monkeys made perceptual decisions based on either the retinal image size of the stimuli or the distance with which the stimuli were presented. These findings are important as it suggests that the monkeys could still compute retinal image size and distance but could not integrate the two types of information.

In the second study, Ungerleider et al. (1977) revealed that removing the inferior temporal cortex or the extrastriate cortex (of both the ventral and dorsal streams) affected abilities in size constancy while removing the pulvinar nucleus of the thalamus did not. Interestingly, an analysis of the types of mistakes that were made suggested that the monkeys could still compute retinal image size but not distance information following lesions to the extra-striate cortex. On the other hand, an analysis of the types of mistakes that were made after lesions to the inferior temporal cortex did not reveal anything conclusive as to the type of deficit that could explain the impairments in size constancy. Another difference between the two studies is that all the brain-lesioned monkeys in the Ungerleider et al. (1977) study could eventually relearn the task, whereas those monkeys with lesions to the inferior temporal cortex in the Humphrey and Weiskrantz (1969) could not. Although the reasons for this discrepancy are unclear, Ungerleider et al. (1977) suggest, on the basis that their brain-lesioned monkeys could re-learn the task, that size constancy could be retrained using compensatory mechanisms that work independently from those typically requiring the inferior temporal cortex.

Case studies of brain-damaged patients paint a similar picture. Frassinetti et al (1999) examined a patient with impaired visual size perception following a right occipital stroke. MRI examination revealed that the lesion resulting from this stroke appeared confined to the lateral and ventral portions of the extrastriate cortex in the right hemisphere (based on their figure it looked to us like it encompassed the lateral occipital complex). Her perceptual deficits consisted of a reduction of the apparent size of objects presented in the left but not the right hemifield, which supports the notion that the extrastriate cortex is important for visual size perception if one considers that retinal information from a hemifield is processed by the contralateral hemisphere. Previous work in patient DF, a well-studied patient with visual form agnosia resulting from bilateral damage to the posterior portions of the ventral stream, leads to similar conclusions. In one study, Carey et al. (1998) investigated patient DF’s abilities to perceive depth and to grasp objects placed at various viewing distances from her eyes. In agreement with Goodale and Milner’s two-stream hypothesis (Goodale et al., 1991; Goodale and Milner, 1992; Milner and Goodale 2008), patient DF’s grasping movements were always metrically accurate and indistinguishable from the control subjects but her depth perception was worse relative to the controls. Namely, she underestimated distance
by 30% under binocular viewing conditions and by 70% under monocular viewing conditions. This result also fits well with the Ungerleider et al.'s (1977) findings in the monkey; showing a deterioration in size constancy mechanisms due to deficits in judging distance.

Another study performed by Servos (2006) shows a decline in size constancy in patient DF. In this study, Servos (2006) had patient DF and a control subject judge the distance and the width of objects presented at the same visual angle (i.e. the visual stimulus always had the same retinal image size). Patient DF had great difficulty judging both viewing distance and object size, and her performance was worse on these tasks compared to the control subject. In a different experiment, Servos (2006) asked DF and the control subject to judge the size of afterimages as opposed to real objects. Interestingly, this experiment revealed that patient DF’s size perception of the afterimages was significantly better than her size perception for the real objects, although this was still worse than the control subject.

Recently, we had a cortically blind patient with Riddoch syndrome, known as patient MC (Dutton 2003), judge the size of objects at different viewing distances (Sperandio et al., 2012c; 2013b). Riddoch syndrome is a condition that is sometimes manifested in cortically-blind patients and is characterised by residual sight for movement in which the patient can still consciously perceive objects if they move or an object moves around them (Riddoch, 1917; Zeki and ffytche, 1998). Structural MRI performed on patient MC revealed widespread damage encompassing both the occipital lobes, the posterior portions of the temporal lobes bilaterally, as well as the right posterior parietal lobe. As a result of this damage, patient MC’s estimates of object size were poorly scaled to the physical size of the stimuli but were correlated instead with their retinal image size. In a separate experiment, we repeated similar procedures but had patient MC reach and grasp for the objects with her dominant (right) hand as opposed to having her estimate their size. Remarkably, the patient had no difficulty performing the necessary size-distance scaling computations to configure her grip according to the physical size of the objects, despite the fact she could not perceptually discern their physical size correctly. These findings suggest that there are multiple size-distance scaling mechanisms in the brain, each operating independently from one another for different purposes.

Taken together, the neuropsychological evidence seems to indicate that the ventral stream is necessary for the conscious visual perception of depth as well as object size. It also shows how certain sectors of the visual system, namely the dorsal stream of visual processing, can still process visual features such as object size for the online visuomotor control of movements independently of perceptual awareness.

On the basis of neuropsychological evidence, Naccache (2005) argues the following principles govern all forms of visual phenomenal consciousness. First, some sectors of the visual system never participate in conscious visual perception as shown by blindsight (Weiskrantz, 1996) and by patients DF and MC’s abilities in object grasping. Second, a visual stimulus is consciously perceived only if that stimulus has been coded by the ventral stream, which is supported by the fact that many forms of visual agnosia are the result of damage to the ventral but not the dorsal stream of visual processing (for review, see Farah 2004). This is not to say that all processing in the ventral stream leads to visual awareness but rather that visual awareness is dependent on ventral stream processing, which leads to the next principle. Third, the ventral stream of visual processing is necessary but not sufficient for conscious visual perception and that certain levels of attention are required to amplify the visual signal enough for visual
awareness to occur (Dehaene et al., 2001; Sergent et al., 2005). Evidence for this latter principle is seen in inattentional blindness during distracted driving. Studies simulating the use of cell phones while driving have shown that the dual task can cause people to fail to recognise important traffic signs, such as a stop sign or a red light, even though these stimuli are in plain sight of them and projected onto their retina (Drews et al., 2009; Regan and Strayer, 2014). Given the evidence, we believe that these general principles governing visual phenomenal consciousness are difficult to refute. We will argue in this paper for yet a fourth principle. Namely, V1 is an important node for the integration of multiple signals for the purposes of experiencing a conscious visual percept. The evidence we will present in the next two sections collectively supports this notion for size perception.

V1 activation can reflect perceptual outcome: Evidence from electrophysiology

Some V1 neurons respond in a way that reflects perceptual outcome as opposed to retinal input. Marg and Adams (1970) mapped receptive fields of neurons in the occipital lobe in patients with intractable epilepsy. Micro-electrodes had been placed there for the primary purpose of locating epileptic foci and the patients had consented to participate in an additional receptive field mapping research study. The diameter of the receptive field for different neurons was measured at different viewing distances. The authors found that for some neurons, but not all, the diameter of their receptive fields increased as viewing distance decreased – which runs contrary to what one would expect for receptive field maps obeying the principles of retinal stimulation and the laws of optics. Instead, the receptive fields for this class of neurons changed in a manner that kept the ratio between receptive-field size and retinal-image size constant at all viewing distances – reflecting characteristics of what the authors called a neurological zoom mechanism. A few years later, the same research group reported finding a similar class of neurons in V1 of the monkey (Smith and Marg, 1975).

More recently, a series of experiments performed in the monkey by Ni and colleagues (2014) also yielded remarkable findings with regards to the firing profile of some V1 neurons (Fig. 7). The authors inserted electrodes into V1 neurons and recorded their responses to the presentation of different sizes of concentric rings over a plain background or a Ponzo background placed in either the top position where the rings looked bigger or in the bottom position where the rings looked smaller. The authors found a number of V1 neurons responding in a way that reflected perceptual outcome as opposed to retinal input. For example, for a neuron responding preferentially to a medium-sized ring over the plain background (i.e. this size of ring passed through most of the neuron’s receptive field whereas the others did not), a smaller ring had to be presented in the top position of the Ponzo background and a larger ring had to be presented in the bottom position of the Ponzo background to invoke an equally strong response. From these results, combined with results from various control experiments, the authors concluded that the pictorial depth information from the Ponzo background was extracted at later stages of the visual system and that this depth information was then used, via feedback mechanisms, to cause some of the neurons in V1 to fire according to the perceived as opposed to the retinal size of the rings.

Can extra-retinal information about depth also modulate V1? The answer appears to be yes, which means that V1 can serve as an important hub for integrating size and distance information for the purposes of size constancy under multiple contexts. In a study performed by Dobbins and colleagues (1998), monkeys were presented with different-sized stimuli at different viewing distances, while the
authors recorded responses in neurons in V1, V2, and V4. Importantly, the stimuli were scaled with distance so that the retinal image size was the same with changing distance. In each of these areas, the authors found neurons that were modulated by retinal image size and others by viewing distance. They also found that 40% of the neurons they sampled in either V1 or V2 and 54% of the neurons they sampled in V4 responded preferentially to the physical (and presumably to the perceived) size of the stimuli, which would be indicative of a neural correlate for size constancy. Unlike the Ni et al. (2014)'s study with the Ponzo illusion in which depth information was only provided by pictorial cues, a number of binocular depth cues, both retinal as well as extra-retinal, provided information to the monkeys about distance. These cues are known to exert strong influences in V1. A number of studies have found that firing rates in V1 neurons can be modulated by vergence, retinal disparity, and gaze direction (Trotter et al., 1992; Trotter et al., 2004; Trotter and Celebrini, 1999; for a review see DeAngelis, 2000). In cases of changes in eye positioning, this modulation can happen quickly, reflecting corollary discharge from extraocular muscles (Trotter and Celebrini, 1999).

V1 activation can reflect perceptual outcome: Evidence from fMRI

In a landmark study, Murray and colleagues (2006) carried out an fMRI study to investigate how V1 responded to the size of flickering checkerboard spheres presented over a Ponzo-like illusory background (Fig. 8; Table 1). The background consisted of a corridor with converging lines that produced a vivid sense of depth. As expected, the participants perceived the sphere presented at the top as being larger (by 17% on average) than the one at the bottom, even though both had the same retinal image size. The authors also found that the increase in perceived size for the top sphere coincided with an anterior shift in the peak response of V1, where larger stimuli are retinotopically represented, while the decrease in perceived size for the bottom sphere coincided with a posterior shift in the peak response of V1, where smaller stimuli are retinotopically represented. Equally important, the spatial shift in the peak response of V1 predicted the perceptual magnitude of the illusion as determined psychophysically in a separate experiment performed outside of the fMRI scanner. This study, as far as we can discern, was the first to show that the retinotopic activity in V1 evoked by the presentation of a visual stimulus can be modulated by its perceived size, even when the size of the retinal image is held constant.

In a different fMRI study, Fang and colleagues (2008) performed similar experiments using the same Ponzo-like illusory background but incorporated two important modifications to the procedures (Table 1). First, the authors presented rings with spokes as opposed to spheres. Second, the study had participants attend to luminance perturbations on the ring or inside the centre of the ring where fixation was maintained. There were a number of advantages to these modifications. First, Fang et al. (2008) could examine V1 activation to visual stimulation at a particular eccentricity as opposed to within an eccentricity. Second, they could examine the effects of attention on V1 activation more directly. In agreement with the Murray et al. (2006) study, the authors showed that the flickering stimulus was perceived as being larger (by 15% on average) when presented at the top position relative to the bottom position and that this enlargement in size perception was accompanied by activation in the more eccentric retinal representations of V1. Moreover, they also showed that shifts in the location of activation in V1 were greater not only when the rings were attended to but also in participants who were more susceptible to the illusion, as assessed in a separate psychophysics study.
This earlier fMRI work was carried out in optical illusions whose perceptual rescaling of object size with apparent distance is much smaller (typically between 10 to 45%) than the perceptual rescaling of afterimage size with real distance, which can be near-perfect (Emmert, 1881; Leibowitz et al, 1969; Sperandio et al, 2010; Gregory, 2008). As explained earlier, this is likely due to the conflicting nature of different depth cues in the illusory pictures: those that inform the brain that there is depth and those that inform the brain that the eyes are focusing on a flat surface. We therefore carried out an fMRI study to examine how V1 responds when participants perceive afterimages of various sizes projected on a screen placed at different viewing distances (Sperandio et al, 2012a; Fig. 9; Table 1). We observed that the response of afterimages in V1 changed in accordance with its perceived size in a manner similar to those described by Murray et al. (2006) and Fang et al. (2008). We also examined BOLD responses in V2 and V3. As the lesion studies reviewed might suggest, we predicted that extra-striate areas in the ventral stream would also show evidence of neural responses reflecting size perception. Contrary to expectations, we could not find any evidence for this type of modulation in either V2 or V3 in the ventral stream. This may relate to differences in the sensitivity of the BOLD signal for measuring top-down modulation in striate versus extra-striate areas, which we will talk more about at the end of this section.

Pooresmaeili et al. (2013) used a different approach to examine size constancy mechanisms in the brain (Table 1). Their approach was based on the notion that adaptation leads to an illusory perception that does not correspond to the physical stimulus; namely, that the perception of a subsequent stimulus is influenced by a previously presented stimulus. For example, if one were to present a small circle followed by a big circle then the big circle would appear to be bigger than it actually is, and vice versa. This is precisely what the authors carried out in an fMRI experiment using Craik–O'Brien–Cornsweet (Cornsweet, 1970) circles (Fig. 10). A number of interesting findings came out of this experiment. First, the authors found both a reduction in the surface area of activated V1 cortical tissue when the adapter (i.e. the previous stimulus) caused the test stimulus (i.e. the stimulus being measured) to appear smaller and an augmentation of activated V1 cortical tissue when the adapter caused the test stimulus to appear larger. Second, the authors found that the amount of cortical tissue in V1 activated by the stimuli was strongly correlated with perceived size as determined in a separate visual psychophysics experiment. Third, when V1 activation was extracted from different regions-of-interest and plotted as a function of retinal eccentric representation, the authors found that the shape of this curve narrowed and widened when subjects perceived the stimuli as being smaller and larger respectively. Taken together, it seems clear that the BOLD response profile in V1 reflected the perceived as opposed to the retinal size of the stimuli. The authors argued that the effects could reflect primarily bottom-up processing in V1. However, as one of us has argued before (Chouinard and Ivanowich, 2014), having an adaptor and a test circle separated by 3 seconds provides far too much time between two stimuli to rule out effects of higher-order expectancies and other forms of cognitive supervision from feeding back into V1 and modulating its activity in response to the second stimulus. Indeed, using magnetoencephalography (MEG), which has far better temporal resolution than fMRI, Bar et al (2006) have shown that executive supervision by the frontal lobes influences activity in visual areas in a top-down manner. In addition, changes in task demands such as from naming versus passive viewing of objects (Chouinard et al, 2008) or directing versus not directing attention to visual stimuli (Murray and Wojciulik, 2004) can profoundly affect fMRI adaptation in visual areas.

Another interesting fMRI study is one on the moon illusion carried out by Weidner and colleagues (2014; Table 1). As explained earlier, the moon appears larger when it is located low on the
horizon then when it is high in the sky, although we know that the moon does not change in size (Fig. 4; Ross and Plug, 2002). In order to recreate this illusion inside the scanner, the authors used virtual reality goggles and the principles of relative disparity to present pictures of the moon with stereoscopic depth impressions; presenting two-dimensional photographs of the moon along different parts of its trajectory does not invoke the illusion. Using this virtual reality set-up, the authors were able to find an increase in perceived size by 3% (on average) when the moon was positioned on the horizon compared to when the moon was in its zenith position. The increased perception of size on the horizon relative to the zenith position was associated with an increase in neural activity in a number of areas in the ventral stream as well as in V1. The authors also carried out a second experiment. In this experiment, they made use of the principles of retinal disparity to manipulate changes in perceived depth and consequently size perception. Two conditions were presented: the first consisted of different perceived image sizes of the moon that did not vary in retinal image size while the second consisted of different retinal image sizes of the moon that appeared to be the same size to the participants. When the authors subtracted the BOLD response to the second condition from the first, they observed a release in adaptation for the perceived image size of the moon in the left V3 in the ventral stream, which they interpreted as reflecting neural processes related to the integration of size and distance information.

Hemodynamic responses, such as the BOLD response, provide valuable information about human brain function. So far, we reviewed a number of fMRI studies that have investigated how V1 responds in a way that reflects more the perceived as opposed to the retinal size of images in the context of optical illusions and afterimages. This brain function obviously depends on a neural substrate. Imaging techniques such as voxel-based morphometry and diffusion-tensor imaging have been developed to quantify relationships between abilities in cognition and inter-subject variability in grey matter volume (Maguire et al. 2000), cortical thickness (Lerch and Evans 2005), and diffusion anisotropy (Scholz et al. 2009). Thus, is it possible that the structural fabric of V1 may differ in individuals who are more susceptible to optical illusions? This question was answered in a study by Schwarzkopf et al. (2011). The study consisted of a visual psychophysics experiment aimed at calculating an individual’s susceptibility to the Ebbinghaus and Ponzo illusions, a retinotopic fMRI experiment for demarking an individual’s boundaries for V1, V2, and V3, and a structural MRI experiment to measure grey-matter volume in each of these visual areas. The authors found that susceptibility to the Ponzo and the Ebbinghaus illusions correlated negatively with the physical extent of V1 (Ponzo illusion: R^2 = 0.23; Ebbinghaus: R^2 = 0.14). In contrast, similar correlations were not observed for either V2 or V3. Why were these correlations negative as opposed to positive? The reasons for this are not entirely clear. Schwarzkopf et al. (2011) speculated in their paper that horizontal connections, a principal component of intrinsic cortical circuitry necessary for processing information and maintaining cortical representations, may be less effective if information has to span longer distances and hence larger V1 may in turn result in less susceptibility to optical illusions. The reasons for the lack of correlations in V2 and V3 are also unknown.

As discussed, electrophysiological and lesion studies provide clear evidence for an involvement of the ventral stream in size constancy. So why is it the case that fMRI studies frequently find these effects more consistently in V1? We propose that the integration of size and distance information in V1 for the purposes of a size-distance scaling is most likely the result of top-down mechanisms, for which the BOLD response is a more sensitive measurement for detecting in striate relative to extra-striate areas. Logothetis and collaborators have shown in monkeys performing binocular rivalry tasks that firing activity for perceived images is stronger in higher associative areas in the ventral stream relative to V1.
(Leopold and Logothetis, 1996; Sheinberg and Logothetis, 1997; Wilke et al., 2006) while others consistently show the reverse patterns of results with fMRI in humans (Polonsky et al., 2000); (Haynes and Rees, 2005). It was later demonstrated that this incompatibility between electrophysiological and neuroimaging results was due to the BOLD response being more sensitive to changes in sustained top-down synaptic input than neuronal firing activity measured by cell recording studies (Maier et al., 2008).

As we reviewed in this section, BOLD activation in V1 reflects the perceived outcome more than the retinal input of stimuli in the context of optical illusions and afterimages, similar to fMRI studies on binocular rivalry. These types of responses are also stronger in V1 than in higher-order visual areas of the ventral stream, perhaps reflecting the importance of top-down feedback into V1 for the integration of size and distance information.

**Closing remarks**

There is more to sight than meets the eye. Sight can also be the product of multisensory processing. As highlighted in this review, there exists plenty of evidence that our visual perception mediating size constancy is not driven solely by retinal signals but also by extra-retinal signals originating from the eyes and other sensory organs. What is increasingly apparent is how activity in V1 can reflect the outcome of this processing and is not restricted to the integration of information from the retina. It then follows that V1 serves as an important node in the integration of multiple signals for the purposes of creating a stable visual experience of perceived size. It would be peculiar that this role would be confined to object size and not include other physical features of objects as well. Other investigations, which we did not review, have reported similar response profiles in V1 that reflect the perceived outcome as opposed to the retinal input for motion (Whitney et al., 2003), colour (Sasaki and Watanabe, 2004; Bannert and Bartels, 2013), contrast (Ress and Heeger, 2003), and luminance (Boyaci et al., 2007), which may equally reflect mechanisms of perceptual constancy. Thus, an important take-home message from this review is that V1 fulfills much more than an entry position role in a series of hierarchical cortical events leading up to visual perception. Future investigations should examine these mechanisms in greater depth to understand more precisely how V1 contributes to visual perception and perceptual constancies.

**Acknowledgements:**

We wish to thank D. Vendramini for his help with the figures, O. Landry and C. Fisher for reading later drafts of this manuscript and providing us with their insightful comments. We thank Scott O. Murray for providing us with the corridor display shown in figures 7 and 8, which he and his colleagues used for the studies described in the figures. This work was supported by a La Trobe University's Research Focus Area (RFA) Understanding Disease grant.
### Table 1.

**Overview of fMRI experiments presented in the paper**

<table>
<thead>
<tr>
<th>Study</th>
<th>Nature of cues</th>
<th>Nature of experiments</th>
<th>Nature of fMRI approach</th>
<th>General results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murray et al. (2006)</td>
<td>• Size: Flickering checkerboard spheres • Depth: Pictorial cues in the context of the Ponzo illusion</td>
<td>• Passive viewing • Compared two versions of the Ponzo: One rich and the other poor in pictorial depth cues</td>
<td>• ROI approach in V1 • ROIs for V1 were defined by a retinotopic localiser, each one representing a different level of retinotopic eccentricity</td>
<td>• BOLD was greatest in eccentric ROIs for perceived size; effect was modulated by richness in pictorial depth cues</td>
</tr>
<tr>
<td>Fang et al. (2008)</td>
<td>• Size: Flickering checkerboard rings • Depth: Pictorial cues in the context of the Ponzo illusion</td>
<td>• Attended to ring: Detect pause in flicker • Attended to fixation inside spoke: Detect luminance change in fixation</td>
<td>• ROI approach in V1, LO, PPA • ROIs for V1 were defined by a retinotopic localiser, each one representing a different level of retinotopic eccentricity</td>
<td>• V1-BOLD was greatest in eccentric ROIs for perceived size; effect was modulated by attention • Spatial shift in peak BOLD response correlated with degree of behavioural susceptibility • Main effect of attention on LO and PPA (i.e. attend-to-ring &gt; attend-to-fixation)</td>
</tr>
<tr>
<td>Sperandio et al. (2012c)</td>
<td>• Size: Afterimages • Depth: Range of monocular and binocular cues. Projection of afterimages on a surface at different viewing distances. Presentation of circular stimuli at different viewing distances</td>
<td>• Projection of afterimages on a surface at different viewing distances • Circular stimuli viewed at different distances • Participants manually indicated onset, offset, and size of afterimages</td>
<td>• ROI approach in V1, V2v, V3v, V2d, V3d • ROIs for these areas were defined by a retinotopic localiser, each one representing a different level of retinotopic eccentricity</td>
<td>• V1-BOLD to afterimages greatest in eccentric ROIs for perceived size of afterimages and circles • V1-BOLD correlated with duration of afterimages • These effects were not evident in the other visual areas</td>
</tr>
<tr>
<td>Pooresmaeili et al. (2013)</td>
<td>• Size: Craik O’Brien–Cornsweet circles • Depth: Depth cues not manipulated</td>
<td>• Pairs of circles repeated under passive viewing; adaptation to specific stimulus size affected perception of second stimulus</td>
<td>• Voxel count in V1 • ROI approach in V1, V2, V3, LO • ROIs for V1, V2, V3 were defined by a retinotopic localiser; each one representing a different level of retinotopic eccentricity • ROIs for LO were defined by a standard localiser</td>
<td>• Voxel count in V1 correlated with perceived size • V1-BOLD and V2-BOLD were greatest in eccentric ROIs for perceived size • Adaptation for retinal size in LO</td>
</tr>
<tr>
<td>Weidner et al. (2014)</td>
<td>• Size: Moon presented in a virtual 3D environment • Depth: Stereoscopic depth cues delivered with virtual 3D goggles</td>
<td>• Performed a size matching task for moon in zenith and horizon positions • Adaptation paradigm for repeated perceived size versus repeated retinal size</td>
<td>• Whole-brain voxel-wise approach was used • ROIs were obtained from resulting clusters derived for the voxel-wise approach</td>
<td>• V1-BOLD and Fusiform-BOLD were greater when moon was in horizon • Increased left V3v-BOLD amplitude in response to perceived size</td>
</tr>
</tbody>
</table>
Figure 1. Retinotopic mapping of human visual areas using the phase-encoding approach. This approach to mapping retinal responses in the visual cortex is based on the principle that when a stimulus is presented in a cyclical manner, the BOLD signal will be modulated in a similar cyclical pattern (Fig 1a). A movie of a stimulus moving across different parts of the visual field is played to the participant in a repeating loop (panel B: for polar angle mapping; panel C: for eccentricity mapping) while the BOLD signal is collected over time. The BOLD signal is then analysed using a Fourier analysis to determine when and where in the visual field the response is the highest. As shown in panel A, the BOLD signal is highest as the stimulus passes through the receptive field for a specific cortical region while the BOLD signal in this same cortical region is lowest when the stimulus is no longer in the receptive field. The results from these kinds of stimulation typically give rise to what is shown in panels D and E. As shown in panel D, green denotes the vertical meridians defining the boundaries of V1 while red denotes the horizontal meridian for the right visual field. The upper visual field is represented in inferior structures while the lower visual field is represented in superior structures. As shown in panel E, purple / red denotes the fovea while cooler colours represent more peripheral parts of the retina. Note how the fovea is represented in proportionally more cortex relative to the more peripheral parts of the retina.
Figure 2. The relationship between retinal image size and distance. If the same object is placed at different distances, such as the mouse illustrated in panel A, then its image on the retina will vary in size following Euclid’s law. According to Euclid’s law, the size of the image on the retina will decrease proportionally as a function of distance. Different sized objects, such as the mouse and the elephant in panel B, can have the same retinal image size if the relative distance between the two is just right. Namely, a mouse close by can have the same retinal image size as an elephant far away. An inducing stimulus, such as the yellow circle in panel C, will bleach a certain part of the retina (also in yellow). The result of this bleaching will cause a subsequent afterimage to appear several seconds later (in dark purple). According to Emmert’s law, the perceived size of the afterimage will increase proportionally as a function of the distance with which one projects the afterimage on. The diagrams are not drawn to scale.
Figure 3. Examples of optical illusions. The Ebbinghaus illusion (A) is an example of a size-contrast illusion whereby the size of the inner circles is perceived relative to the size of the surrounding context. Specifically, the inner circle looks bigger when surrounded by smaller circles and smaller when surrounded by bigger circles. Figure adapted from Sperandio et al. (2012b). In this example of the Ponzo illusion (B), the length of the two red lines is identical but the converging lines in the background make the upper one appear longer. These types of pictorial cues about distance are frequently present in our environment. For example, they can be observed when we look at railroad tracks, hall corridors, roads, and sidewalks.
Figure 4. The moon illusion. The moon appears larger when it is located on the horizon then when it is in its zenith position in the night sky, even though we know that the moon does not change in size. Interestingly, the larger-looking moon on the horizon will not appear so large any more if the moon is observed through a peephole such that the rest of the visual scene is occluded from view. Although this phenomenon has been studied for centuries, there is still no consensus on its explanation. The diagram is not drawn to scale.
In a classic study, Holway and Boring (1941) had participants judge the size of circles at different viewing distances. The circles were always presented with a visual angle of 1 degree. Specifically, the physical size of the circles increased proportionally as a function of distance such that the retinal image size was always constant. As the availability of distance cues was reduced, participants began to judge the size of the circles according to their retinal image size as opposed to their actual size. Many of the studies we review in this paper used a similar paradigm in order to present stimuli with the same retinal image size at different viewing distances (e.g. Humphrey and Weiskrantz, 1969; Ungerleider et al., 1977; Sperandio et al., 2010; Combe and Wexler, 2010; Servos, 2006; Dobbins et al, 1998). The diagram is not drawn to scale.
Figure 6. The Taylor illusion. To experience the illusion, an afterimage of the hand or of a hand-held object (e.g. a ring of lights as illustrated here) is generated in complete darkness by means of a flash of light. The hand is then moved right after the induction of the afterimage either towards or away from the body in a continuous and smooth fashion. One can readily observe that the afterimage will progressively decrease in size as the hand approaches the body (panel B) and vice versa increase in size as the hand is moved away from the body (panel C). Figure adapted from Sperandio et al. (2013a).
Figure 7. Firing properties of neurons that reflect perceived size. The study performed by Ni and colleagues (2014) examined receptive field responses in V1 neurons in the monkey while they looked at rings over a Ponzo illusory background like the one shown here. In these simplified examples, the neuron in either upper or lower panel would fire equally poorly to the presentation of either ring presented without the illusory background. When the smaller ring was presented in the top position of the illusory background (upper panel), the authors showed that this would produce a stronger response because the context made it look bigger and its apparent size would cover a greater surface area of the neuron’s receptive field. When the bigger ring was presented in the bottom position of the illusory background (lower panel), the authors showed that it would produce a stronger response because the context made it look smaller and its apparent size would cover a greater surface area of the neuron’s receptive field.
Figure 8. An fMRI study on the Ponzo illusion. The study performed by Murray and colleagues (2006) examined BOLD responses to the presentation of flickering checkerboard spheres on a Ponzo illusory background like the one shown here (A). The BOLD responses were extracted from various pre-defined eccentric ROIs in V1. The top circle, which appeared larger to participants, invoked larger BOLD responses in an ROI for a larger eccentricity (drawn schematically in purple in panel B) than the one for the retinal image size of the stimulus (drawn schematically in yellow in panel B). The bottom circle, which appeared smaller to participants, invoked larger BOLD responses in an ROI for a smaller eccentricity (drawn schematically in green in panel B) than the one for the retinal image size of the stimulus (drawn schematically in yellow in panel B).
Figure 9. An fMRI study on afterimages. In our fMRI study presented in this review, we examined the role of V1 in the size perception of afterimages (Sperandio et al., 2012). A) BOLD signal in V1 was extracted from eccentricity ROIs for the closest viewing distance (distance 1) vs. the furthest viewing distance (distance 5). See the legend on the inflated brain in panel B for colour coding of each ROI along the calcarine sulcus. At the beginning of the time course, when the inducing light was presented, we observed the greatest BOLD response in the smallest eccentricity ROI (red curve) for both viewing distances. In fact, this ROI subtended 4.1° which corresponded to the retinal size of the inducing light. In contrast, during the perception of the afterimage (indicated by the two solid lines), we observed the greatest BOLD response in the biggest eccentricity ROIs (blue and purple curves) at the furthest viewing distance, when participants experienced the biggest afterimages. The emergence of activity in regions that were never stimulated by light is a clear demonstration that V1 responds to the perceived size of an afterimage. B) Correlation between perceived duration of the afterimage at the furthest viewing distances and area under the curve of the BOLD signal extracted from the biggest eccentricity ROIs. Every individual is represented as a data point while the solid line indicates the best-fitting line for the observed data. Figures adapted from Sperandio et al. (2012a).
Figure 10. Craik–O'Brien–Cornsweet circles. In an adaptation study, Pooresmaeili et al (2013) presented Craik–O'Brien–Cornsweet circles similar to the ones shown here. When presenting an adapter followed by a smaller test stimulus, participants perceived the test stimulus as being smaller, and vice versa. In Craik–O'Brien–Cornsweet circles, the overall luminance inside is the same as the background yet we perceive the inside as having a slightly different luminance. In this case, it is perceived as lighter.
References


