

Steady-state visually evoked potentials reveal partial size constancy in early visual cortex

Jing Chen*

School of Psychology, Shanghai University of Sport,
Shanghai, China
Abteilung Allgemeine Psychologie,
Justus-Liebig-Universität Gießen, Gießen, Germany



Meaghan McManus*

Centre for Vision Research, York University, Toronto,
Ontario, Canada



Matteo Valsecchi

Abteilung Allgemeine Psychologie,
Justus-Liebig-Universität Gießen, Gießen, Germany



Laurence R. Harris

Centre for Vision Research, York University,
Toronto, Ontario, Canada



Karl R. Gegenfurtner

Abteilung Allgemeine Psychologie,
Justus-Liebig-Universität Gießen, Gießen, Germany



Our visual system maintains a stable representation of object size when viewing distance, and thus retinal size, changes. Previous studies have revealed that the extent of an object's representation in V1 shows systematic deviations from strict retinotopy when the object is perceived to be at different distances. It remains unknown, however, to what degree V1 activity accounts for perceptual size constancy. We investigated the neural correlates of size-constancy using steady-state visually evoked potentials (SSVEP) known to originate in early visual cortex. Flickering stimuli of various sizes were placed at a viewing distance of 40 cm and stimuli twice as large were shown at 80 cm. Thus both sets of stimuli had identical retinal sizes. At a constant viewing distance, SSVEP amplitude increased as a function of increasing retinal size. Crucially, SSVEP was larger when stimuli of a given retinal size were presented at 80 cm compared with at 40 cm independent of flicker frequency. Experiments were repeated and extended in virtual reality. Our results agree with previous findings showing that V1 activity plays a role in size constancy. Furthermore, we estimated the degree of the neural correction for the SSVEP as being close to 50% of the perceptual size constancy. This was the case in all experiments, independent of the effectiveness of perceptual size constancy. We conclude that retinotopy in V1 does get quite massively adjusted by perceived size, but not to the same extent as perceptual judgments.

Introduction

Size constancy is one of the most remarkable features of the human visual system. We are capable of perceiving the “true” constant, physical size of an object despite massive changes in the size its image subtends on the retina when it is seen at different distances (e.g., Holway & Boring, 1941). The neural basis of size constancy is still relatively unclear, but there have been several studies suggesting that a relatively size invariant representation of the world might arise as early as in primary visual cortex (Murray, Boyaci, & Kersten, 2006; Fang, Boyaci, Kersten, & Murray, 2008; Sperandio, Chouinard, & Goodale, 2012; Ni, Murray, & Horwitz, 2014; He, Mo, Wang, & Fang, 2015). This would be remarkable because early cortical areas have generally been described as having a strictly retinotopic organization (e.g., Engel, Glover, & Wandell, 1997; Tootell, Hadjikhani, Mendola, Marrett, & Dale, 1998; Wandell, Dumoulin, & Brewer, 2007). A representation that includes size constancy is not compatible with a fixed one-to-one mapping between the retina and cortex because the area of cortex allocated to a given object would need to vary with perceived size.

An open question is the degree to which the representation in early visual cortex does show size

Citation: Chen, J., McManus, M., Valsecchi, M., Harris, L. R., & Gegenfurtner, K. R. (2019). Steady-state visually evoked potentials reveal partial size constancy in early visual cortex. *Journal of Vision*, 19(6):8, 1–15, <https://doi.org/10.1167/19.6.8>.

<https://doi.org/10.1167/19.6.8>

Received October 30, 2018; published June 11, 2019

ISSN 1534-7362 Copyright 2019 The Authors



constancy. In studies using pictorial illusions to induce perceived size changes a relatively good agreement between perceptual and physiological effects was found (Murray et al., 2006; Fang et al., 2008; Ni, Murray, & Horwitz, 2014; He et al., 2015). However, typically the perceptual effects used were modest, up to 30% change in perceived size, and the perceptual and physiological data were difficult to compare. Sperandio et al. (2012) used afterimages to study size constancy at different physical distances, rather than at distances that differed only perceptually. In accordance with Emmert's law (Emmert, 1881) they obtained good size constancy and found a proportional correction in primary visual cortex (V1). Under Sperandio et al.'s conditions, however, a quantitative comparison of neural and physical size changes was difficult because the response to afterimages cannot easily be compared to responses caused by real physical stimuli. Our approach uses identical stimuli presented at different distances to allow us to get an excellent degree of size constancy and perceptual and neurophysiological measurements that can be directly compared.

The degree of cortical correction for size constancy is important because any deviation from retinotopy is bound to lead to changes in the topology of the neural representation of the scene. For example, if, nonoverlapping objects of equal size were simultaneously viewed at different distances, then an increase in the size of the cortical representation of the more distant objects might lead to the representations of two objects overlapping in V1. This would be in stark contrast to our veridical perception of such scenes.

We addressed these issues by systematically varying the degree of perceptual size constancy in real-world scenes and in virtual reality (VR), from perfect size constancy to almost none. This allowed us to assess any potential correlation between cortical activity and perceptual constancy. We assessed early cortical activity by measuring steady-state visual evoked potentials (SSVEP), an oscillatory brain response to periodic visual stimulation (see Norcia, Appelbaum, Ales, Cottureau, & Rossion, 2015, for a review) that is thought to originate largely from the primary visual cortex (Di Russo et al., 2007; Müller, Teder, & Hillyard, 1997; Wittevrongel et al., 2018).

In Experiment 1 we measured SSVEPs to stimuli presented at two physical viewing distances (40 cm and 80 cm), and in three locations in the visual field. The retinal sizes of the stimuli were kept constant. In Experiment 2 we varied flicker frequency to investigate the level at which the visual representation becomes size-constant, under the assumption that the temporal response at higher flickering frequencies is severely attenuated in higher visual areas due to increased spatial and temporal averaging (e.g., Hawken, Shapley, & Grosf, 1996; Lennie, 1998). In Experiment 3, we

measured size constancy in VR. VR allowed us to manipulate depth cues and include a control condition where all binocular and pictorial cues to distance were removed, which would be difficult to achieve in a classical psychophysical setup on a monitor, where the edges of the displays and other features of the room are inevitably illuminated. In Experiment 4 we asked observers to do a perceptual adjustment task in both the real-world setup used in Experiments 1 and 2, and in the VR environment used in Experiment 3.

Methods

Three SSVEP experiments and a perceptual adjustment task were conducted in agreement with the Declaration of Helsinki and followed guidelines approved by the local ethics committee (Giessen IRB#2017-0028). All observers signed informed consent forms before taking part in the experiment and were naïve as to the purpose of the study at the time of testing. They had normal or corrected-to-normal vision and had no known neurological or oculomotor diseases.

Apparatus

In Experiments 1 and 2, stimuli were displayed using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007) in MATLAB (MathWorks, Natick, MA), on a 120 Hz Samsung SyncMaster 2230R7 22-inch monitor (Samsung Group, Seoul, South Korea). The screen had a spatial resolution of $1,680 \times 1,050$ pixels and extended 61° horizontally and 38° vertically at a viewing distance of 40 cm. Experiments 1 and 2 were conducted in a dimly lit room.

In Experiment 3, stimuli were presented in an Oculus Rift CV1 virtual headset. The Oculus Rift had a field of view that extended approximately $\pm 110^\circ$ diagonally. The screen had a $1,080 \times 1,200$ -pixel resolution per eye and a 90 Hz refresh rate. The stimuli were created in Unity (Version 5.5.2f1, Unity Technologies, San Francisco, CA).

Methods specific to Experiment 1: Two distances

Participants: Eight observers (six females and two males, age 19–33, average: 27 years) participated in this experiment.

Stimulus and procedure: Filled circles, whose luminance was modulated sinusoidally between black (0.2 cd/m^2) and white (215.8 cd/m^2) at 8 Hz, were presented either in the left or right visual field (10° from the center) or at the center of the screen. The background was black

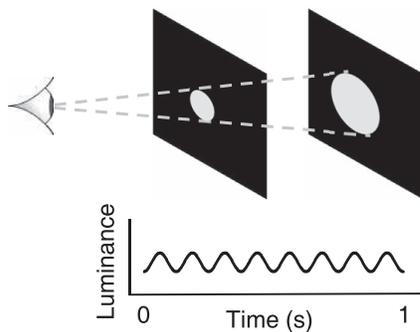


Figure 1. Stimulus displays. Participants viewed filled circles on a monitor placed at a distance of either 40 cm or 80 cm in different experimental blocks, with the filled circles matched in retinal size. The luminance of the circles was modulated sinusoidally at 8 Hz to elicit SSVEPs.

(Figure 1). A blue dot (radius = 0.2°) was always displayed in the center as a fixation spot. In each trial, the diameter of the filled circle was stepped through 10 different sizes from 1° to 10° in steps of 1° , in an increasing or decreasing order (balanced across trials). Each size was presented for 5 seconds, except the first size that was presented for 7 seconds. Thus, each trial took 52 s of which the first 2 s were excluded from the analysis to remove the abrupt visual response due to stimulus onset.

Participants were required to passively view the flickering, size-stepping stimuli at either a distance of 40 cm or 80 cm, in two experimental blocks, while maintaining fixation on the fixation circle. To change the distance between two blocks, participants were moved while the monitor remained at the same location. To match the retinal sizes, the physical sizes of the stimuli on the monitor at 80 cm were doubled compared with the stimuli at 40 cm (Figure 1). Half of the participants started the experiment at the 40 cm distance, and the other half started at the 80 cm distance. Participants conducted six trials for three locations on screen (left visual field, center, right visual field) at each distance (40 cm, 80 cm), resulted in 36 trials in total. For each distance, the order of trials were randomized. See Table 1 for a summary of the conditions.

Methods specific to Experiment 2: Two frequencies

Participants: Eight observers (four females and four males, age 24–32, average: 27 years) took part in this experiment. None of them participated in Experiment 1. **Stimulus and procedure:** The experimental settings were the same as in Experiment 1. The differences were that (1), both 8 and 30 Hz stimulation frequency were used; (2), stimuli were only presented in the center, not in the periphery; and (3), the step sizes ranged from 2° to 8° in an increment of 1° (seven different sizes), which resulted in a trial length of 37 s (5 s for each size, plus the extra 2 s for the first size) presented either stepping up or down in counterbalanced order. See Table 1 for a summary of the conditions.

Methods specific to Experiment 3: SSVEP in VR

Participants: Twenty observers (13 females and seven males, age 19–32, average: 26 years) took part in this experiment. Two of them had participated in Experiment 2.

Stimulus and procedure: The main target stimulus was a very thin cube (>1 cm in depth) that flickered between black and white at 5 Hz. We changed from the 8 Hz used in Experiment 1 and 2 because with a 90 Hz refresh rate the Oculus Rift cannot display an 8 Hz pattern-reverse flickering stimulus. The flickering target was bisected by a black plane (47.5 cm \times 30 cm) such that it was just visible. The plane mimicked the monitor used in Experiment 1 and 2 (see Figure 2). Due to the angle of view and how it was placed against the plane, the cube looked like a square. A square shape was used in Experiment 3 instead of a circle because it was easier to ensure that Unity was rendering the object to the correct size. If a sphere had been used instead, the front of the sphere would have been closer to the participants than other parts of the sphere. As well, circles, or flattened spheres, tend to look like a many sided shape such as an octakaidecagon.

Two different environments were used in VR: a 3D hallway condition and a featureless condition that had significantly reduced distance cues. In the 3D hallway condition (Figure 2A), distance cues including the 3D context and binocular disparities were available. The

Experiment	Distance near	Sizes near	Distance far	Sizes far	Other comparisons
1	40 cm	1° to 10° in steps of 1°	80 cm	Retinally matched to near	8 Hz, Visual field location comparison: Center vs Periphery
2	40 cm	2° to 8° in steps of 1°	80 cm	Retinally matched to near	Frequency comparison: 8 Hz vs 30 Hz
3	40 cm	2° to 8° in steps of 2°	80 cm	Retinally matched to near	5Hz, Environment with 3D cues vs Featureless environment

Table 1. Summary of the main experimental conditions for each experiment.

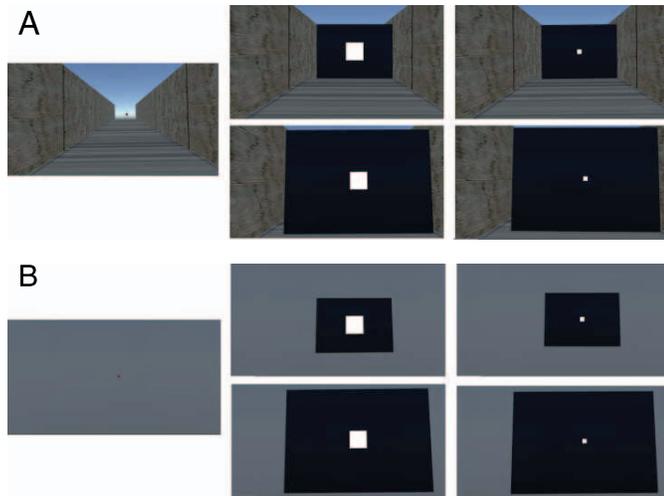


Figure 2. Representation of what the participants saw in Experiment 3. The images are screen shots from the view on the computer. (A) The 3D hallway condition. The far left image is the fixation spot at the far distance. The top left image is the target at its largest size, and the top right image is the target at its smallest size, both at far distance (80 cm). The bottom left image is the target at its largest size and the bottom right image is the target at its smallest size, both at near distance (40 cm). Stimuli were displayed binocularly. (B) The target descriptions for the featureless control condition are the same as A but with a featureless background and monocular viewing, where both eyes receive identical stimuli.

featureless control condition (Figure 2B) consisted of a gray environment (lack of 3D context). Binocular cues were also removed in this condition by making the right and left eye's displays identical. Parallax was available but head movements were limited by means of a chin rest.

The square could be either simulated as close to the participant (40 cm) or far from them (80 cm). Before each trial, a red fixation spot was presented at the stimulus location. The red fixation spot was more noticeable in the 3D environment than the blue fixation spot used previously. After a button press, the fixation spot disappeared, and the flickering target square appeared. We stepped through four different retinal sizes (2° , 4° , 6° , and 8° in increments of 2°). At the 40 cm distance, this corresponded to a side length of 1.4, 2.8, 4.2, and 5.6 cm. At the 80 cm distance, the side lengths were 2.8, 5.6, 8.4, and 11.2 cm. The square was always initially presented at the minimum size (1.4 cm near, or 2.8 cm far) or the maximum size (5.6 cm near, or 11.2 cm far), with the starting size counterbalanced between trials. It was then increased or decreased in size respectively every 5 s until the end of the range, at which point the square reversed its stepping direction back to its initial size. There were 40 trials in total with 10 trials for each condition (two environments, two distances). See Table 1 for a summary of the conditions.

Methods specific to Experiment 4: Perceptual adjustment task

Observers performed a perceptual adjustment task in the real-world setup used in Experiments 1 and 2 (four observers), and in the VR environment used in Experiment 3 (eight observers). In the adjustment task, the stimuli were flickering as in the SSVEP experiments but no electroencephalogram (EEG) recordings were made. For the real-world setup, a comparison-filled circle (diameter = 2° , 4° , 6° , 8° , or 10°) was presented at one distance (e.g., 40 cm) on one monitor, while participants were required to press keyboard buttons to adjust the size of the other filled circle displayed at the other distance (e.g., 80 cm) on another monitor, until both circles had identical perceived sizes. The adjusted target started at a random size in the range of 50% smaller to 50% larger than the comparison size. Four trials were conducted for each size and each distance, resulting in 40 trials. For the VR environment, participants viewed the stimulus at one distance for 5 s, and subsequently were required to adjust the size of the stimulus at the other distance. Two sizes were tested (4° and 8°) at each distance. There were 40 trials in total (five for each test condition).

EEG recordings and analyses

In Experiment 1, the EEG was recorded from 32 scalp sites according to the international 10–20 system (FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Pz, Oz, FC1, FC2, CP1, CP2, FC5, FC6, CP5, CP6, TP9, TP10, HLeo, Veo, HReo). Signals were amplified (Brain Products GmbH, Munich, Germany) and sampled at 1000 Hz. The ground electrode was placed at the AFz location, and the on-line reference electrode at the Cz location. Electrode impedances were kept below 5 k Ω .

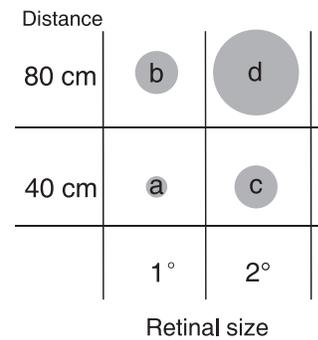
In Experiment 2, we switched to an active electrodes EEG system. The EEG was recorded from 32 electrodes (actiCAP, Brain Products) at 5000 Hz sampling rate. The ground electrode was placed at FPz, and the on-line reference electrode at FCz location. Electrode impedances were below 25 k Ω (the lowest impedance signaled by the actiCAP system).

Experiment 3 had identical EEG setup, except that only Oz, the ground electrode and the reference electrode were placed on the head. As participants had to wear the Oculus Rift headset, it was not possible to record from more electrodes. We chose to record at Oz, which showed the highest SSVEP responses in Experiments 1 and 2 (Figure 4; see also Chen, Valsecchi, & Gegenfurtner, 2017a, 2017b). Note that the VR headset inevitably causes some additional baseline noise in EEG signals. It is, however, less of a problem for the present study, as we measured the amplitude of responses at the stimulus frequency relative to the

background noise assessed at nearby frequency bins. The SSVEP technique has been successfully used in VR, mostly for brain computer interfaces (Royer, Doud, Rose, & He, 2010; see Kerouš & Liarokapis, 2016, for a recent review).

Analyses were carried out using EEGLAB toolbox (Delorme & Makeig, 2004) and customized scripts in MATLAB (MathWorks). EEG signals were re-referenced to a common average reference in Experiments 1 and 2 (in Experiment 3, re-referencing was not possible because only Oz electrode was recorded). EEG epochs lasting 5 s (corresponding to a flickering stimulus with a certain size) were cut out. Each epoch was first de-trended by removing the linear fit (Bach & Meigen, 1999), and multiplied by a Tukey window (i.e., tapered cosine window, $\alpha = 0.2$). Fast Fourier transform (*fft.m* in MATLAB) was used to obtain the amplitude spectrum. At each frequency (e.g., 8 Hz), we subtracted from the peak amplitude the average amplitude of the four nearby bins (e.g., 7.6, 7.8, 8.2, 8.4 Hz), so that the background noise was removed (e.g., Liu-Shuang, Torfs, & Rossion, 2016). To calculate the total SSVEP amplitude, we summed all harmonics below 45 Hz. Note that including harmonics without significant powers would not change the result, as their amplitudes were close to zero after subtracting the background noise. All the SSVEP analyses were identical in all three experiments, except that we did not include the second harmonic (i.e., 10 Hz) of the 5 Hz stimulus in Experiment 3. The reason was that 10 Hz lies exactly at the peak of spontaneous alpha oscillations, which tends to increase throughout the experiment as participants get tired. The main results remained unchanged regardless of whether the 10 Hz harmonic was included or not. For analyses of SSVEP phases we used the circular statistic toolbox in MATLAB (Berens, 2009).

SSVEP amplitudes differ largely between the central and peripheral visual field, between different frequencies, and across participants. We used standard procedures to normalize the amplitudes (Andersen, Fuchs, & Müller, 2011; Andersen, Müller, & Hillyard, 2015), by transforming them into z scores for each participant, separately for locations in visual field (Experiment 1) and separately for each stimulus frequency (Experiment 2). We excluded data points with z scores exceeding the range of $[-2.5, 2.5]$ (0.59% of total trials for Experiment 1; 0.67% for Experiment 2; 0.97% for Experiment 3). Results were similar regardless of whether outliers were removed or not. In Experiments 1 and 2, the average amplitudes at O1, Oz, and O2 electrodes were used for statistics, as the SSVEP responses were confined to these electrodes (Figure 4). In Experiment 3, the amplitudes at Oz were used, which was the only electrode recorded.



Perceptual size constancy index

$$= \frac{P(b) - P(a)}{P(c) - P(a)} \times 100\%$$

SSVEP size constancy index

$$= \frac{S(b) - S(a)}{S(c) - S(a)} \times 100\%$$

$P(x)$: perceptual size of object x
 $S(x)$: SSVEP response to object x

Figure 3. Calculation of perceptual and SSVEP size constancy indices (see text for details). The four circles (i.e., a , b , c , d) are drawn to reflect their physical size rather than retinal size. Note that “ d ” is not used in the calculation but is kept here as a demonstration of the physical size ratios used.

In all figures in the results, we report *within-subject* confidence intervals calculated with the method provided by Cousineau (2005), as between-subject variability measurements are misleading for a within-subject design such as the current study.

Calculation of the size constancy indices

Based on the results from the adjustment task a *perceptual size constancy index* can be computed, the rationale of which was borrowed from the literature on color constancy (Arend, Reeves, Schirillo, & Goldstein, 1991). Figure 3 illustrates the calculation. Supposing there were four objects (a , b , c , d), their physical sizes were in a ratio of 1:2:2:4, as displayed in Figure 3. Objects b and d were placed at 80 cm distance, while objects a and c were placed at 40 cm distance. As a result, objects a and b (1° in visual angle), and objects c and d (2° in visual angle) were matched for retinal size. If size constancy were perfect, b should be perceived as the same size as c . If there were no size constancy, and perception was based purely on retinal projections, b should be perceived as the same size as a . Therefore, a perceptual size constancy index can be calculated as

$$\text{Perceptual size constancy index} \\ = \frac{P(b) - P(a)}{P(c) - P(a)} \times 100\%,$$

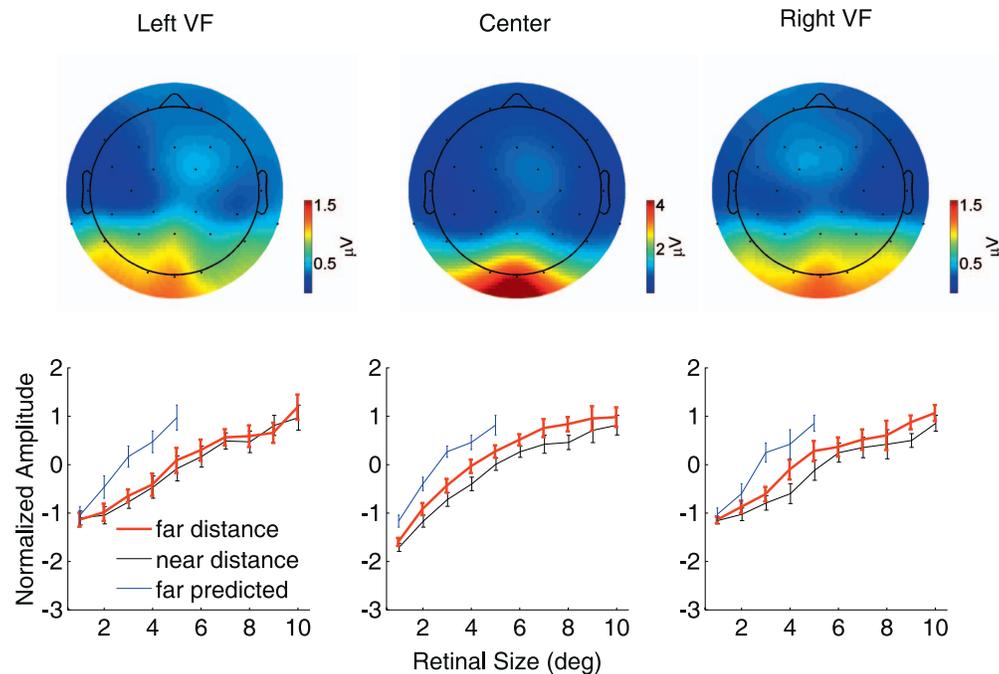


Figure 4. SSVEP amplitudes in Experiment 1 shown separately for stimuli in the left visual field, in the center, and in the right visual field. The top portion of the figure shows topographic plots, with the responses mainly confined to occipital electrodes (O1, Oz, and O2). SSVEPs to stimuli in the center were much stronger than those in the periphery (note the different scale). At the bottom, normalized SSVEP amplitudes (i.e., z score transformed, separately for left VF, center, and right VF) as a function of retinal size for the far distance (red curve) and near distance (gray curve). The blue curve (“far predicted”) draws the predicted SSVEP responses for far stimuli if SSVEPs depend only on object sizes (e.g., the predicted SSVEP to a 1° stimulus at far is equal to the SSVEP to a 2° stimulus at near, as they have identical object sizes). Error bars represent within-participant 95% confidence intervals (Cousineau, 2005).

where $P(a)$ denotes the perceptually matched size of object a (the same for $P(b)$, $P(c)$, $P(d)$). The computed index has a meaningful range from 0 to 100%, with 0 indicating no constancy, and 100% indicating perfect constancy. In the perceptual adjustment task, observers adjusted both the near object to match the far one, and the far object to match the near one. The average of both adjustments was used for the calculation of constancy index.

We applied the same calculation to the SSVEP responses to object a , b , c , and d , as opposed to the perceived sizes determined in the adjustment task. Therefore, the

$$\text{SSVEP size constancy index} = \frac{S(b) - S(a)}{S(c) - S(a)} \times 100\%,$$

where the $S(a)$ denotes the SSVEP amplitude in response to object a (the same to $S(b)$, $S(c)$, $S(d)$). As individual EEG responses tended to be noisy, the calculation was done only on SSVEP amplitudes averaged across observers.

Results

Experiment 1: SSVEPs are correlated with retinal size as well as with perceived size

We recorded SSVEP responses to flickering stimuli with 10 different retinal sizes, at two distances and at different locations in the visual field (see Table 1). As SSVEPs to stimuli in the center were much stronger than to those in the periphery (Figure 4 top), we did a z-score transformation to normalize the SSVEP amplitudes separately for each part of the visual field. Figure 4 shows a strong increase in SSVEP amplitude with retinal size. If the response depended only on retinal size, the response curves at the two distances should be identical. This was not the case. The red curve is consistently above the gray curve, indicating that more distant objects lead to a stronger response. These more distant objects are also typically perceived to be larger, in line with their larger physical size. If size constancy were complete and the SSVEP response depended only on physical or perceived size, then the

red curve should be identical to the blue curve in Figure 4. This is also not the case. The SSVEP response falls somewhere between the two predictions and thus indicates an intermediate degree of size constancy.

These results were reflected in the statistical analysis. A 2 (distances) \times 3 (visual field locations) \times 10 (retinal sizes) analysis of variance (ANOVA) revealed a significant main effect of retinal size, $F(9, 63) = 157.49$, $p < 0.001$, $\eta_p^2 = 0.96$, indicating a strong dependence of SSVEPs on retinal size. There was also a significant main effect of distance, $F(1, 7) = 6.74$, $p = 0.036$, $\eta_p^2 = 0.49$. SSVEPs for stimuli at further distances (physically and perceptually bigger) were larger than SSVEPs to stimuli at near distance (physically and perceptually smaller), though the size of their retinal projections were matched. This result suggests that SSVEPs provide a neural correlate of size constancy instead of reflecting retinal stimulation per se. In Figure 4, it seems that size constancy was less pronounced in the left visual field, but there was no significant interaction between visual field and distance, $F(2, 14) = 2.11$, $p = 0.16$, $\eta_p^2 = 0.23$, or between visual field, distance, and retinal sizes, $F(18, 126) = 0.71$, $p = 0.62$, $\eta_p^2 = 0.09$.

Experiment 2: SSVEPs represent perceived size independent of flicker frequency

The results of our first experiment indicate some degree of size constancy in visual cortex. Since the largest activity was at the central occipital electrode, we are quite confident that this mainly reflects activity in early areas (V1, V2) (Di Russo et al., 2007; Müller et al., 1997; Wittevrongel et al., 2018). However, the relatively low flicker frequency of 8 Hz that we used could potentially have stimulated neurons in higher cortical areas, where eventually a size-constant representation would be reached. Due to the progressive averaging and low-pass filtering in the visual pathway (e.g., Hawken et al., 1996; Lennie, 1998), we hypothesized that the use of a much higher temporal frequency might reduce the contributions from higher cortical areas. We therefore compared responses to 8 Hz and 30 Hz flicker. If the size constancy seen in Experiment 1 was mainly due to responses in higher visual cortical areas, it should be diminished with the 30 Hz flickering stimuli.

Figure 5 shows that this was not the case. The difference in responses to near and far stimuli was quite similar for both frequencies. A 2 (distance) \times 2 (stimulus frequency) \times 7 (retinal sizes) ANOVA revealed a main effect of retinal sizes, $F(6, 42) = 49.89$, $p < 0.001$, $\eta_p^2 = 0.88$, and a main effect of distance, $F(1, 7) = 37.24$, $p < 0.001$, $\eta_p^2 = 0.84$, replicating the result in Experiment 1. Stimulus frequency did not interact with

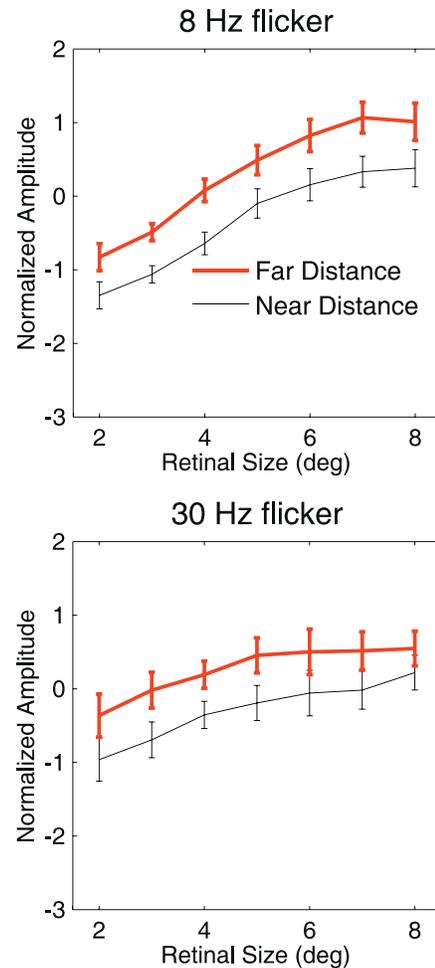


Figure 5. Normalized SSVEP amplitudes recorded in Experiment 2 shown separately for 8 Hz and 30 Hz stimuli at both distances. Error bars represent within-participant 95% confidence intervals (Cousineau, 2005). Note that the y-axes are scaled to be the same as in Figure 4 to facilitate comparisons between them.

distance ($p = 0.70$), suggesting that SSVEPs represent size constancy independently of stimulus frequency.

Experiment 3: SSVEPs code perceived size in virtual 3D environment

In the experiments so far, we used real stimuli with a rich variety of available depth cues. Under these conditions perceptual size constancy should be close to perfect (Holway & Boring, 1941; Stanley & Furedy, 1966). Since we wanted to investigate the relationship between perceptual size constancy and the degree of neuronal size constancy in the early visual system, we needed to be able to manipulate the degree of perceptual constancy. VR is ideally suited for such manipulations of different visual cues.

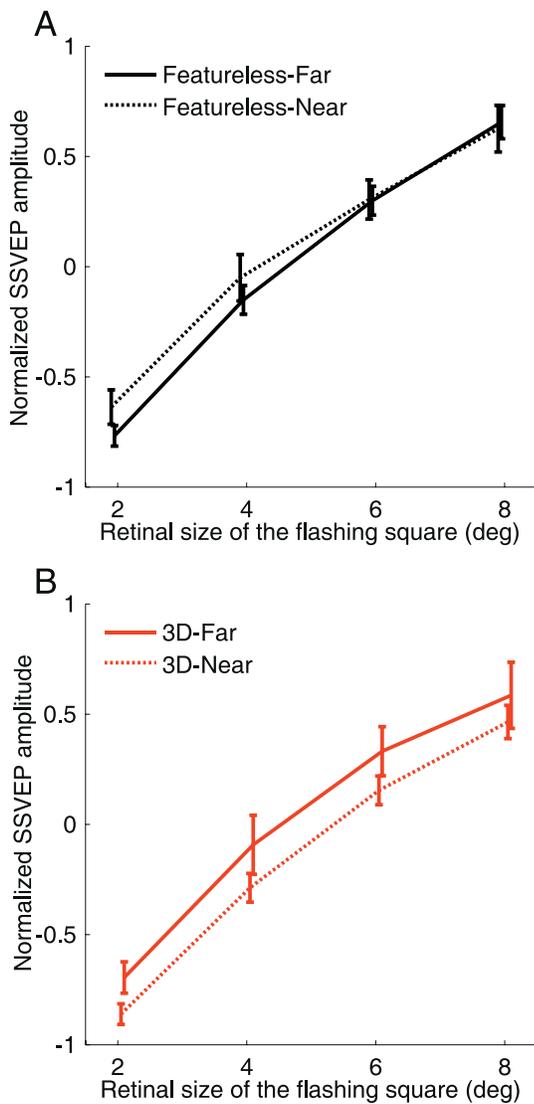


Figure 6. Normalized SSVEP amplitudes recorded in Experiment 3. SSVEP responses depended on retinal size (horizontal axis), distance (dotted vs solid lines), as well as the environment in a virtual reality setup (A: in a featureless environment; B: in a 3D hallway). Error bars represent within-participant 95% confidence intervals (Cousineau, 2005).

Using a similar design as in Experiments 1 and 2, we contrasted SSVEP responses in a virtual 3D hallway environment against the SSVEPs in a featureless control condition (Figure 2b, Table 1). As before, SSVEP responses increased with retinal size under all conditions. In the virtual hallway, where many 3D cues were available, SSVEP responses were larger for distant objects (Figure 6). For the featureless environment, size constancy was greatly diminished, with the SSVEP amplitude to the near stimuli being actually slightly larger than that to the far stimuli. A 2 (environment) \times 2 (distance) \times 4 (retinal size) ANOVA showed a main effect of retinal size, $F(3, 57) = 109.18, p < 0.001, \eta_p^2 = 0.85$. Crucially, we observed an interaction between

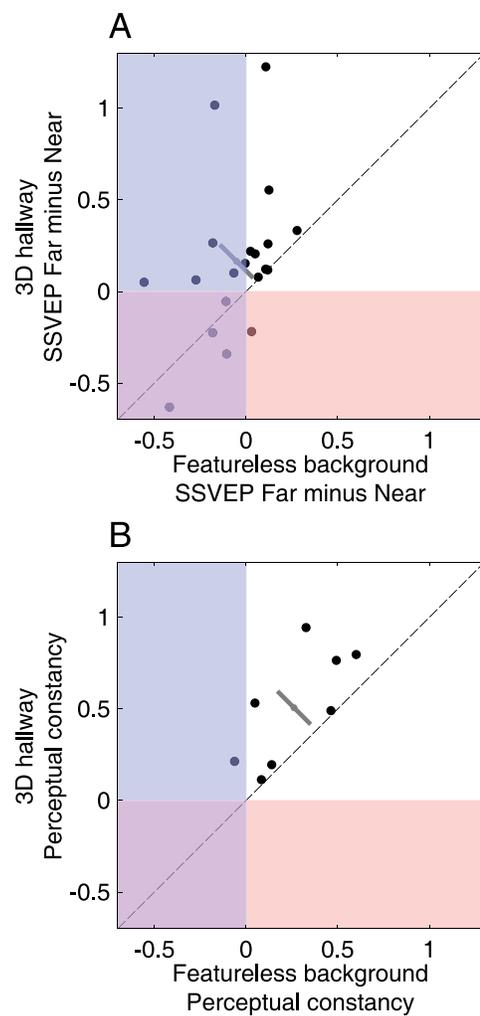


Figure 7. (A) The size constancy effect on SSVEPs (SSVEPs to far stimuli minus SSVEPs to near stimuli) is plotted for the 3D hallway environment against the same effect for SSVEPs measured in the featureless background condition. (B) The perceptual constancy index is plotted for the 3D VR environment against the perceptual index in the featureless background condition, which was measured in the perceptual adjustment task for eight observers. Error bar shows the 95% confidence interval of the difference between two conditions.

environment and distance, $F(1, 19) = 6.17, p = 0.023, \eta_p^2 = 0.25$ (Figure 6). This shows that it is indeed the visual cues available in a 3D environment, rather than other factors such as eye vergence, that lead to the size constancy responses we observed.

Figure 7A shows the effect of distance for each participant (SSVEP amplitude in response to far stimuli minus SSVEP amplitude to near stimuli) for the 3D hallway environment, plotted as a function of the responses in the featureless control condition. Most data points fall above the diagonal line, showing an enhanced size constancy response in the 3D hallway condition. This is reflected in the statistical analysis.

The effect of distance was significantly larger ($M = 0.203$, $SD = 0.419$) in the 3D environment than in the featureless control ($M = -0.052$, $SD = 0.165$), $t(19) = 2.48$, $p = 0.023$ (two tailed), Cohen's $d_z = 0.56$. Overall, these results suggest that it was indeed the distance cues that induced the size constancy responses in SSVEPs.

Experiment 4: Perceptual size constancy

Although our neural data indicate some degree of size constancy in early visual cortex, the degree of such constancy has to be evaluated with respect to perceptual measures. We therefore tested the degree of perceptual size constancy in all our settings. In the real-world environment, four observers adjusted the size of a flickering filled circle at one distance (e.g., 40 cm) to match the physical size of another filled circle at the other distance (e.g., 80 cm). Their size constancy (set size/actual size of the reference) was 95.9% ($SE = 2.9%$): close to perfect. In our setup, this confirmed the classical result that size constancy in the real world is nearly perfect (Holway & Boring, 1941; Stanley & Furedy, 1966). For the virtual environment, eight observers did the same size adjustment task, where they matched object sizes at two simulated distances (40 cm and 80 cm) within the same VR environments as Experiment 3. They showed a size constancy index of 50.7% ($SE = 11.0%$) in the 3D hallway environment and 26.2% ($SE = 9%$) in the featureless background condition, with the constancy in 3D hallway condition significantly larger than that in the featureless condition, $t(7) = 3.19$, $p = 0.015$, Cohen's $d_z = 1.13$ (Figure 7B). Participants still showed a fair amount of size constancy (26.2%) in the featureless condition, which might be because that they took usage of the relative size of the black board (Figure 2B) for the adjustment task.

Comparison of perceptual and neurophysiological indices of size constancy

As SSVEP responses were related to both retinal and perceived size for identical retinal inputs (Figures 4 through 6), we can obtain an estimate to what extent SSVEPs code the perceived size by calculating the neural SSVEP size constancy index using averaged and normalized SSVEP amplitudes. In the real-world setup (Experiments 1 and 2), we combined the data from the central visual field in Experiment 1 and the data with 8 Hz stimuli in Experiment 2. This resulted in an estimated SSVEP size constancy index of 47.6%. In the VR setup in Experiment 3, the SSVEP size constancy index was 27.5% (3D hallway) and $-18.6%$ (featureless background). Figure 8 summarizes the size constancy

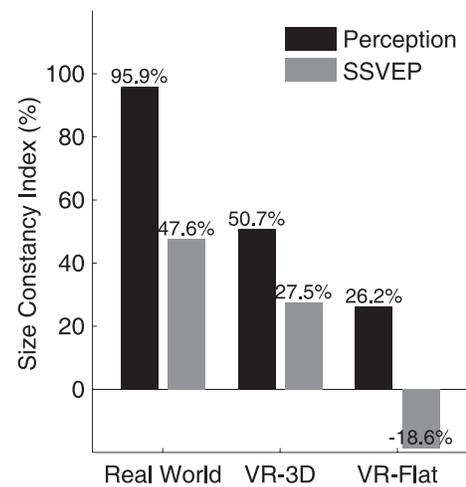


Figure 8. Size constancy indices calculated from the perceptual adjustment task and SSVEP responses for the real-world setup (Experiments 1 and 2, left two bars) and VR 3D environment (middle two bars) and flat featureless condition (right two bars). No error measures are provided for SSVEP index, as the index is calculated using averaged data of all subjects.

indices calculated from perception and SSVEPs. Size constancy indexed by SSVEP responses cannot fully account for the perceptual size constancy in either the real-world setup or the VR environment. Instead, in both environments the SSVEP index accounted for, at most, about half of the perceptual size constancy (49.6% in the real world; 54.2% in the 3D VR environment; 0% in featureless flat environment, since in the featureless flat environment the SSVEP effect was in fact in the opposite direction as the perceptual effect, which means that the SSVEP result could not explain any perceptual effect).

SSVEP phase is not affected by stimulus size or distance

Analyses aforementioned concerned only the amplitude of SSVEPs. In this section we examine the phase of SSVEP responses. We focus on the data from central visual field in Experiment 1, and from Experiment 2, as the SSVEP amplitudes recorded in these experiments showed the highest level of size constancy. In Experiment 1, we carried out a 2 (distance) \times 10 (retinal size) ANOVA on SSVEP phases, and did not find a main effect of distance, $F(1, 7) = 0.096$, $p = 0.77$; retinal size, $F(9, 63) = 0.065$, $p = 0.86$; or any interaction, $F(9, 63) = 0.56$, $p = 0.55$. The overall average phase was 15.0° ($SD = 16.0^\circ$) for stimuli at the near distance and 16.4° ($SD = 20.2^\circ$) for stimuli at the far distance. In Experiment 2, a 2 (distance) \times 2 (frequency) \times 7 (retinal size) ANOVA showed that there was no main effect of distance, $F(1, 7) = 4.94$, $p = 0.062$; retinal size, $F(6, 42) = 1.90$, $p =$

0.19, or any interaction, $F(6, 42) = 1.26$, $p = 0.31$. For 8 Hz flicker, the average phase was 27.8° ($SD = 15.9^\circ$) for near and 10.2° ($SD = 27.7^\circ$) for far distances. For 30 Hz flicker, the average phase was 92.5° ($SD = 49.3^\circ$) for near and 89.0° ($SD = 50.3^\circ$) for far distances. Overall, our result suggests that SSVEP phase is not significantly affected by stimulus size or distance.

Discussion

We measured SSVEP responses to stimuli with various retinal sizes located at two distances (near: 40 cm; far: 80 cm). When stimuli at two distances were matched in retinal size, the farther stimuli (physically and perceptually larger) consistently evoked larger SSVEP responses than the near stimuli (physically and perceptually smaller). The effect was independent of stimulus flicker frequency and held in both a real-world setup and a virtual reality environment. Using size constancy indices, we found that SSVEP responses could account for at most half of perceptual size constancy.

Effect of frequency

We used 5 Hz (Experiment 3), 8 Hz (Experiments 1 and 2), as well as 30 Hz (Experiment 2) flickering stimuli to evoke SSVEPs. The 30 Hz stimulus was expected to lead to phasic responses mainly in low-level visual areas such as V1, due to the progressive averaging and low-pass filtering along the visual pathway (e.g., Hawken et al., 1996; Lennie, 1998). In addition, previous SSVEP studies showed that higher frequency inputs are not reflected in higher levels of visual processing, such as face perception or exacting emotional content from images (Alonso-Prieto, Van Belle, Liu-Shuang, Norcia, & Rossion, 2013; Bekthereva & Müller, 2015).

However, we observed size constancy responses, i.e., larger SSVEPs for distant stimuli than near stimuli, despite identical retinal sizes for all the stimulus frequencies we used. It seems that it is not necessary for inputs to go beyond early visual cortex to be modulated by size constancy mechanisms. This result is consistent with the idea that low-level visual cortex such as V1 is an important site for size constancy coding.

Neural correlates of size constancy

Our result is consistent with previous fMRI studies showing that V1 activity changes with perceived size, when retinal size remains identical. Previous studies

investigated the neural basis of size constancy with illusory size induced by pictorial cues (Murray et al., 2006; Fang et al., 2008; Liu, Wu, Yang, Campos, Zhang, & Sun, 2009; Ni et al., 2014; He et al., 2015), size adaptation (Pooresmaeili, Arrighi, Biagi, & Morrone, 2013), and with afterimages projected at different distances (Sperandio et al., 2012). We displayed real stimuli, rather than afterimages, at different physical distances. The major advantage of our approach is that it results in close-to-perfect size constancy (see the results of experiment 4) and that identical stimuli can be used to measure the magnitude of the cortical correction for size constancy. Our results showed that neural responses measured by SSVEPs increased for stimuli of the same retinal image size when viewed at far distances, compared to stimuli viewed at near distances. Although EEGs do not provide the spatial resolution to measure the eccentricity of the activity, this result is consistent with that of Sperandio et al. (2012) and other studies (Murray et al., 2006; Fang et al., 2008) in revealing stronger neural responses to perceptually larger stimuli in the early visual system.

To what extent does the early visual system support our perceptual constancy? Does it respond to perceived size alone, thus accounting for 100% of perceptual size constancy? Alternatively, does the early visual system represent a mixture of perceived and retinal size? Some results from previous studies may give a hint. In the Ponzo illusion, previous studies measured the perceived size change induced by pictorial distance cues and the resulted neural activity change in V1. By comparing the V1 neural effect against the perceptual effect, one gets an idea of whether V1 supports size constancy, at least in the context of the Ponzo illusion.

Murray et al. (2006) reported comparable effects for perception and BOLD response, at about 20% of an effect for both. However, they used a variety of perceptual measures in supplementary experiments, showing effects between 20% and 35%. This indicates that their neuronal size constancy, relative to perception, was somewhere between $20/35 = 60\%$ and $20/20 = 100\%$. More recently, He et al. (2015) found relatively low, but comparable effects for perception (6.3%) and for V1 population receptive fields (6.8%). Ni et al. (2014) also investigated the Ponzo illusion, but looked at a corresponding shift of receptive fields in area V1 of macaque monkeys. The shift of neuronal receptive field due to Ponzo illusion was measured at 2.3%, while the perceptual effect in the same two monkeys was 5.2%, i.e., the near object needed to be 5.2% larger to be judged as the same size as the far object. Thus, the difference in V1 responses could explain only 44% of the perceptual effect, suggesting that V1 neurons may only partially account for perceptual size constancy. See Table 2 for a summary of results from these studies.

	Murray et al., 2006, human fMRI	He et al., 2015, human fMRI	Ni et al., 2014, monkey single cell recordings	Present study, SSVEP in real world	Present study, SSVEP in VR
Perceptual effect constancy	20%–35%	6.3%	5.2%	95.9%	50.7%
Neural effect constancy	20%	6.8%	2.3%	47.6%	27.5%
Percentage explained	60%–100%	100%	44%	50%	54%
Coding	Partial/Full	Full	Partial	Partial	Partial

Table 2. A summary of studies that measured perceptual and neural correlates of size constancy.

In general, the Ponzo illusion leads to striking effects on size constancy, but the corrections observed are typically less than 100% (Murray et al., 2006; Ni et al., 2014; He et al., 2015). This may be partly due to the use of monocular as opposed to stereo cues to depth. Sperandio et al. (2012) used afterimages to achieve higher, close to perfect, levels of size constancy and found corrections for size in the V1 BOLD response. However, they could not directly compare the after-image response to an equivalent response to physical stimuli of different distances and sizes. Therefore, the exact degree of corrections to the V1 retinotopy in their experiments remains unclear.

In our experiments we specifically tested how much activity in the early visual system accounts for perceptual size constancy by manipulating physical size and retinal size, and by manipulating the degree of size constancy in different experiments. Table 2 compares our results with those of these other studies.

To answer the question quantitatively, we derived an index of size constancy (Figure 3), inspired by the color constancy index (Arend et al., 1991). The size constancy index ranges from 0 (no constancy) to 100% (perfect constancy) and allows for a direct numerical comparison between perception and neural activity. The perceptual size constancy was found to be high in the real world (95.9%, $SE = 2.9\%$), but reduced in the virtual world (50.7%, $SE = 11.0\%$). This difference might be due to differences in depth cues in the virtual environments compared to real-world environments, though which cues cause the misestimation remains largely unknown (Renner, Velichkovsky, & Helmert, 2013; Carnegie & Rhee, 2015; Rousset, Bourdin, Goulon, Monnoyer, & Vercher, 2015; Langbehn et al., 2016). Further studies should evaluate to see the extent of size constancy in head-mounted VR displays. We found that the neural size constancy index was approximately half the perceptual index, both when perceptual constancy was high, i.e., in natural viewing, and when it was artificially reduced in a VR environment. This result suggests that activity in the early

visual system signaled by SSVEPs partially codes retinal inputs and partially represents perceived size.

Even if the compensation found in the early visual pathway does not fully account for perceptual size, any such correction would potentially interfere with V1 retinotopy (Horton & Hoyt, 1991; Sereno et al., 1995; for review, see Wandell et al., 2007). The representation of parts of the visual field corresponding to objects perceived as distant would become bigger, at the expense of the representation of neighboring objects, implying that the map of the visual field in V1 is much less stable than is assumed. Distant objects are often actually occluded by closer ones, so that their potentially increased neural representation may not be a problem. But conditions easily come to mind where a perfect scaling of the V1 representation would lead to topological contradictions. For example, if there were many closely spaced distant objects, their cortical representations would have to overlap, if cortical were properly scaled for distance. The partial correction of about 50% that we observed might be a compromise just to prevent or at least reduce the impact of such troubling cases.

Note that not only the magnitude, but also the shape of SSVEPs, differs from prediction under perfect size constancy. Specifically, given perfect size constancy, the slope of SSVEPs for far objects should be doubled compared with the slope for near objects. Based on our data, we cannot give a conclusive answer. The slope agreeing with the 50% correction we found would lie in between the near responses and the predicted far responses. Based on the data in Figure 4 (especially the center and right panels), the slope for the far responses might indeed be a bit higher than for the near responses. However, given the current data set, we cannot distinguish statistically between the different outcomes. The problem gets further complicated by the saturation of the neural responses at increasing retinal size, which would further reduce potential differences in slope.

What could be the neural mechanism underlying the modulation of SSVEPs in early visual cortex? Most

previous studies point to the hypothesis that alterations of neural responses in early visual cortex related to size constancy are due to feedback signals from higher visual areas. Such feedback could be used to adjust the response gain, as well as the retinotopic mapping (Fang et al., 2008; Liu et al., 2009; Ni et al., 2014; Sperandio et al., 2012). A recent preliminary report (Chen, Sperandio, Henry, & Goodale, 2018) used ERPs to examine the temporal evolution from retinal size coding to perceived size coding, and found that size constancy took at least 150 ms to emerge in the visual cortex. This result suggests that size constancy requires subsequent processing from the higher visual cortices. Our results are consistent with the notion that retinal signals in early visual cortex are modulated by information about object distance that is present in higher visual areas.

Effect of eccentricity

There is also the possibility that the correction for size constancy might be smaller in peripheral vision, thus alleviating some of the problems with topological “errors” in the cortical representation. Although we did not observe a statistically significant effect of peripheral presentation in our Experiment 1, it is apparent from Figure 4 that there was hardly any correction for distance in the left parafoveal visual field. Remarkably, size perception and size constancy in the peripheral visual field have been only explored relatively rarely, but there are reports of a decreased perceived size in peripheral vision (Baldwin, Burleigh, Pepperell, & Ruta, 2016; Valsecchi & Gegenfurtner, 2016). From Figure 4, the response to stimuli in the periphery seems to increase linearly with retinal size. Foveally, there was a saturation, presumably because the increase in stimulus size gradually becomes less magnified outside the fovea.

Role of attention and gaze direction

Attention is another factor that has been shown to affect neural response strongly (Peterson & Posner, 2012) and to modulate size constancy responses in V1 (Fang et al., 2008). Is it possible that observers paid more attention to stimuli at the far distance than the near one in our experiments, which led to the enhanced SSVEPs for the far stimuli in the present study? On one hand, it is not plausible that attention would induce such massive effects (up to ~50% in the present study), as previous studies examined attentional effect on SSVEPs generally showed an effect of 10%–20% (e.g., Andersen, Hillyard, & Müller 2013; Chen et al., 2017b). On the other hand, previous studies actually have shown that people not unreasonably tend to attend

preferentially to closer objects (Lang et al., 2012). Therefore, it is quite unlikely that our results could be explained by attention.

Another possibility is that the SSVEP responses change due to a change in the direction of eye gaze rather than in the perceived object size. It is well known that gaze direction modulates neural responses (“gain fields,” e.g., Andersen & Mountcastle, 1983). However, in our study the difference in vergence between the two viewing distances is rather small, giving rise to a change of ~2° in gaze direction of each eye. Furthermore, if you consider that vergence consists of two opposite shifts for the two eyes, the overall gain change should be really balanced. Most importantly, different neurons will be affected in different ways. Some will fire more and others will fire less. The eye position is reflected in a population code, while the overall rate of firing is not affected (e.g., Morris, Bremner, & Krekelberg, 2016). This way the aggregated response in the SSVEPs should not necessarily change. This is confirmed by the finding that accurate decoding of eye positions from fMRI BOLD signals depends on information in the spatial distributions of responses, but not in the mean amplitude (e.g., Merriam, Gardner, Movshon, & Heeger, 2013). Gaze change is therefore unlikely to explain our results.

Conclusions

Our approach using SSVEPs allowed us to introduce a new way to investigate the neural mechanisms of size constancy, one that is more amenable to use in real and immersive 3D environments compared with other techniques such as fMRI. This allowed us to show that the degree of size constancy in the early visual system is linked to, but does not fully account for, perceptual constancy. In further studies we can now isolate specific depth cues to shed light on the neural structures and the specific mechanisms responsible for generating our conscious experience of size.

Keywords: size constancy, size perception, EEG, SSVEP, virtual reality

Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 222641018 – SFB/TRR 135 Projects A1 and A8, and by the DFG International Research Training Group 1901. M. McManus held a research studentship from the NSERC CREATE program. J. Chen was supported by the Shanghai Sailing Program

(19YF1445900). All data reported in this paper are available on zenodo.org at DOI: 10.5281/zenodo.2636083.

*Jing Chen and Meaghan McManus contributed equally to this work and are shared first authors. Commercial relationships: none. Corresponding author: Karl R. Gegenfurtner. Email: gegenfurtner@uni-giessen.de. Address: Abteilung Allgemeine Psychologie, Justus-Liebig-Universität Gießen, Gießen, Germany.

References

- Alonso-Prieto, E., Van Belle, G., Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2013). The 6 Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia*, *51*(13), 2863–2875.
- Andersen, R. A., & Mountcastle, V. B. (1983). The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *Journal of Neuroscience*, *3*, 532–548.
- Andersen, S. K., Fuchs, S., & Müller, M. M. (2011). Effects of feature-selective and spatial attention at different stages of visual processing. *Journal of Cognitive Neuroscience*, *23*, 238–246.
- Andersen, S. K., Hillyard, S. A., & Müller, M. M. (2013). Global facilitation of attended features is obligatory and restricts divided attention. *Journal of Neuroscience*, *33*(46), 18200–18207.
- Andersen, S. K., Müller, M. M., & Hillyard, S. A. (2015). Attentional selection of feature conjunctions is accomplished by parallel and independent selection of single features. *Journal of Neuroscience*, *35*(27), 9912–9919.
- Arend, L. E., Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America. A, Optics and Image Science*, *8*(4), 661–672.
- Bach, M., & Meigen, T. (1999). Do's and don'ts in Fourier analysis of steady-state potentials. *Documenta Ophthalmologica*, *99*(1), 69–82.
- Baldwin, J., Burleigh, A., Pepperell, R., & Ruta, N. (2016). The perceived size and shape of objects in peripheral vision. *i-Perception*, *7*(4), 2041669516661900.
- Bekhtereva, V., & Müller, M. M. (2015). Affective facilitation of early visual cortex during rapid picture presentation at 6 and 15 Hz. *Social Cognitive and Affective Neuroscience*, *10*(12), 1623–1633.
- Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*, 257–266.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Carnegie, K., & Rhee, T. (2015). Reducing visual discomfort with HMDs using dynamic depth of field. *IEEE Computer Graphics and Applications*, *35*(5), 34–41.
- Chen, J., Sperandio, I., Henry, M. J., & Goodale, M. A. (2018). Temporal evolution from retinal image size to perceived size in human visual cortex. *bioRxiv*, 455139.
- Chen, J., Valsecchi, M., & Gegenfurtner, K. R. (2017a). Enhanced brain responses to color during smooth pursuit eye movements. *Journal of Neurophysiology*, *118*, 749–754.
- Chen, J., Valsecchi, M., & Gegenfurtner, K. R. (2017b). Attention is allocated closely ahead of the target during smooth pursuit eye movements: Evidence from EEG frequency tagging. *Neuropsychologia*, *102*, 206–216.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*, 42–45.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.
- Di Russo, F., Pitzalis, S., Aprile, T., Spitori, G., Patria, F., Stella, A., & Hillyard, S. A. (2007). Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Human Brain Mapping*, *28*(4), 323–334.
- Emmert, E. (1881). Größenverhältnisse der Nachbilder. *Klinische Monatsblätter für Augenheilkunde und für augenärztliche Fortbildung*, *19*, 443–450.
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*(2), 181–192.
- Fang, F., Boyaci, H., Kersten, D., & Murray, S. O. (2008). Attention-dependent representation of a size illusion in human V1. *Current Biology*, *18*(21), 1707–1712.
- Gregory, R. L. (2015). *Eye and brain: The psychology of seeing*. Princeton, NJ: Princeton University Press.
- Hawken, M. J., Shapley, R. M., & Gross, D. H.

- (1996). Temporal-frequency selectivity in monkey visual cortex. *Visual Neuroscience*, 13(3), 477–492.
- He, D., Mo, C., Wang, Y., & Fang, F. (2015). Position shifts of fMRI-based population receptive fields in human visual cortex induced by Ponzo illusion. *Experimental Brain Research*, 233(12), 3535–3541.
- Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *The American Journal of Psychology*, 54(1), 21–37.
- Horton, J. C., & Hoyt, W. F. (1991). The representation of the visual field in human striate cortex: A revision of the classic Holmes map. *Archives of Ophthalmology*, 109(6), 816–824.
- Kerous, B., & Liarokapis, F. (2016). Brain-computer interfaces-a survey on interactive virtual environments. In *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)* (pp. 1–4). IEEE.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1–16.
- Lang, C., Nguyen, T. V., Katti, H., Yadati, K., Kankanhalli, M., & Yan, S. (2012). Depth matters: Influence of depth cues on visual saliency. In *European conference on computer vision* (pp. 101–115). Berlin, Heidelberg: Springer.
- Langbehn, E., Raupp, T., Bruder, G., Steinicke, F., Bolte, B., & Lappe, M. (2016). Visual blur in immersive virtual environments: Does depth of field or motion blur affect distance and speed estimation? In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (pp. 241–250). ACM.
- Lennie, P. (1998). Single units and visual cortical organization. *Perception*, 27(8), 889–935.
- Liu, Q., Wu, Y., Yang, Q., Campos, J. L., Zhang, Q., & Sun, H. J. (2009). Neural correlates of size illusions: An event-related potential study. *NeuroReport*, 20(8), 809–814.
- Liu-Shuang, J., Torfs, K., & Rossion, B. (2016). An objective electrophysiological marker of face individualisation impairment in acquired prosopagnosia with fast periodic visual stimulation. *Neuropsychologia*, 83, 100–113.
- Merriam, E. P., Gardner, J. L., Movshon, J. A., & Heeger, D. J. (2013). Modulation of visual responses by gaze direction in human visual cortex. *Journal of Neuroscience*, 33, 9879–9889.
- Morris, A. P., Bremmer, F., & Kregelberg, B. (2016). The dorsal visual system predicts future and remembers past eye position. *Frontiers in Systems Neuroscience*, 10, 9.
- Müller, M. M., Teder, W., & Hillyard, S. A. (1997). Magnetoencephalographic recording of steady state visual evoked cortical activity. *Brain Topography*, 9(3), 163–168.
- Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, 9(3), 429–434.
- Ni, A. M., Murray, S. O., & Horwitz, G. D. (2014). Object-centered shifts of receptive field positions in monkey primary visual cortex. *Current Biology*, 24(14), 1653–1658.
- Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottureau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, 15(6):4, 1–46, <https://doi.org/10.1167/15.6.4>. [PubMed] [Article]
- Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35, 73–89.
- Pooresmaeili, A., Arrighi, R., Biagi, L., & Morrone, M. C. (2013). Blood oxygen level-dependent activation of the primary visual cortex predicts size adaptation illusion. *Journal of Neuroscience*, 33(40), 15999–16008.
- Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, 46(2), 23.
- Rousset, T., Bourdin, C., Goulon, C., Monnoyer, J., & Vercher, J. L. (2015). Does virtual reality affect visual perception of egocentric distance? In *Virtual Reality (VR), 2015 IEEE* (pp. 277–278). IEEE.
- Royer, A. S., Doud, A. J., Rose, M. L., & He, B. (2010). EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18, 581–589.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., ... Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212), 889–893.
- Sperandio, I., Chouinard, P. A., & Goodale, M. A. (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience*, 15(4), 540–542.
- Stanley, G., & Furedy, J. J. (1966). Size constancy and Emmert's law of apparent sizes. *Australian Journal of Psychology*, 18(3), 266–270.
- Tootell, R. B., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy

- to recognition: fMRI in human visual cortex. *Trends in Cognitive Sciences*, 2(5), 174–183.
- Valsecchi, M., & Gegenfurtner, K. R. (2016). Dynamic re-calibration of perceived size in fovea and periphery through predictable size changes. *Current Biology*, 26(1), 59–63.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2), 366–383.
- Wittevrongel, B., Khachatryan, E., Hnazaee, M. F., Carrette, E., De Taeye, L., Meurs, A., . . . Van Hulle, M. M. (2018). Representation of steady-state visual evoked potentials elicited by luminance flicker in human occipital cortex: An electrocortigraphy study. *NeuroImage*, 175, 315–326.