Behavioral/Cognitive

Visual Size Processing in Early Visual Cortex Follows Lateral Occipital Cortex Involvement

ⒾHang Zeng,[1] ⒾGereon R. Fink,[1,2] and ⒾRalph Weidner[1]

[1]Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3), Research Center Jülich, Jülich 52425, Germany, and [2]Department of Neurology, University Hospital Cologne, Cologne University, Cologne 50937, Germany

Neural activation in the early visual cortex (EVC) reflects the perceived rather than retinal size of stimuli, suggesting that feedback possibly from extrastriate regions modulates retinal size information in EVC. Meanwhile, the lateral occipital cortex (LOC) has been suggested to be critically involved in object size processing. To test for the potential contributions of feedback modulations on size representations in EVC, we investigated the dynamics of relevant processes using transcranial magnetic stimulation (TMS). Specifically, we briefly disrupted the neural activity of EVC and LOC at early, intermediate, and late time windows while participants performed size judgment tasks in either an illusory or neutral context. TMS over EVC and LOC allowed determining whether these two brain regions are relevant for generating phenomenological size impressions. Furthermore, the temporal order of TMS effects allowed inferences on the dynamics of information exchange between the two areas. Particularly, if feedback signals from LOC to EVC are crucial for generating altered size representations in EVC, then TMS effects over EVC should be observed simultaneously or later than the effects following LOC stimulation. The data from 20 humans (13 females) revealed that TMS over both EVC and LOC impaired illusory size perception. However, the strongest effects of TMS applied over EVC occurred later than those of LOC, supporting a functionally relevant feedback modulation from LOC to EVC for scaling size information. Our results suggest that context integration and the concomitant change of perceived size require LOC and result in modulating representations in EVC via recurrent processing.

Key words: objects' size information; TMS; visual processing

Significance Statement

How we perceive an object's size is not entirely determined by its physical size or the size of its retinal representation but also the spatial context. Using transcranial magnetic stimulation, we investigated the role of the early visual cortex (EVC) and the higher-level visual area, lateral occipital cortex (LOC), known to be critically involved in object processing, in transforming an initial retinal representation into one that reflects perceived size. Transcranial magnetic stimulation altered size perception earlier over LOC compared with EVC, suggesting that context integration and the concomitant change in perceived size representations in EVC rely on feedback from LOC.

Introduction

Size representations are coded by different stages of the visual hierarchical stream. More specifically, early cortical visual areas are known to represent the visual field in a retinotopic fashion preserving spatial arrangements as coded on the retina (Wandell et al., 2007). Higher visual areas represent size information in a more abstract sense (e.g., reflecting objects' real-world canonical sizes independent of retinal size variations) (Konkle and Oliva, 2012). This functional organization suggests feedforward processing in the visual system, where incoming information is gradually analyzed while ascending through hierarchies of the visual system, eventually resulting in a complex percept.

However, according to various psychological and neurobiological models (Bullier, 2001; Lamme, 2001; Pollen, 2003), visual information processing might also involve feedback from higher to lower visual areas. It was demonstrated that feedback from higher visual regions alters response characteristics in lower regions (Lamme, 1995), hence changing representations in the early visual cortex (EVC) in the course of visual processing. For instance, several functional imaging studies have demonstrated that V1 not rigidly represents retinotopic size but that size representations are modulated by context and thereby code the perceived size (S. O. Murray et al., 2006; Fang et al., 2008; Sperandio et al., 2012), possibly via receptive field shifts (Ni et al., 2014; He

et al., 2015). Accordingly, feedforward signals (e.g., a bottom-up retinal size signal) can be transformed by feedback, most likely from extrastriate regions, hence generating a neural code that represents perceived rather than retinal size. Consistently, a study using EEG revealed that the amplitude of the visual P2 component is affected by the illusory size, suggesting that EVC is indeed sensitive to size illusions and that this sensitivity is related to later processing stages (Liu et al., 2009). Similarly, a recent EEG study investigated the neural mechanisms underlying size constancy and found that neural activity in the visual cortex reflected retinal size merely before 150 ms. Only afterward, the EEG signal reflected size information independent of distance (Chen et al., 2019). Although ample evidence suggests that the feedback from high-level visual areas to EVC modulates size representations, little is known regarding the underlying neural mechanisms. A brain area that is possibly responsible for the feedback modulation concerning perceived size is the lateral occipital cortex (LOC). It is known to be critically involved in object perception and particularly in object size perception as indicated, for example, by patient studies showing that lesions of LOC cause hemimicropsia, an apparent reduction of the size of an object when presented in one hemifield (Cohen et al., 1994). In addition, an fMRI study using the Müller-Lyer illusion showed that neural activation in LOC is sensitive to the degree of size scaling involved (Weidner and Fink, 2007).

In the present study, we aimed to investigate the causal and temporal relationships between size representations in EVC and LOC using transcranial magnetic stimulation (TMS). We briefly disrupted ongoing neural activity in EVC and higher-level visual area LOC in an early, intermediate, and late time window while participants performed a size judgment task. If feedback signals from LOC to EVC are crucial for generating perceived size and hence altered size representations in EVC, then a specific temporal pattern of TMS effects for the two regions is expected. If LOC constitutes the early essential node for generating perceived size, then this should be the region that is affected first by TMS; therefore, TMS stimulation should decrease size illusion effects relatively early. Similarly, if feedback information from LOC determines illusion-related size processing in EVC, then the effects of TMS applied over EVC on illusion-related effects of size processing are expected to occur later than those observed following TMS over LOC. Overall, we expected that context integration and concomitant changes in perceived size occur in LOC, which then provides feedback modulation on the processing in the EVC.

Materials and Methods

Participants
Twenty-six participants were enrolled in the current experiment. Five of them were excluded because they could not experience any phosphenes (see details in the TMS protocol part). One participant was excluded because the algorithm that was used to determine the perceived size (see details in the procedure part) failed to converge in more than one-third of all blocks. Therefore, the data of 20 participants were used for further analyses (7 female, 13 male, mean ± SD age: 29.79 ± 2.56 years, age range: 26–37 years). The sample size was chosen based on previous visual TMS studies (Mancini et al., 2011; Li et al., 2019). None of the participants reported a history of neurologic or psychiatric disorders. All participants included had normal or corrected-to-normal vision. Written informed consent was obtained before the experiment following the Declaration of Helsinki. Participants were remunerated for their time. The ethics committee of the German Society of Psychology had approved the study.

Stimuli
Two different types of scene background images were used, and each type was associated with a specific experimental condition. In an illusion condition, the corresponding background image depicted a hallway scene consisting of a straight hallway made up of two brick sidewalls and a brick floor inducing the Ponzo illusion, a visual size illusion that is based on visual perspective (Fig. 1A). In the baseline condition, the background image consisted of phase-scrambled versions of the hallway scene. Phase-scrambling removed any perspective information from the hallway scene and was hypothesized to remove the illusion effect. In particular, four different phase-scrambled versions of the hallway scene background image were generated and used as neutral background images where no Ponzo illusion effect was expected (Fig. 1B). All the low-level stimulus attributes, such as luminance, contrast, and spatial frequency, were matched for the background stimuli using the SHINE toolbox (Willenbockel et al., 2010). In both conditions, two ellipses (3.25 cd/m^2) were embedded and located at 25% and 75%, respectively, of the screen height from the bottom of the screen (corresponding to 7° of visual angle away from the center of the screen). The lower ellipse was defined as the standard stimulus, and its size was kept constant throughout the experiment (11.2° of visual angle in width, 2.3° of visual angle in height). The upper ellipse served as the probe with its size changed based on participants' responses, which in the end matched the perceived size of the standard ellipse. Its initial size randomized and ranged from 87.5%, 90%, 93.75%, 106.25%, 110%, to 112.5%, relative to the standard ellipse. A fixation point (1.8° of visual angle) positioned at either the left or right side of the screen (5.6° of visual angle away from central) was added to the display. The location of the fixation point varied across participants only and was hence constant for each participant throughout the experiment (see details in the TMS protocol part). In general, the stimuli setup was identical for both the illusion and the baseline condition, with the only difference being the scene background.

Stimuli were presented on a 22-inch SyncMaster 2233RZ (Samsung Electronics) LCD screen at a distance of 57 cm. The resolution of the screen was 1680 × 1050 pixels with a refresh rate of 60 Hz. This monitor was shown to be suitable for visual research with sufficiently precise timing (Wang and Nikolić, 2011). The distance was preserved by a chin and forehead rest. Stimuli were presented with Presentation software (Neurobehavioral Systems).

Experimental design and statistical analyses
Experimental design and procedure
To investigate the roles of EVC and LOC in illusory size rescaling, we applied TMS over EVC, LOC, or the vertex at three different time intervals after stimulus onset (for details, see TMS protocol). Stimulation sessions for each TMS location comprised two parts: an illusion part (i.e., illusion condition with the hallway background) and a baseline part (i.e., baseline condition with the phase-scramble neutral background). For both the illusion and the baseline part, a series of six blocks was implemented for each time interval condition. This was done for each stimulation site. The baseline part was conducted to provide baseline size estimates for each participant, which were later used for baseline correction.

The experiment started with one practice block excluding TMS pulses, followed by another practice block with TMS pulses included, so that participants could get accustomed to the task as well as to the TMS pulses. Each block started with a 1000 ms presentation of one of the two scene backgrounds and a fixation point, which was either presented at the left or right side of the screen (Fig. 1A,B). Participants were instructed to keep their eyes at the fixation point and to attend the two ellipses at the same time. Background scene and fixation were present throughout the whole length of each block. Each trial started with a 100 ms (6 frames) presentation of both ellipses and disappeared for 2000 ms, during which responses were recorded (Fig. 1A). Participants' task was to indicate the (apparently) larger of the two ellipses via button press using their right hand. A response with the right index finger indicated the upper ellipse to appear larger, whereas a response with the right thumb indicated the lower of the two ellipses to appear larger. The locations of the stimuli corresponded to the physical arrangement

of the response buttons; that is, participants pressed the upper button to indicate the upper ellipse as larger and the lower button to indicate the lower ellipse as larger. When participants failed to perceive a size difference between the two ellipses, they were instructed to guess (two-alternative forced-choice task). The size of the upper ellipse (i.e., the probe) was changed based on an adaptive algorithm as a function of the participants' responses. We used the original version of Parametric Estimation by Sequential Testing (PEST) developed by Taylor and Creelman (1967; see also Macmillan and Creelman, 1991). This algorithm was designed to determine the point of subjective equality, that is, the point at which no observable differences between responses to a variable stimulus (here: the probe ellipse) and a standard stimulus (here: the standard ellipse) can be detected. The PEST algorithm determines the point of subjective equality by sequentially varying the size of the probe, depending on the participant's response. In particular, size changes follow the five rules of the PEST algorithm: (1) After each reversal, the step size is halved. A reversal is defined as a step in the opposite direction from the previous step. (2) A step in the same direction as the last keeps step size unchanged, with the following exceptions. (3) A third step in the same direction calls for a doubled step size, and each successive step in the same direction is also doubled until the next reversal. This rule has its own exceptions. (4) If a reversal follows a doubling of step size, then an extra same size step is taken after the original two before doubling. (5) A maximum step size is specified, at least 8 or 16 times the size of the minimum step. For example, if the participant consistently judges the probe to be larger than the test, its size is decreased at large steps to make it more similar to the test stimulus. Steps become gradually smaller as the participants' responses approach the chance level. Once the step size falls below a predetermined value (here: one pixel on the screen), that is, once the size of the probe ellipse is assumed to reflect the apparent size of the test ellipse, the PEST algorithm converges and finishes. The exact number of trials in an algorithm block was determined by the speed at which the adaptive algorithm converged with a minimum of five trials and a maximum of 20 trials. In the present experiment, once the PEST algorithm converged, a new block was started, and a new PEST sequence was initiated.

TMS protocol

To investigate the role of EVC and LOC in scaling size information, TMS stimulation was applied over EVC and LOC in two distinct sessions (Fig. 1C), which were separated from each other by at least 24 h. Vertex stimulation was included as a control for any effects induced by pulse noise or cutaneous stimulation. Moreover, to test whether subjects per se showed any systematic effects across the two different sessions, we
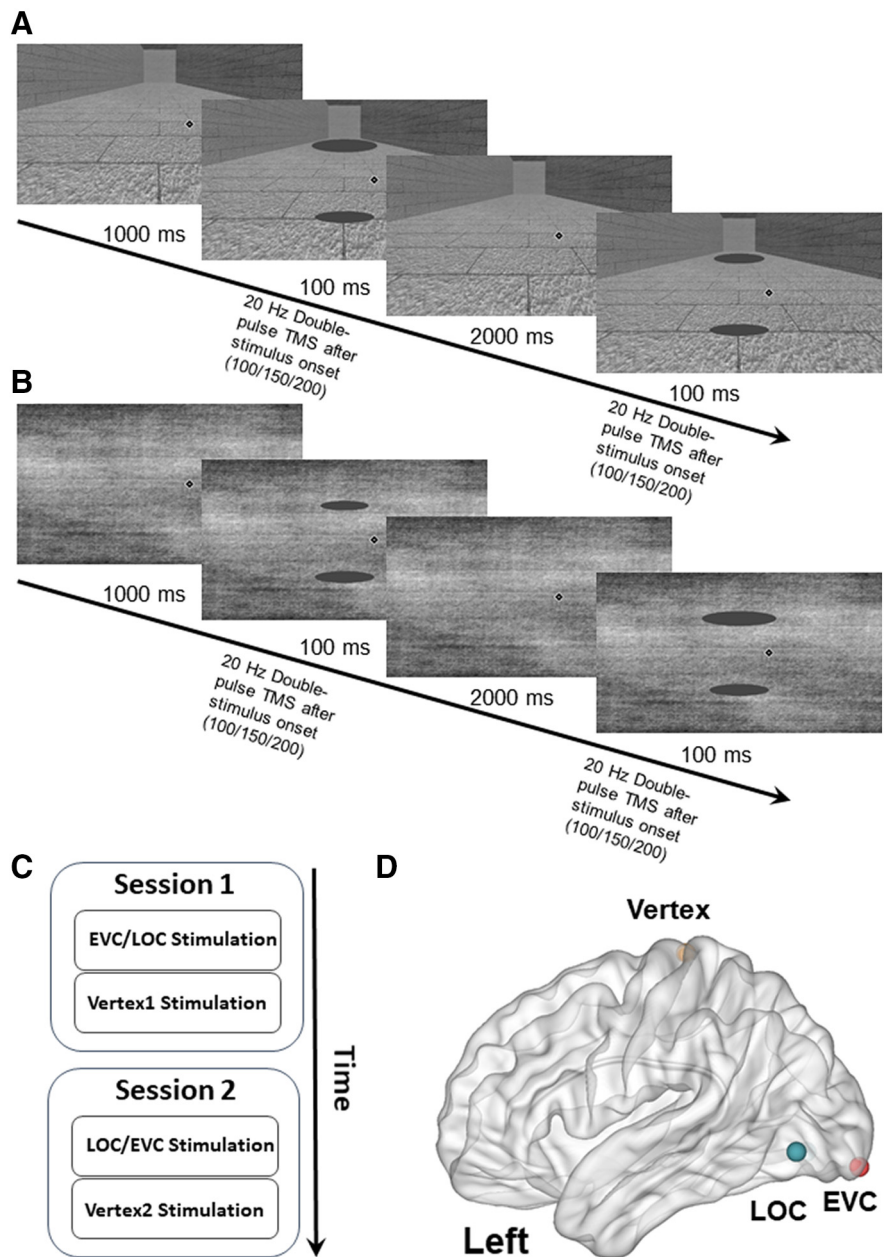

Figure 1. The experimental design. **A**, The procedure of the illusion background. Each block of the size discrimination task started with a 1000 ms presentation of the illusory backgrounds and fixation. The fixation was either presented at the left or right side of the screen. After that, each trial started with a 100 ms (6 frames) presentation of two ellipses; 20 Hz double-pulse TMS was delivered at 100/150/200 SOA. Then, the ellipse disappeared for 2000 ms during which responses were recorded. Participants' task was to indicate the (apparently) larger of the two ellipses via button press. The size of the upper ellipse (i.e., the probe) was changed based on the PEST algorithm as a function of the participants' responses, whereas the size of the lower one (i.e., the standard stimulus) was kept constant throughout the experiment. **B**, The procedure was identical for the neutral background. **C**, Participants received an EVC or LOC stimulation in the first session, followed by a vertex stimulation. After at least 24 h, LOC or EVC stimulation was delivered followed by another vertex stimulation. **D**, TMS location: EVC (red), LOC (blue), and the vertex (control site, yellow).

implemented vertex stimulation in both sessions after EVC or LOC stimulation to make it as comparable as possible across the two sessions. Specifically, in the first session, participants received TMS pulses over EVC or LOC and the vertex. In the second session, they received TMS pulses over LOC or EVC and the vertex. Half of the participants received stimulation on the left hemisphere and the other half on the right hemisphere. TMS stimulation of EVC and LOC was always applied contralaterally to the target stimuli. To this end, the fixation symbol was presented either on the left or right, while the targets were always presented centrally. Hence, the target stimuli fell into either the participant's

right or left visual field. The order of the TMS stimulation site (i.e., first EVC, then LOC or first LOC, then EVC) was randomly assigned and balanced across participants.

Before the actual experiment started, phosphene thresholds were determined for each participant. Phosphene sites were localized by applying single TMS pulses to the occipital lobe, with the participants' eyes closed. Stimulation intensity was increased from 45% of maximal stimulator output in 5% steps until a phosphene was reported or 85% maximal stimulator output was reached. As mentioned above, participants ($n = 5$) were excluded if no phosphene was reported after reaching 85% of stimulation over 5 points around the occipital pole. The phosphene threshold (50% of the pulses resulted in the perception of a phosphene) was then determined using the PEST algorithm. A stimulation intensity of 80% of the phosphene threshold was then used for all experimental sessions (mean: 40.3; range: 30-58). The exact coordinates of EVC were determined by choosing the coil location that reliably induced phosphenes. The coordinates of the left and right LOC were chosen from an fMRI study of Weidner and Fink (2007), as the local maximum of increased neural activity associated with the size illusion (MNI coordinate: left LOC, $x = -46$, $y = -74$, $z = -6$; right LOC, $x = 36$, $y = -78$, $z = -6$; Fig. 1D). The coordinates were back-normalized from the MNI template to the individual participant's brain, using the inverse normalization parameters from the normalization of each participant's brain to the MNI template. The back-normalization was performed using SPM12 (Wellcome Department of Imaging Neuroscience, London; http://www.fil.ion.ucl.ac.uk). Although EVC is located relatively close to LOC, TMS has been demonstrated to be spatially specific enough to stimulate LOC and EVC separately (Koivisto et al., 2011; Wokke et al., 2013). Vertex as a control site was identified as the point which had an equal distance to the left and the right preauricular, and at the same time an equal distance to the nasion and the inion as seen in the participants' anatomic images (Fig. 1D).

T1-weighted anatomic MRI scans for each participant were acquired using a standard T1-weighted MPRAGE sequence (voxel size of $1 \times 1 \times 1$ mm^3) with a Siemens Trio 3.0-T whole-body scanner (Siemens), and later used to perform neuro-navigated TMS over EVC, LOC, and the vertex. Double-pulse (20 Hz) stimulation was delivered by Magstim Super Rapid (Magstim) stimulators with one power supply unit attached in combination with a 70 mm figure-8 coil. The exact location of the coil was guided and continuously monitored throughout the whole experimental session via frameless stereotaxic neuronavigation system (Brainsight2, Rogue Research). The TMS coil was held manually. Hence, the coil position could be adjusted easily when head movements occurred. Two consecutive TMS pulses were delivered with an interstimulus interval of 50 ms in early (100 ms), intermediate (150 ms), and late (200 ms) time windows after display onset of the stimuli to determine the temporal sequence of contributions of EVC, LOC, and the vertex during the size judgment task. Each time window condition was presented six times (i.e., 6 PEST algorithm blocks) in both background conditions (i.e., the baseline and illusion conditions) for all TMS locations. The orders of the time window condition and the background condition were randomly assigned and counterbalanced across participants. In the current study, stimulation at each site lasted, on average, <20 min. Since we only tested two sites per session (one session per day, in total two sessions), a session lasted ~40 min.

Data analysis
The free statistical software R (R Foundation for Statistical Computing, Vienna, Austria; www.r-project.org) was used to analyze behavioral data. The dependent variables (i.e., the size of the probe stimulus in the size discrimination task across different experimental conditions) were recorded and averaged. The values of the baseline condition (i.e., the neutral, phased-scrambled background condition) served as size estimates of the standard stimulus assessing each participant's general perceptual and response biases. Illusion effects in the illusion condition were corrected, taking into account the values obtained in the baseline condition. Specifically, baseline-corrected illusion effects were calculated as a ratio of the difference between the size estimation of the illusion condition and the size estimation of the baseline condition to the size

estimation of the baseline condition, as shown in the following formula: Baseline corrected illusion effects = ((Size estimation$_{illusion}$ − Size estimation$_{baseline}$)/Size estimation$_{baseline}$) × 100.

First, the perceived size estimations in the baseline condition were tested for any difference across TMS locations and time windows. Second, the baseline-corrected illusion effects of two vertex sessions were tested for any significant difference across sessions and time windows to verify whether participants performed comparably across sessions. The results showed no significant difference across sessions. Third, the baseline-corrected illusion effects were then entered into a 2 (visual field: left and right) × 4 (TMS location: EVC, LOC, vertex1, vertex2) by 3 (time window: early, intermediate, and late) repeated-measures ANOVA to test for differences of illusion effects in all conditions. Further pairwise comparisons using Bonferroni correction (i.e., $p \times$ the number of comparisons) were performed to test for specific differences between time windows in EVC, LOC, and vertex1 and vertex2 separately (all results of pairwise comparisons were Bonferroni-corrected unless otherwise stated). We further subtracted illusion effects related to EVC and LOC stimulation from those found during the same vertex stimulation to test the TMS effects. All size judgments were made based on the participants' subjective decisions. Therefore, there are no correct or wrong responses. Hence, no accuracy data can be reported.

We expected TMS to have a differential effect at different time windows, depending on the location of stimulation. If the feedback signal is critical for processing size scaling, TMS applied over EVC should affect performance at a later time window relative to LOC stimulation.

Eye movement tracking
Eye positions were monitored using the monocular eye-tracking system EyeLink 1000 (SR Research) at a sampling rate of 1000 Hz. Eye movement data were recorded from the left eye. A 5 point calibration and validation procedure was performed to map the eye positions to the screen coordinates. Drift correction was performed before the actual main experiment. A quadratic area around the fixation, whose height and width were both 2° of visual angle, served as an ROI. The eye movement data between the appearance and the disappearance of the target stimuli (i.e., a 100 ms time window) were analyzed by calculating the percentage of time that the participant's gaze was inside the ROI while the target display was present on the screen. This allowed evaluating the quality of fixation during target presentation across all sessions.

Results
Eye movement data
Because of technical problems of the eye-tracker, eye movement data from 3 participants could not be recorded. Hence, eye movement data were available from only 17 participants. The relative amount of time that the participants maintained fixation (in percent) in the different conditions was entered into a 2 (background: illusion and baseline) × 4 (TMS location: EVC, LOC, vertex1, and vertex2) × 3 (time window: early, intermediate, and late) repeated-measures ANOVA.

On average, the participants maintained fixation on 95.20% of the target presentation time. None of the main effects of the factors was significant: background condition ($F_{(1,16)} = 1.993$, $p = 0.177$, $\eta_p^2 = 0.111$), TMS location ($F_{(3,48)} = 0.750$, $p = 0.528$, $\eta_p^2 = 0.045$), and time window ($F_{(2,32)} = 0.287$, $p = 0.753$, $\eta_p^2 = 0.017$). There was no significant two-way or three-way interaction between any of the factors (all $p > 0.05$). The eye tracking data hence indicate that participants fixated comparably well across the different conditions.

Behavioral data
First, to test whether size perception and size judgments were per se affected by TMS stimulation at different stimulation sites and different stimulation intervals, a 4 (TMS location: EVC, LOC, vertex1, and vertex2) × 3 (time window: early,
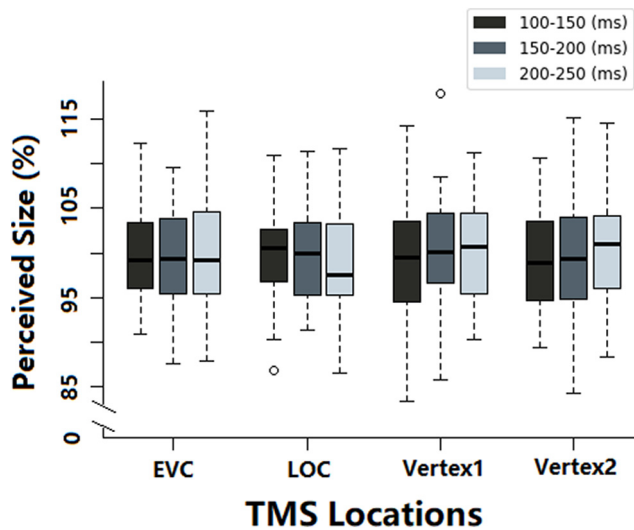
Figure 2. Box plot represents percentages of perceived size in the baseline condition under stimulations of EVC, LOC, vertex1, and vertex2. The size judgments on the perceived size in the baseline condition were not systematically altered, neither by TMS at different stimulation sites nor different time windows.

intermediate, and late) repeated-measures ANOVA was conducted using the estimates of perceived size in the baseline condition as dependent variable (Fig. 2). Neither the main effect of the TMS location ($F_{(3,57)} = 0.350$, $p = 0.789$, $\eta_P^2 = 0.018$), nor the main effect of the time window ($F_{(2,38)} = 0.590$, $p = 0.559$, $\eta_P^2 = 0.003$) was significant. There was also no significant interaction between TMS location and time window ($F_{(6,114)} = 1.306$, $p = 0.260$, $\eta_P^2 = 0.064$). These results indicate that size judgments on the perceived size in the baseline condition were not systematically altered, neither by TMS at different stimulation sites nor different time windows.

Second, to test whether participants per se showed any systematic effect across the two different sessions (Fig. 1C) concerning different time interval conditions, baseline-corrected illusion effects of the vertex stimulation in both sessions were submitted to a 2 (session: first and second) × 3 (time window: early, intermediate, and late) repeated-measures ANOVA. Neither the main effect of the session ($F_{(1,19)} = 0.977$, $p = 0.335$, $\eta_P^2 = 0.049$) nor the main effect of the time window ($F_{(2,38)} = 0.312$, $p = 0.734$, $\eta_P^2 = 0.016$) was significant. There was also no significant interaction between TMS session and time window ($F_{(2,38)} = 1.093$, $p = 0.345$, $\eta_P^2 = 0.054$). This pattern of results suggests that participants performed comparably in the two sessions and independent of the stimulation time window. Accordingly, any significant differences in EVC and LOC stimulations can be considered to be caused by specific TMS effects.

Third and most importantly, to test the effects of TMS stimulation site at different TMS time windows, baseline-corrected illusion effects were submitted to a 2 (visual field: left and right) × 4 (TMS location: EVC, LOC, vertex1, and vertex2) by 3 (time window: early, intermediate, and late) repeated-measures ANOVA (Fig. 3A). The main effect of visual field was not significant ($F_{(1,18)} = 0.440$, $p = 0.515$, $\eta_P^2 = 0.024$). There were no significant two-way or three-way interactions between visual field and other factors (all $p > 0.05$), indicating no lateralization effect. The main effect of the TMS location was significant ($F_{(3,54)} = 5.345$, $p = 0.003$, $\eta_P^2 = 0.229$). Further pairwise comparisons indicated that the illusion effect of the vertex1 (14.40%) and vertex2

(13.94%) stimulation were significantly stronger compared with stimulation over EVC (10.40%), and LOC (10.48%) (all $p < 0.005$), which suggests that TMS over EVC and LOC both impaired illusion-related processing. There was no significant difference between EVC and LOC stimulations. In addition, the main effect of the time window turned out to be significant ($F_{(2,36)} = 3.357$, $p = 0.046$, $\eta_P^2 = 0.116$). Further pairwise comparisons revealed a marginal significant reduction of illusion effects in the intermediate (11.19%) relative to the early time window (12.62%): $p = 0.123$ (before Bonferroni correction, $p = 0.041$). No other pairwise comparison revealed a significant difference.

Most importantly, the interaction between the TMS location and the time window was significant ($F_{(6,108)} = 3.803$, $p = 0.006$, $\eta_P^2 = 0.151$), suggesting that for the different stimulation locations the time interval within which the stimulation occurred was mattering (Fig. 3A). Further planned two-tailed t test with Bonferroni correction showed that when TMS was applied over EVC, illusion effects were significantly weaker in the late TMS time window (9.21%) than in the early TMS time window (11.62%): $t_{(19)} = 2.951$, $p = 0.024$, 95% CI (0.69%, 4.10%). In contrast, stimulation over LOC, resulted in significantly weaker illusion effects in the intermediate TMS time window (8.95%) compared with the late TMS time window (11.76%): $t_{(19)} = -3.337$, $p = 0.009$, 95% CI (-4.57%, -0.10%); and marginally weaker than in the early time window (10.74%): $t_{(19)} = 2.193$, $p = 0.123$ ($p = 0.041$ before Bonferroni correction), 95% CI (0.08%, 3.49%). No significant differences across the three time windows were found when stimulation was applied over the vertex1 and vertex2 (all $t < 1.358$, all $p > 0.05$). No further significant effects were found. Accordingly, the largest reduction of size scaling processing in LOC was found in a time window that was earlier than the corresponding time window for EVC.

We further subtracted illusion effects related to EVC and LOC stimulation from those found during vertex stimulation of the same session (i.e., TMS effects relative to the vertex; Fig. 3B). The effects were submitted to a 2 (TMS location: EVC and LOC) × 3 (time window: early, intermediate, and late) repeated-measures ANOVA. Neither the main effect of TMS location ($F_{(1,19)} = 0.004$, $p = 0.949$, $\eta_P^2 = 0.001$), nor the main effect of the time window was significant ($F_{(2,38)} = 2.811$, $p = 0.072$, $\eta_P^2 = 0.128$). However, the interaction between the TMS location and time window was significant ($F_{(2,38)} = 3.616$, $p = 0.037$, $\eta_P^2 = 0.160$). Furthermore, planned two-tailed t tests with Bonferroni correction revealed that, when applied over EVC, TMS reduced the illusion strength marginally significantly more in the late (5.05%) compared with the early time window (2.64%): $t_{(19)} = -2.176$, $p = 0.102$ (before Bonferroni correction, $p = 0.034$), 95% CI (-4.73%, -0.01%). In contrast, when stimulation was applied over LOC, TMS in the intermediate time window (5.52%) reduced illusion strength significantly more than in the early (3.13%) and late stimulation time window (2.67%) (all $p < 0.048$). This again confirms that the critical time window in LOC was found earlier than the one for EVC.

We further compared the TMS effects of EVC and LOC in the same time window. Since different brain regions may react differently to the TMS, we used the value of the first time window as a baseline to correct for it. We then perform a single-tailed t test to test whether the TMS effects of EVC are smaller than LOC in the intermediate time window, and greater than LOC in the late time window. The results showed that the TMS effects in EVC were not significantly smaller than LOC in the intermediate time window: $t_{(19)} = -1.165$, $p = 0.129$, 95% CI (-Inf,
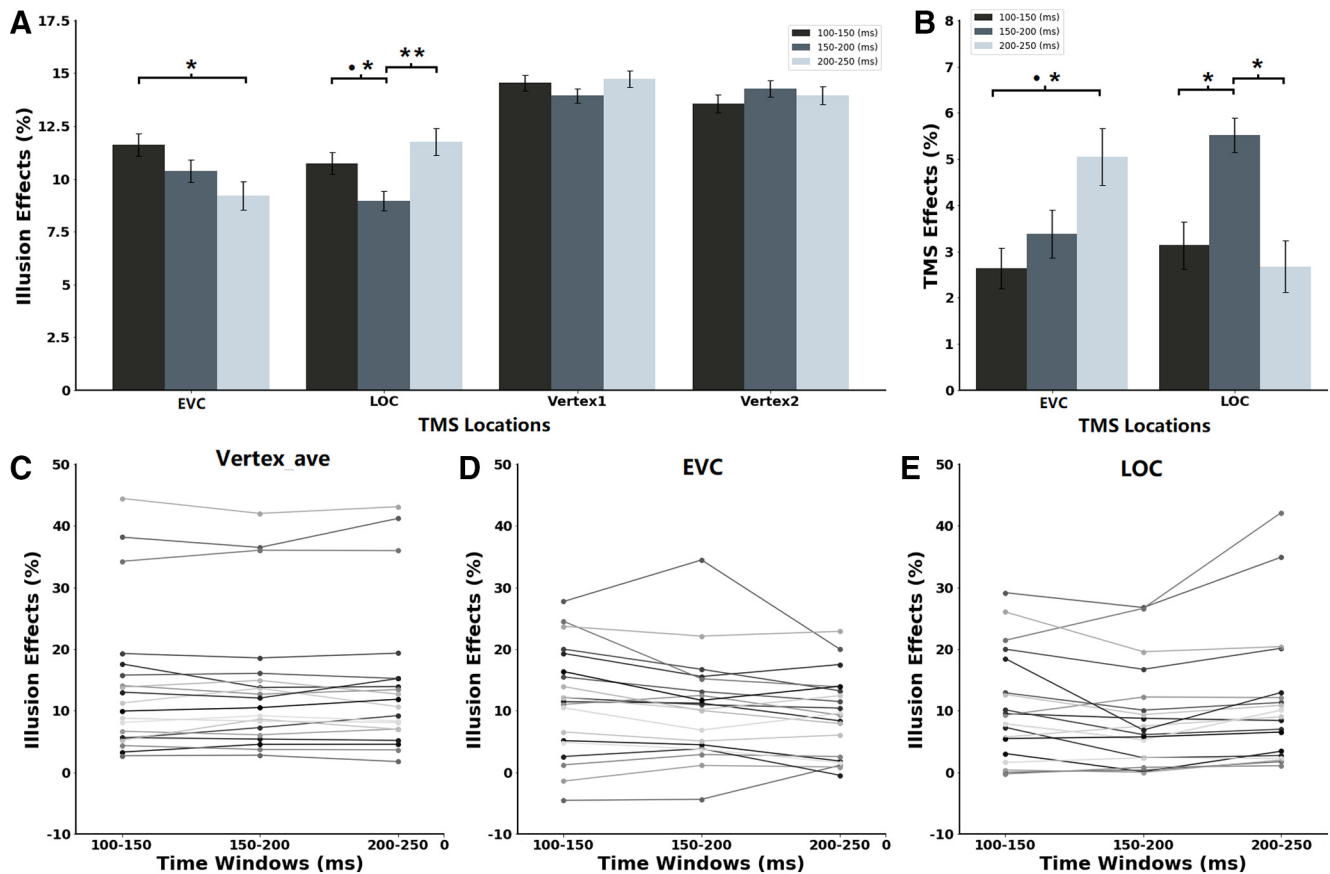
Figure 3. Behavioral results of the experiment. **A**, Percentages of baseline-corrected illusion effects of the EVC, LOC, and vertex1 and vertex2 in three time windows are plotted. EVC: 11.62%, 10.36%, and 9.21%; LOC: 10.74%, 8.95%, and 11.76%; vertex1: 14.55%, 13.93%, and 14.73%; vertex2: 13.57%, 14.28%, and 13.97%. Pairwise comparisons with Bonferroni correction revealed a significantly weaker illusion effect in the late time window than in the early one in EVC ($p = 0.024$). Meanwhile, the illusion effect was significantly weaker in the intermediate time window than in the late one in LOC ($p = 0.009$). **B**, TMS effects relative to the vertex of EVC and LOC in three time windows are plotted. EVC: 2.64%, 3.38%, and 5.05%; LOC: 3.13%, 5.52%, and 2.67%. Pairwise comparisons revealed a similar pattern as in **A**. **C**, Individual values of illusion effects in the vertex stimulation. **D**, Individual values of illusion effects in EVC stimulation. **E**, Individual values of illusion effects in LOC stimulation. ** $p < 0.01$; * $p < 0.05$; •* $p < 0.05$; before Bonferroni correction for multiple comparisons. Error bars indicate $\pm$ within-subject SEM. Method from O'Brien and Cousineau (2015).

0.7%); but significantly stronger than LOC in the later time window: $t_{(19)} = 4.721$, $p < 0.001$, 95% CI (7.6%, Inf).

Given the variability in individual participants, we additionally used the lme4 package in R (Bates et al., 2015) to perform a linear mixed-effects analysis to investigate the data. As fixed effects, we entered the visual field, TMS location, and time window (with interaction term) into the model. As random effects, we had intercepts for subjects, as well as by-subject random slopes for the effect of TMS locations and time windows. The results showed a similar pattern. Specifically, the main effect of the visual fields was not significant ($F_{(1,18)} = 0.440$, $p = 0.515$). The main effect of the TMS location was significant ($F_{(3,198)} = 5.345$, $p = 0.002$). The main effect of the time window was not significant ($F_{(2,198)} = 2.000$, $p = 0.139$). Most importantly, the interaction between TMS location and time window was still significant ($F_{(6,198)} = 3.645$, $p = 0.002$).

Discussion

Both EVC and LOC have previously been suggested to constitute core regions of the network underlying object size perception (S. O. Murray et al., 2006; Konkle and Oliva, 2012). Importantly, however, it was hitherto unknown how these two regions jointly process object size information. The goal of the present study was, therefore, to investigate the temporal dynamics of the relevant contributions provided by the two regions. Of particular

interest was the question of whether size representations reflecting perceived size in EVC are formed based on feedback signals from LOC.

We used double-pulse TMS to briefly interfere with neural activity related to object size perception in EVC and LOC. Double-pulse TMS creates a broader time window of disruption while maintaining the high temporal resolution associated with single-pulse TMS. In the current experiment, we delivered 20 Hz double-pulse TMS at stimulus onset asynchrony (SOA) of 100, 150, and 200 ms to cover the relevant processing period, as shown in our previous MEG study (Weidner et al., 2010). There is a temporal overlap of the effects induced in the different stimulated time windows. In particular, the last pulse of the first time window was presented at an SOA of 150 ms. The timing was, therefore, identical to the first pulse of the next time window. Accordingly, the temporal overlap of these effects corresponds to the duration of the effect induced by a single pulse. Although the time windows were not entirely separate, the experimental design, nevertheless, permits a distinction between early and late processes that allows inference on the chronological order of effects in two different brain regions. Results indicated that TMS over EVC and LOC both impaired the strength of the Ponzo illusion. The reduction of illusion strength by TMS was different for EVC and LOC, considering the critical stimulation intervals. The strongest TMS effects were found later when stimulating EVC

compared with LOC. The present findings indicate that EVC is relevant for perceiving an object's size but that the relevant processes underlying object size judgment occur after higher visual regions have been critically involved. This finding is difficult to reconcile with a framework that emphasizes strict feedforward processing in which the contribution of EVC occurs exclusively before those coming from higher visual areas (Lamme and Roelfsema, 2000). The different timing patterns most likely reflect the dynamics of interactions between the two regions. In particular, if processing in EVC and LOC is mutually dependent, then the temporal order in which both are critically involved reflects the direction of information flow. In other words, the fact that LOC is involved earlier than EVC suggests that feedback signals from LOC to EVC are likely to play an essential role. In this case, our findings would support a size scaling neural mechanism that incorporates feedback modulation or recurrent interactions from LOC to EVC. This interpretation is in line with previous studies showing the existence of feedback signals from higher- to lower-level visual areas. For instance, Sterzer et al. (2006) demonstrated the critical role of feedback connection from specialized higher-level visual areas, such as hMT$^+$/V5 in generating a neural representation of the illusory percept in EVC (Sillito et al., 2006). Furthermore, Shpaner et al. (2013), using pathfinder displays, discovered that contour integration initially relies on information processing in higher-order visual cortices (i.e., LOC) before the contextual effects involve early visual regions. Our findings are also consistent with previous findings related to other processes in visual perception, such as awareness of motion (Silvanto et al., 2005) and extrafoveal perception (Chambers et al., 2013).

Our data suggest that both LOC and EVC are essential for processing size information. In particular, to generate perceived size, processing in EVC is required; and according to the present data, it becomes clear that these processes are implemented later than those related to LOC. This allows us to determine the contributions of EVC and LOC further. LOC as a high-level visual area is known to be involved in object perception (M. M. Murray et al., 2002, 2004; Pegna et al., 2002; Ritzl et al., 2003; Mancini et al., 2011), object recognition (Grill-Spector et al., 2001), and object size perception (Cohen et al., 1994; Konkle and Oliva, 2012). Based on the current findings, LOC contributes to these functions rather than constituting an area primarily responsible for forming object representations, including object size as we perceive it. Our findings are thus in line with those from Koivisto et al. (2011), who used a natural-scene categorization task to demonstrate the importance of feedback from LOC, suggesting that LOC may not only be responsible for object processing per se but may also be involved in contextual analysis and scene perception. Concerning size perception and the Ponzo illusion, LOC most likely is involved in extracting scene information that can then be combined and integrated with early visual representations. These early representations may be located in EVC and may account for the TMS effects observed at 200–250 ms SOA since integrating scene information relevant for size perception requires access to early visual representations.

LOC has also been shown to be involved in processing the size information of a single object without any spatial context when no scene information was to be analyzed (Konkle and Oliva, 2012; Chiou and Lambon Ralph, 2016). However, objects, even without a specific spatial context, may involve implicit size information, such as real-world canonical size information, which constitutes one form of top-down context information. LOC might combine different sources of context information, including spatial and conceptual ones, and supports forming the final perceived size representation. This can be tested by manipulating various kinds of context information in one experiment and test the neural response in LOC.

Furthermore, our finding of a late TMS effect in EVC also corresponds well with previous studies (Heinen et al., 2005; Dugué et al., 2011; Koivisto and Silvanto, 2012; Chambers et al., 2013; Allen et al., 2014; see review, see de Graaf et al., 2014). For example, Wokke et al. (2013) showed that EVC was functionally relevant for perceiving illusory contours induced by Kanizsa figures. However, its involvement was found to be later than that of LOC. Camprodon et al. (2010) also reported a TMS effect at 220 ms SOA in a visual discrimination task. In that study, participants were required to perform discriminations between natural images of birds and large mammals. Task performance was significantly impaired when TMS was applied over EVC at SOAs after 100, as well as after 220 ms. While the early TMS masking effect at 100 ms is most likely based on interference with feedforward processing (Amassian et al., 1989), the late effect was attributed to a disruption of feedback projections to EVC. In the present study, we could not find an object-size specific effect in the earliest time window. This may, however, indicate that the early feedforward sweep was already finished before TMS stimulation at 100 ms started. This issue warrants further investigation.

Our findings are also in line with previous reports from patient studies. Cohen et al. (1994) reported 2 patients with lesions of LOC afflicted with hemimicropsia, an apparent reduction of the size of an object when presented in one hemifield. Similarly, Frassinetti et al. (1999) reported altered size perception in the contralesional hemifield of a patient suffering from lesions in the right occipital prestriate area. Our data suggest that size perception emerges as in interaction between LOC and EVC. Importantly, the ability to perceive size was not disturbed by TMS over LOC or EVC. Despite TMS, our participants were able to perform the size comparisons required in the current study. However, TMS had a very specific effect on the impact of context information on perceived size. If perceived size emerges only gradually in the course of an interaction between higher and lower visual areas, then initial size representations should be largely unaffected by higher-level context information. These representations might be based on default context parameters that are later altered by higher-level regions. If, however, the interaction between higher- and lower-level regions is prevented (e.g., due to lesions in LOC), then the initial standard parameters will continue to be effective. Hence, size perception will be independent of context information. In other words, an object would still be perceived but with some standard or default value. This initial standard representation might either be larger or smaller than the object's veridical size and might generate hemimicropsia (Cohen et al., 1994; Frassinetti et al., 1999).

Together, EVC and LOC are essential for generating perceived size. A possible scenario consistent with the present findings is that EVC receives low-level stimulus properties and passes the information to the higher-level visual cortex (e.g., LOC) via feedforward connections. LOC extracts relevant contextual information in the scene, such as texture, stereopsis, as well as familiarity and perspective cues. This information is then fed back to EVC and integrated with initial representations established in the feedforward sweep. In sum, size perception

emerges from the interaction between higher- and low-level visual areas, and feedback information flow plays an essential role.

References

Allen CP, Sumner P, Chambers CD (2014) The timing and neuroanatomy of conscious vision as revealed by TMS-induced blindsight. J Cogn Neurosci 26:1507–1518.

Amassian VE, Cracco RQ, Maccabee PJ, Cracco JB, Rudell A, Eberle L (1989) Suppression of visual perception by magnetic coil stimulation of human occipital cortex. Electroencephalogr Clin Neurophysiol 74:458–462.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.

Bullier J (2001) Integrated model of visual processing. Brain Res Brain Res Rev 36:96–107.

Camprodon JA, Zohary E, Brodbeck V, Pascual-Leone A (2010) Two phases of V1 activity for visual recognition of natural images. J Cogn Neurosci 22:1262–1269.

Chambers CD, Allen CP, Maizey L, Williams MA (2013) Is delayed foveal feedback critical for extra-foveal perception? Cortex 49:327–335.

Chen J, Sperandio I, Henry MJ, Goodale MA (2019) Changing the real viewing distance reveals the temporal evolution of size constancy in visual cortex. Curr Biol 29:2237–2243.e4.

Chiou R, Lambon Ralph MA (2016) Task-related dynamic division of labor between anterior temporal and lateral occipital cortices in representing object size. J Neurosci 36:4662–4668.

Cohen L, Gray F, Meyrignac C, Dehaene S, Degos JD (1994) Selective deficit of visual size perception: two cases of hemimicropsia. J Neurol Neurosurg Psychiatry 57:73–78.

de Graaf TA, Koivisto M, Jacobs C, Sack AT (2014) The chronometry of visual perception: review of occipital TMS masking studies. Neurosci Biobehav Rev 45:295–304.

Dugué L, Marque P, VanRullen R (2011) Transcranial magnetic stimulation reveals attentional feedback to area V1 during serial visual search. PLoS One 6:e19712.

Fang F, Boyaci H, Kersten D, Murray SO (2008) Attention-dependent representation of a size illusion in human V1. Curr Biol 18:1707–1712.

Frassinetti F, Nichelli P, Di Pellegrino G (1999) Selective horizontal dysmetropsia following prestriate lesion. Brain 122:339–350.

Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vision Res 41:1409–1422.

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. Exp Brain Res 233:3535–3541.

Heinen K, Jolij J, Lamme VA (2005) Figure-ground segregation requires two distinct periods of activity in VI: a transcranial magnetic stimulation study. Neuroreport 16:1483–1487.

Koivisto M, Railo H, Revonsuo A, Vanni S, Salminen-Vaparanta N (2011) Recurrent processing in V1/V2 contributes to categorization of natural scenes. J Neurosci 31:2488–2492.

Koivisto M, Silvanto J (2012) Visual feature binding: the critical time windows of V1/V2 and parietal activity. Neuroimage 59:1608–1614.

Konkle T, Oliva A (2012) A real-world size organization of object responses in occipitotemporal cortex. Neuron 74:1114–1124.

Lamme VA (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 15:1605–1615.

Lamme VA (2001) Blindsight: the role of feedforward and feedback cortico-cortical connections. Acta Psychol (Amst) 107:209–228.

Lamme VA, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci 23:571–579.

Li Y, Wang Y, Li S (2019) Recurrent processing of contour integration in the human visual cortex as revealed by fMRI-guided TMS. Cereb Cortex 29:17–26.

Liu Q, Wu Y, Yang Q, Campos JL, Zhang Q, Sun HJ (2009) Neural correlates of size illusions: an event-related potential study. Neuroreport 20:809–814.

Macmillan NA, Creelman CD (1991) Adaptive methods for estimating empirical thresholds. In: Detection theory: a user's guide, pp 183–209. Cambridge: Cambridge UP.

Mancini F, Bolognini N, Bricolo E, Vallar G (2011) Cross-modal processing in the occipito-temporal cortex: a TMS study of the Müller-Lyer illusion. J Cogn Neurosci 23:1987–1997.

Murray MM, Wylie GR, Higgins BA, Javitt DC, Schroeder CE, Foxe JJ (2002) The spatiotemporal dynamics of illusory contour processing: combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. J Neurosci 22:5055–5073.

Murray MM, Foxe DM, Javitt DC, Foxe JJ (2004) Setting boundaries: brain dynamics of modal and amodal illusory shape completion in humans. J Neurosci 24:6898–6903.

Murray SO, Boyaci H, Kersten D (2006) The representation of perceived angular size in human primary visual cortex. Nat Neurosci 9:429–434.

Ni AM, Murray SO, Horwitz GD (2014) Object-centered shifts of receptive field positions in monkey primary visual cortex. Curr Biol 24:1653–1658.

O'Brien F, Cousineau D (2015) Representing error bars in within-subject designs in typical software packages. TQMP 11:126–126.

Pegna AJ, Khateb A, Murray MM, Landis T, Michel CM (2002) Neural processing of illusory and real contours revealed by high-density ERP mapping. Neuroreport 13:965–968.

Pollen DA (2003) Explicit neural representations, recursive neural networks and conscious visual perception. Cereb Cortex 13:807–814.

Ritzl A, Marshall JC, Weiss PH, Zafiris O, Shah NJ, Zilles K, Fink GR (2003) Functional anatomy and differential time courses of neural processing for explicit, inferred, and illusory contours: an event-related fMRI study. Neuroimage 19:1567–1577.

Shpaner M, Molholm S, Forde E, Foxe JJ (2013) Disambiguating the roles of area V1 and the lateral occipital complex (LOC) in contour integration. Neuroimage 69:146–156.

Sillito AM, Cudeiro J, Jones HE (2006) Always returning: feedback and sensory processing in visual cortex and thalamus. Trends Neurosci 29:307–316.

Silvanto J, Cowey A, Lavie N, Walsh V (2005) Striate cortex (V1) activity gates awareness of motion. Nat Neurosci 8:143–144.

Sperandio I, Chouinard PA, Goodale MA (2012) Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. Nat Neurosci 15:540–542.

Sterzer P, Haynes JD, Rees G (2006) Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT$^+$/V5. Neuroimage 32:1308–1316.

Taylor MM, Creelman CD (1967) PEST: efficient estimates on probability functions [erratum and note]. J Acoust Soc Am 42:1097–1097.

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366–383.

Wang P, Nikolić D (2011) An LCD monitor with sufficiently precise timing for research in vision. Front Hum Neurosci 5:1–10.

Weidner R, Boers F, Mathiak K, Dammers J, Fink GR (2010) The temporal dynamics of the Müller-Lyer illusion. Cereb Cortex 20:1586–1595.

Weidner R, Fink GR (2007) The neural mechanisms underlying the Müller-Lyer illusion and its interaction with visuospatial judgments. Cereb Cortex 17:878–884.

Willenbockel V, Sadr J, Fiset D, Horne GO, Gosselin F, Tanaka JW (2010) Controlling low-level image properties: the SHINE toolbox. Behav Res Methods 42:671–684.

Wokke ME, Vandenbroucke AR, Scholte HS, Lamme VA (2013) Confuse your illusion: feedback to early visual cortex contributes to perceptual completion. Psychol Sci 24:63–71.