

# Human Depth Sensitivity Is Affected by Object Plausibility

Nicole H. L. Wong<sup>1</sup>, Hiroshi Ban<sup>2</sup>, and Dorita H. F. Chang<sup>1</sup>

## Abstract

■ Using behavioral and fMRI paradigms, we asked how the physical plausibility of complex 3-D objects, as defined by the object's congruence with 3-D Euclidean geometry, affects behavioral thresholds and neural responses to depth information. Stimuli were disparity-defined geometric objects rendered as random dot stereograms, presented in plausible and implausible variations. In the behavior experiment, observers were asked to complete (1) a noise-based depth task that involved judging the depth position of a target embedded in noise and (2) a fine depth judgment task that involved discriminating the nearer of two consecutively presented targets. Interestingly, results indicated greater behavioral sensitivities of depth judgments for implausible versus

plausible objects across both tasks. In the fMRI experiment, we measured fMRI responses concurrently with behavioral depth responses. Although univariate responses for depth judgments were largely similar across cortex regardless of object plausibility, multivariate representations for plausible and implausible objects were notably distinguishable along depth-relevant intermediate regions V3 and V3A, in addition to object-relevant LOC. Our data indicate significant modulations of both behavioral judgments of and neural responses to depth by object context. We conjecture that disparity mechanisms interact dynamically with the object recognition problem in the visual system such that disparity computations are adjusted based on object familiarity. ■

## INTRODUCTION

Binocular disparity, which refers to the difference in the images that fall on the left and right retinas, is one of the most important depth cues and provides crucial information for interacting with a 3-D environment. Indeed, disparity information has been shown to be critical for grasping (Watt & Bradshaw, 2002), breaking camouflage (Julesz, 1971), and estimating surface reflectance (Blake & Bühlhoff, 1990), although recent human studies suggest inaccuracies in judging depth intervals from binocular disparity (Norman, Todd, & Orban, 2004; Tittle, Todd, Perotti, & Norman, 1995). Still, the mechanisms and hierarchy of disparity information processing remain largely elusive. Although an increasing body of work has indicated that the retrieval of depth (from disparity) is subserved by nodes in dorsal and ventral cortex of the primate brain (Welchman, 2016; Preston, Li, Kourtzi, & Welchman, 2008; Uka & DeAngelis, 2006; Tsao et al., 2003), this knowledge has been gained through the testing of highly simplified patterns that are far from relevant to those encountered in the everyday world.

Neurophysiological work in the macaque has revealed that middle temporal (V5) is involved in segmentation of coarse disparity targets from noise, but not in an alternative task requiring the discrimination of clear but fine depth positional differences (i.e., a “fine” or “feature”

discrimination task; Uka & DeAngelis, 2006), which implicates ventral regions including V4 and the inferior temporal gyrus (Shiozaki, Tanabe, Doi, & Fujita, 2012; Uka, Tanabe, Watanabe, & Fujita, 2005). A similar dissociation in terms of responses to coarser segmentation of disparity targets from noise versus fine depth information along the dorsal and ventral cortex have been observed in humans (Chang, Mevorach, Kourtzi, & Welchman, 2014).

The few attempts at understanding neural representations of binocular depth information using more complex stimuli appear to hint at more complex cortical involvement in the primate brain. With the use of 3-D convex shapes, Chandrasekaran, Canon, Dahmen, Kourtzi, and Welchman (2007) showed that discrimination of 3-D shapes recruits additional ventral and dorsal visual areas (V7, POIPS), possibly from the need to segment contours and discriminate curvature details. Additionally, V3 and V3A are involved in the processing of 3-D slants and curvature (Ban & Welchman, 2015; Georgieva, Peeters, Kolster, Todd, & Orban, 2009), regions that appear less prominently in previous studies that adopted simplified geometric stimuli. This implies that possible interactions among higher extrastriate areas for stereoscopic depth perception may have been overlooked in previous work. This is an important consideration given the potential behavioral manifestations of such interactions (i.e., is depth sensitivity affected by object identity?). One way to probe this is to test how sensitivity (to depth) and the associated neural responses are affected by varying

<sup>1</sup>The University of Hong Kong, <sup>2</sup>Osaka University

object-level information, varying their relevance to everyday vision. We ask these questions here, choosing to vary object information by manipulating their physical plausibility.

Indeed, the retrieval of many visual dimensions appears to be affected by object context. For example, race appears to modulate luminance judgments of face stimuli (Levin & Banaji, 2006), and the naturalness of an object affects its perceived color (Witzel, Valkova, Hansen, & Gegenfurtner, 2011; Olkkonen, Hansen, & Gegenfurtner, 2008). More relevant, Murray, Kersten, Olshausen, Schrater, and Woods (2002) demonstrated that responses in human primary visual cortex can be modulated by activity in higher object processing areas, in which more complex shape stimuli invoke weaker responses in V1 as compared with lines and 2-D shapes. Furthermore, neurons in macaque V2 appear to use contextual depth information beyond their classical receptive field to integrate occluded contours (Bakin, Nakayama, & Gilbert, 2000). Although not a lot of work has been done to date to understand the neural substrates underlying these object-level modulations of visual sensitivity to lower order visual properties, this body of work renders it imperative to ask whether responses of the stereovision system in fact is subject to object-level modulations that appear to be increasingly common ground in vision. For example, there is further intricate cross-talk between dorsal and ventral extrastriate regions for visual action processing (Rossetti & Pisella, 2002; Chao & Martin, 2000), and even reading (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008) that has broad functional implications.

Our study has two aims: (1) to investigate whether object context modulates depth sensitivity and (2) to elucidate the neural representation of depth in the context of everyday relevant and irrelevant 3-D complex objects. Using both behavioral (Experiment 1) and fMRI (Experiment 2) paradigms, we sought to understand how the physical plausibility of objects affects depth perception. We tested disparity sensitivities under varying “plausibility contexts” using signal-in-noise and feature discrimination task paradigms that have been established to target different aspects of depth perception.

## METHODS

### Experiment 1: Behavior

#### *Participants*

All participants provided written informed consent in line with procedures approved by the Human Research Ethics Board of the University of Hong Kong and methods conformed to the relevant guidelines and regulations. All participants were screened for stereo deficits and had normal or corrected-to-normal vision as screened with the Snellen linear acuity chart. A total of 64 observers participated in this first experiment. Thirty-two participants (age:  $M = 22.69$  years,  $SD = 3.49$  years; 8 men, 24 women) were tested in the first subtask

(signal-in-noise task), and a second group of 32 participants (age:  $M = 23.14$  years,  $SD = 3.47$  years; 10 men, 22 women) were tested in the second subtask (feature task). Education levels (years of education counting from primary 1) and the age of participants were comparable across tasks, education level:  $t(62) = -1.04$ ,  $p = .301$ ; age:  $t(62) = -0.57$ ,  $p = .573$ .

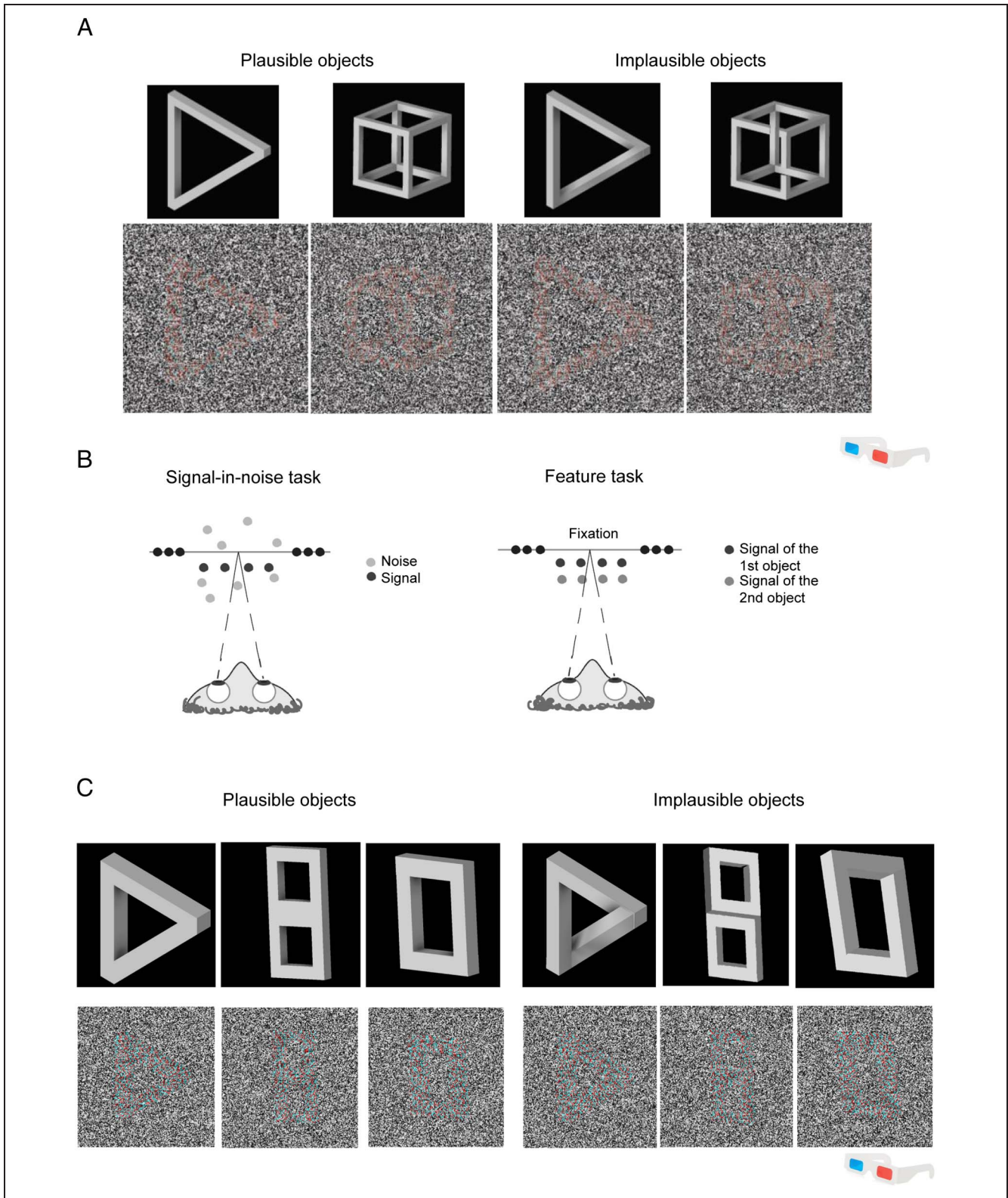
#### *Apparatus*

Stimuli were presented on a mirror stereoscope in which each eye viewed the left or right half of a single 24-in. monitor (resolution:  $1920 \times 1080$ ; 60 Hz refresh frequency) through silver-fronted mirrors mounted and oriented at  $45^\circ$  angles. The viewing distance was 65 cm, as stabilized with a chinrest.

#### *Stimuli*

Stimuli consisted of two classes of 3-D objects (triangle and cube), each with a physically plausible and implausible variation, rendered as random dot stereograms (RDS). 3-D stimuli were first generated using Inventor Studio 2018 (Autodesk, 2018; Figure 1A). The plausible triangle was an intact equilateral triangle, and the plausible cube was a Necker cube. Both classes of plausible objects were projected isometrically (i.e., angles between any two of the  $x$ ,  $y$ ,  $z$  axes were the same). The implausible triangle (i.e., a Penrose triangle) was made from two perpendicular square beams, and an S-shaped beam positioned such that, when projected isometrically, the resulting object resembles an intact, equilateral triangle. The implausible cube was derived from a Necker cube using solid beams as edges. Two small sections of the edges were removed: (1) the section where the top edge of the front square appeared to intersect with the vertical edge of its diagonal and (2) the section where the right edge of the front square appeared to intersect with the lower horizontal edge at the back of the cube. These 3-D structures were then defined in terms of depth maps. To match low-level features between plausible and implausible objects, we ensured that (1) the beam width and surface areas were equivalent across the two variations of objects and (2) the overall disparity across both object variations were equivalent. From the depth maps, RDS were finally generated by computing the corresponding horizontal displacement from the gray-level intensity maps.

We use the term “stimulus plausibility” to refer to the two variations of each stimulus class. Alternatively, these objects could perhaps also be defined in terms of the congruency between its local constituents and global structure. Implausible objects have conflicting local and global structures. That is, structural information from local components would suggest them to be coherent parts of an object; however, interpretation of the entire object becomes incoherent when they are considered globally as parts of a whole.



**Figure 1.** (A) Plausible and implausible variants of the 3-D objects used in the current study. Depth maps (top row) were rendered into RDS (bottom row). (B) The geometry of the stimuli for the SNR depth task and feature task. Participants were asked to judge whether the target was in front of or behind a zero-disparity reference plane (signal-in-noise task) or which of two consecutively presented objects was nearer (feature task). (C) Plausible and implausible variants of the new 3-D objects used in a replication of the behavioral paradigms. Note that grayscale images were shown via a mirror stereoscope in the experiment proper (i.e., red-cyan anaglyphic stimulus examples are shown here for greater clarity).

The RDS were presented using MATLAB (The MathWorks) with extensions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Dots of the RDS were randomly black and white and had a density of 20 dots/deg<sup>2</sup>, with each dot subtending 0.025°. Each RDS was presented on a black background and subtended 7.4° × 7° in size. All objects had a maximum disparity of 4.86 arcmin. The RDS was additionally surrounded by a grid of black and white squares (each 0.44° in size) so as to assist vergence. The RDS were presented on the left- and right-halves of the monitor, which were viewed through the silver-fronted mirrors.

### Tasks

**Signal-in-noise depth task.** For the signal-to-noise ratio (SNR) task, we varied the ratio of signal dots defining the object relative to noise dots. Participants were asked to judge whether the target was in front of or behind the reference plane by pressing one of two keyboard keys. At 100% signal, the object was coherent as all dots were defining the object surface. However, difficulty increases as fewer signal dots relative to the number of noise dots, distributed randomly within a ±5 arcmin range, were defining the object surface. Therefore, at 0% signal, the stimulus consisted purely of noise dots. Percentage signal was adjusted according to QUEST staircase procedure for measuring percentage signal required for each participant to attain 82% correct-level performance (Watson & Pelli, 1983).

**Feature discrimination task.** For the feature task, we manipulated the (depth) distance between a reference and target object presented in consecutive intervals, with a QUEST staircase procedure yielding thresholds at 82% accuracy. We displaced the reference object forward by 2.5 arcmin and varied the disparity of the target object within a ±2.5 arcmin range with respect to the reference. Participants were asked to judge which of two consecutively presented intervals contained the “nearer” object. Figure 1B illustrates the geometry of both tasks.

### Procedures

In the SNR task, on each trial, participants were asked to fixate first on a nonius-type fixation for 500 msec. This fixation comprised left and right image components that, when fused, appeared as an intact square with four lines extending across its center vertically and horizontally. The stimulus was then presented for a duration of 500 msec, after which a response was allowed. Trials were separated by an intertrial interval of 500 msec. Each participant completed 15 demonstration trials before the experiment proper to ensure that they understood the testing requirements. Each participant then completed three runs of trials for each condition (plausible and implausible). A particular run consisted of two interleaved staircases of 60 trials (each staircase with equal number of triangle and cube stimuli), comprising 8 practice trials and 52 experimental

trials. The initial test value was determined by the threshold obtained during the practice trials. The entire task lasted approximately 1 hr.

In the feature task, on each trial, participants were asked to first fixate for 500 msec. Two stimuli, 500 msec each in duration and separated by an ISI of 200 msec, were then presented. Similar to the SNR task, trials were separated by an interval of 500 msec. Each participant completed three runs of trials for each condition (plausible and implausible). Similar to the SNR task, each run consisted of 120 trials with equal number of triangle and cube stimuli. Each run consisted of two interleaved staircases of 60 trials, comprising 8 practice trials and 52 experimental trials. The initial test value was determined as the best threshold approximate from the practice trials. All other parameters and procedures were identical to those for the SNR task.

## Experiment 2: fMRI

### Participants

Nineteen participants (age:  $M = 24.63$  years,  $SD = 4.52$  years; 13 men, 6 women) were tested in the fMRI component of this study. Data of one participant were excluded from the behavioral analysis only since his initial in-bore signal-in-noise depth thresholds were at ceiling (see Procedures below for stimulus sampling) but had consistently attained at least 70% performance accuracies in subsequent functional runs. All had normal or corrected-to-normal vision and provided written informed consent in accordance with procedures approved by the Human Research Ethics Committee of the University of Hong Kong, and the ethics committee of National Institute of Information and Communications Technology, Osaka, Japan.

### Stimuli

Stimuli were generated identically to the main behavioral tasks. They were projected with two projectors (WUX4000, Canon), each equipped with a different polarizing filter (resolution: 1280 × 1024 pixels; 60-Hz refresh rate). Stimuli were back-projected to a screen placed 96 cm from the back of the bore and viewed through a 45° tilted mirror mounted in front of the participant. Participants wore corresponding polarizing lenses.

### Image Acquisition

fMRI data were acquired using a 3-T Siemens Trio MR scanner at Center for Information and Neural Networks, National Institute of Information and Communications Technology (Osaka, Japan) using a half of the 32-channel phase array coil, which covers occipital lobe. EPI data were obtained from 78 axial (slightly oblique along the AC-PC line) slices (whole-brain coverage, repetition time = 2000 msec, echo time = 30 msec, field of view = 192 × 192, flip

angle = 75°, multiband factor = 3,  $2 \times 2 \times 2 \text{ mm}^3$  resolution), collecting 205 volumes for each functional run. The multiband EPI sequence was provided by the University of Minnesota (under a C2P contract). Head movement was limited by foam padding within the head coil. For each participant, a high-resolution T1 image was acquired using a  $1 \text{ mm}^3$  anatomical scan (208 slices, repetition time = 1900 msec, echo time = 2.48 msec, field of view =  $256 \times 256$ , flip angle = 9°) for accurate coregistration of fMRI images to individual anatomy space and for reconstructing cortical surfaces. Localization of ROIs was performed in a separate session.

### *Identification of ROIs*

For each participant, we identified ROIs (V1, V2, V3, V3A, V3B, V4, V7) using standard phase-encoded retinotopic mapping procedures; polar angle representations were obtained with a slowly rotating (clockwise or anticlockwise) checkerboard wedge stimulus (Sereno et al., 1995). LOC was defined as the set of voxels in lateral occipitotemporal cortex that responded significantly ( $p < .01$ ) to intact object images compared with their scrambled (mosaic) counterparts (Kourtzi, Tolias, Altmann, Augath, & Logothetis, 2003). The human motion complex, hMT+, was defined as the region that responded significantly higher ( $p < .01$ ) to coherent dot motions (toward and away from fixation) compared with static dot images (Huk, Dougherty, & Heeger, 2002). For five participants who did not complete MT+ localizer scan because of time constraints, MT+ was defined as a spherical ROI (5 mm rad) centered on Talairach coordinates of [left: -51, -72, 0; right: 51, -69, 3] (Orban et al., 2003).

### *Design and Procedures*

Participants completed the SNR task only for the fMRI session. The task question remained the same; changes to the procedures are detailed below. We elected to test this task as it exhibited relatively smaller individual subject variability as compared with the feature task in Experiment 1, but more critically is the task for which we have shown previously to have substantial learning-related reorganization in cortex and therefore deemed more intriguing (Chang, Kourtzi, & Welchman, 2013). Before the scanning session, participants completed two psychophysics runs inside the bore of the scanner corresponding to one each for the plausible and implausible conditions. Each run consisted of two interleaved staircases with equal number of triangle and cube stimuli. Each staircase consisted of 60 trials, yielding a total of 120 trials per run. For each condition, we then took the last 30 trials of each staircase and generated stimuli with an SNR sampled from a uniform distribution between  $\pm 1$  standard deviation of the mean threshold estimates. This ensured that stimulus difficulty was matched across conditions and across subjects.

fMRI runs were arranged in a block design, with each block lasting 16 sec. One scan run consisted of five main block types comprising four stimulus condition blocks (e.g., plausible triangle, implausible triangle, plausible cube, implausible cube) and a fixation block. Stimulus blocks were interleaved with fixation blocks. On each trial in a stimulus block, the stimulus was presented for 0.500 sec, after which participants were allotted a maximum duration of 1.5 sec to respond. Each block comprised eight trials, with an equal number of “near” and “far” trials. Each stimulus block was repeated three times within a particular run, yielding 24 repetitions of a particular stimulus condition. Stimulus condition (block) order was randomized. Each run consisted of 96 trials and lasted 6.4 min. The entire scanning session lasted 90 min.

### *Imaging Data Analysis*

MRI data were processed using Brain Voyager QX 3.6.0 (Brain Innovations). The initial five volumes of each functional run were discarded to eliminate effects of start-up magnetization transients in the data. Functional data were preprocessed using slice-time correction, 3-D motion correction, high-pass filtering (three cycles per run), and linear trend removal. No spatial smoothing was applied to keep fine-scale spatial response patterns for multivoxel pattern analyses. Functional images were aligned with each participant’s anatomical scan and transformed into Talairach space (Talairach & Tournoux, 1988). Functional data between different runs were aligned to the first functional volume of the first run.

Functional data were further analyzed at two levels: using both a random-effects analysis (general linear model, GLM) and a multivariate pattern analysis (MVPA). Whereas the GLM computes the response amplitude evoked by an experimental condition averaged across voxels in each ROI, MVPA is able to gauge the uniqueness of a response pattern to the experimental conditions in each ROI. In the GLM analyses, design matrices were defined by modeling stimulation periods separately for each stimulus condition by a boxcar model convolved with a canonical hemodynamic function (two-gamma model; Glover, 1999) to provide idealized hemodynamic responses. Six motion parameters (three translation parameters, in millimeters, and three rotation parameters pitch, roll, and yaw, in degrees) were additionally included in the design matrices as nuisance regressors. The time course signal of each voxel was subsequently modeled as a linear combination of the different regressors (least squares fits), and the regressor coefficients were used for contrasts of the different experimental conditions. For each stimulus type, we contrasted responses to the plausible versus implausible variations.

For the MVPA, we used a linear Support Vector Machine (SVM) classifier (libSVM; Chang & Lin, 2011) together with an adaptive feature selection algorithm, Recursive Feature Elimination (RFE), for estimating spatial patterns

(De Martino et al., 2008). The basic principle of the RFE is to start with all voxels in the ROI and to gradually exclude voxels that do not contribute in discriminating patterns from different experimental conditions. The use of RFE provides estimates based on voxel subsets extracted from training data sets (no voxel was used from test data sets in each estimation step) with higher weight values within each ROI, thus eliminating the need of choosing a rather arbitrary fixed number of voxels per ROI. Essentially, all voxels and their time courses were first converted to  $z$  scores and shifted 4s to account for the delay of the hemodynamic response. Specifically, retaining blocks as data units, we took 80% of time-collapsed data as training data to compute SVM weights. The resampling procedure was repeated 20 times within an RFE step (i.e., each voxel has 20 sampled weights, which were averaged ultimately). We ordered voxels based on their weight from the highest to the lowest during each step. The five most uninformative voxels were omitted based on those weights, and the remaining voxels were used to decode the test patterns. This provided us an accuracy at the current voxel count. The procedure was repeated until voxel count reached 50 voxels. Mean prediction accuracies were tested against chance level (0.53), as obtained via permutation tests of the data (i.e., by running 1000 SVMs with shuffled labels).

### Supplementary Experiment

*Supplementary task with new objects.* In a supplementary experiment, we assessed the generalizability of our findings by replicating the SNR and feature tasks in the laboratory using new objects. We retained the original triangle stimuli (as used in Experiments 1 and 2) and newly included two object classes drawn from Freud, Rosenthal, Ganel, and Avidan (2015): cubes and rectangles. For the construction of the new plausible “cube” stimuli, two squares were arranged vertically facing the same azimuthal direction. The implausible version was made by repositioning the top square so that it faced an azimuthal direction opposite to that of the bottom square while tilting the top square, such that the bottom is visible. For the new “rectangle” stimuli, two L-shaped bars were first made. One of the bars was then tilted and positioned such that neither ends were touching the other’s L-shaped component, generating an implausible stimulus (Figure 1C). Procedures were identical to those described for the main behavioral experiment (Experiment 1).

## RESULTS

### Behavioral Results

#### Signal-in-noise Task

Behavioral thresholds for the SNR task are presented in Figure 2A. Inspection of this figure indicates that participants were more sensitive (i.e., thresholds were lower)

for depth judgments of implausible objects and, in particular, implausible triangles relative to the other objects. This observation is confirmed by a 2 (Class: triangle and cube)  $\times$  2 (Plausibility: plausible and implausible) repeated-measures ANOVA that indicated a main effect of Class,  $F(1, 31) = 5.23, p = .029, \eta_p^2 = .144$ , and a main effect of Plausibility,  $F(1, 31) = 4.36, p = .045, \eta_p^2 = .123$ , reflecting the fact that thresholds were on average lower for implausible than for plausible objects. There was also a significant interaction between Class and Plausibility,  $F(1, 31) = 6.21, p = .018, \eta_p^2 = .167$ . Interestingly, thresholds were lower for implausible versus plausible objects, but only for the triangle stimulus class,  $t(31) = 3.70, p = .001, d = 0.653$ . There was no difference in thresholds between plausible and implausible cubes,  $t(31) = 0.53, p = .603, d = 0.09$ . Furthermore, there was no significant difference between thresholds for the plausible triangle and plausible cube,  $t(31) = 0.495, p = .624$ , although thresholds for the implausible triangle were significantly lower relative to those for the implausible cube,  $t(31) = -2.99, p = .0055$ .

#### Feature Discrimination Task

Thresholds for the feature task are presented in Figure 2B. As for the SNR task, thresholds were entered into a 2 (Class: triangle and cube)  $\times$  2 (Plausibility: plausible and implausible) repeated-measures ANOVA that indicated a main effect of Class,  $F(1, 31) = 4.42, p = .044, \eta_p^2 = .125$ , and a main effect of Plausibility,  $F(1, 31) = 11.30, p = .002, \eta_p^2 = .267$ . There was no interaction between Class and Plausibility,  $F(1, 31) = 1.87, p = .182, \eta_p^2 = .057$ . Consistent with the SNR task, participants demonstrated significantly higher sensitivities (i.e., lower thresholds) toward feature depth judgments for implausible objects versus plausible objects.

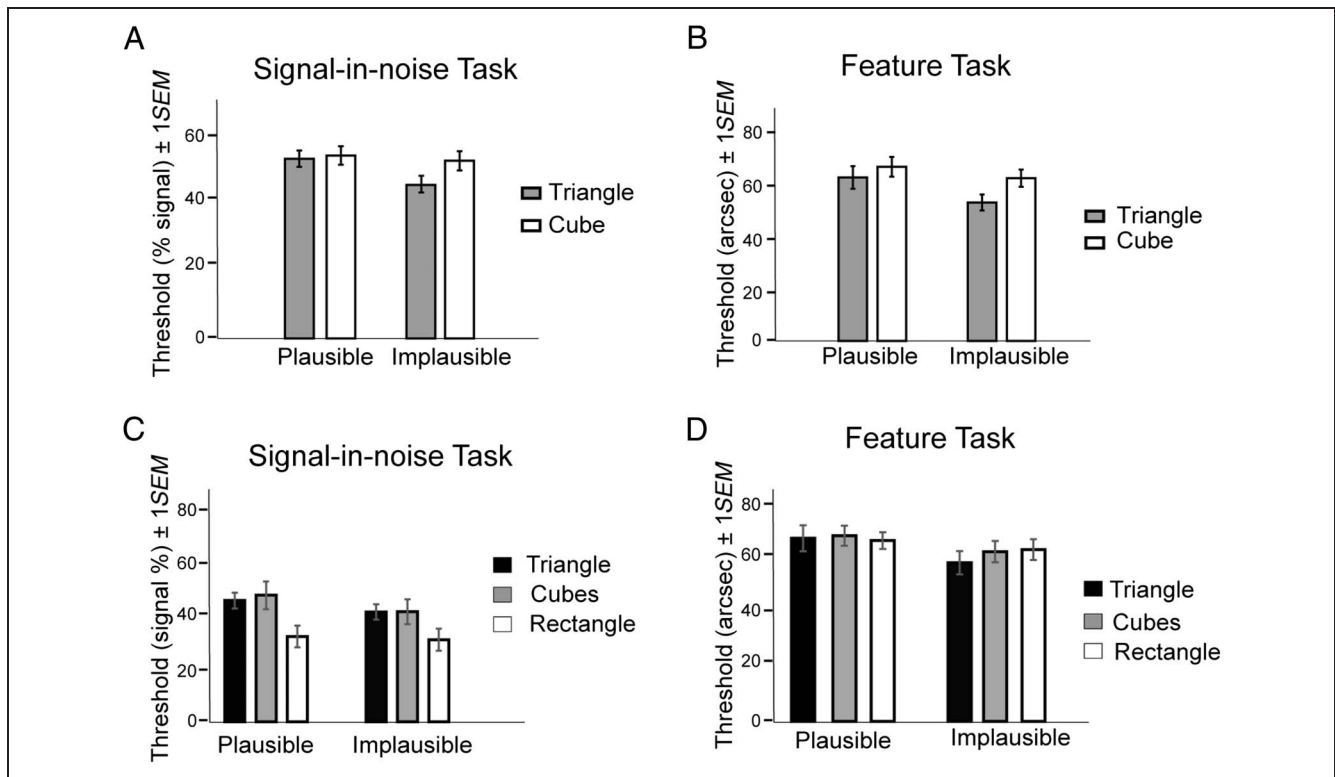
#### Supplementary Task with New Objects

Thresholds were entered into a 3 (Class: triangle, cube and rectangle)  $\times$  2 (Plausibility: plausible and implausible) repeated-measures ANOVA. Results indicated a main effect of Plausibility for both tasks (SNR task ( $n = 23$ ):  $F(1, 22) = 4.30, p = .05$ ; feature task ( $n = 20$ ):  $F(1, 19) = 6.62, p = .019$ ), reflecting the fact that thresholds for implausible objects were lower than those for plausible objects in both tasks. The analysis for the SNR task indicated an additional main effect of Class,  $F(2, 44) = 7.57, p = .001$ , but no interactions (Figure 2C, D).

### fMRI Results

#### fMRI Behavioral Results (SNR Task)

As noted, at the start of each scanning session, participants completed two SNR behavioral task runs while laying inside scanner bore. Each run measured independent



**Figure 2.** Behavioral depth discrimination thresholds for the (A) SNR task and (B) feature depth discrimination task presented separately for plausible (filled) and implausible (open) objects in Experiment 1. Error bars represent  $\pm 1 SEM$  ( $n = 32$  per task). Behavioral performance on Supplementary Experiment 2, testing (C) SNR ( $n = 23$ ) and (D) feature task ( $n = 20$ ) performance using new object exemplars.

thresholds for plausible and implausible objects. Crucially, consistent with behavioral results obtained in the laboratory, thresholds for implausible triangles were again lower than those for plausible triangles; thresholds for the two variants of cubes did not differ.

Thresholds were entered into a 2 (Class: triangle and cube)  $\times$  2 (Plausibility: plausible and implausible) repeated-measures ANOVA that indicated a main effect of Class,  $F(1, 17) = 4.93, p = .04, \eta_p^2 = .225$ , and a significant interaction between Class and Plausibility,  $F(1, 17) = 4.48, p = .049, \eta_p^2 = .208$ . There was no main effect of Plausibility,  $F(1, 17) = 2.56, p = .128, \eta_p^2 = .131$ . Follow-up paired  $t$  tests indicated that thresholds differed significantly between plausible and implausible triangles,  $t(17) = 2.61, p = .018, d = 0.615$ , but not between plausible and implausible cubes ( $p > .05$ ).

#### GLM Random Effects (Beta Weights)

The GLM computes the amplitude averaged across voxels in each ROI. Within each ROI, we first evaluated univariate GLM responses to our stimuli in terms of beta weights (percent signal changes), contrasting the plausible versus implausible variations for the two classes of stimuli. GLM beta weights were entered into a 2 (Class)  $\times$  2 (Plausibility)  $\times$  9 (ROI) repeated-measures ANOVA that indicated a significant interaction between Class and ROI,  $F(3.14, 56.50) = 2.77, p = .048$  (with Greenhouse–Geisser

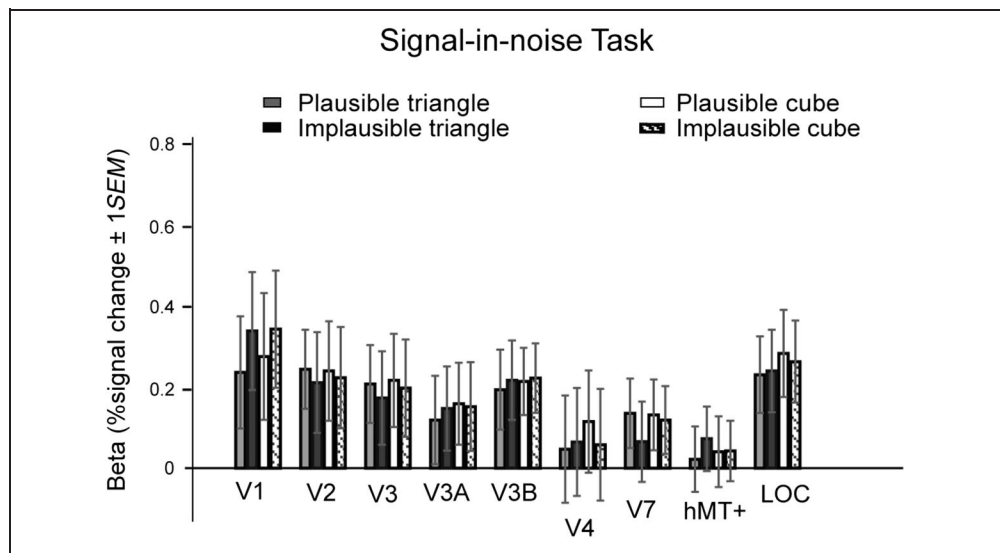
corrections applied for violations of sphericity; Figure 3). Follow-up paired  $t$  test indicates that signals for cubes were slightly higher than that for triangles in V3,  $t(37) = -2.28, p = .028, d = 0.370$ , and V3B,  $t(37) = -3.72, p = .001, d = 0.604$ .

We repeated the analysis, including Hemisphere (left/right) as an additional factor. Critically, the analysis did not indicate any significant main effects or interactions involving Hemisphere,  $F(1, 7) = 0.985, p = .354$ . The interaction among Hemisphere, Class, and Plausibility was not significant,  $F(1, 7) = 0.24, p = .638$ , nor that among Hemisphere, Class, and ROI, Greenhouse–Geisser corrected,  $F(8, 56) = 1.25, p = .317$ . Finally, the interaction among Hemisphere, Class, ROI, and Plausibility was also not significant,  $F(8, 56) = 1.88, p = .139$ .

#### Multivariate Pattern Analysis

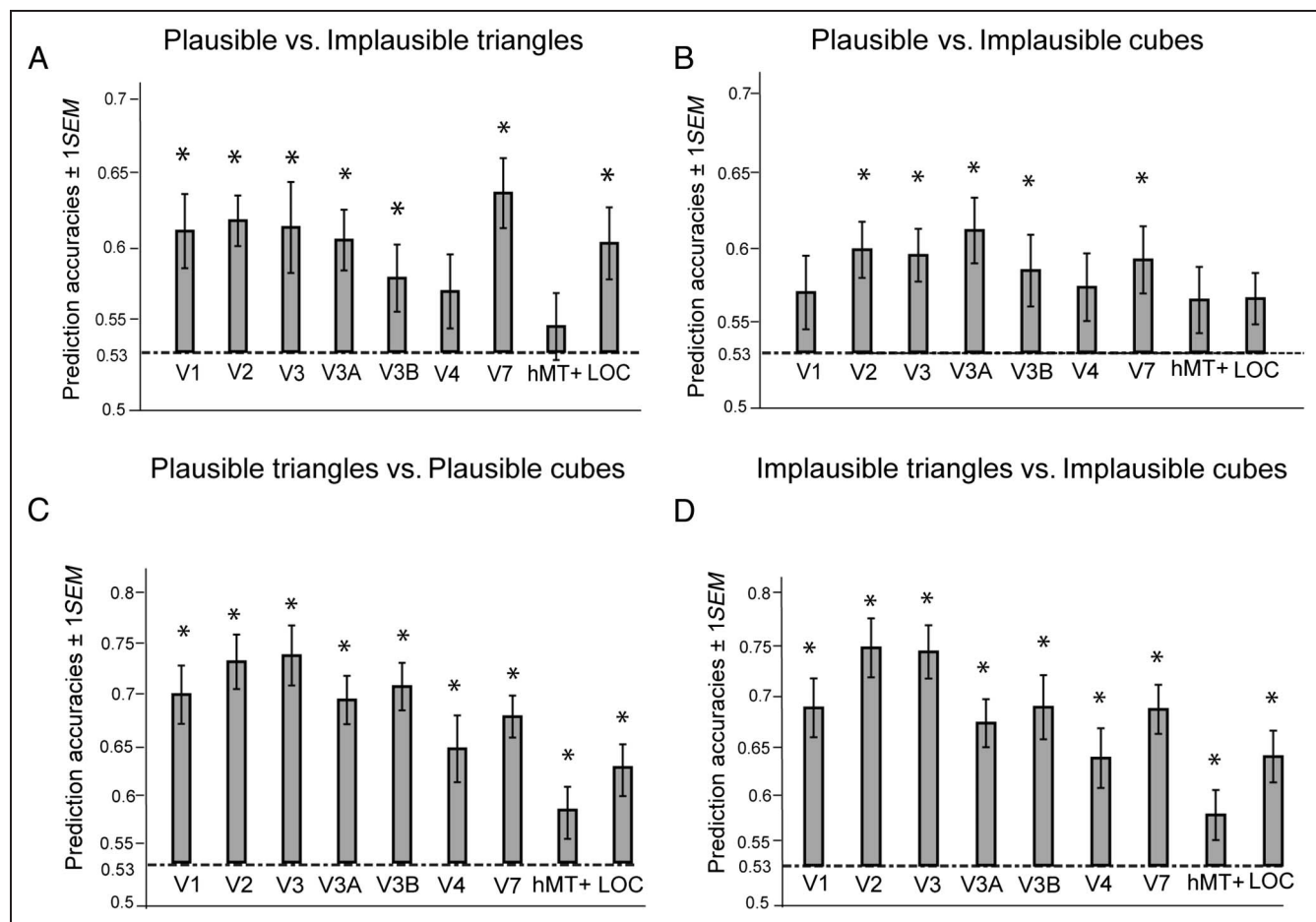
Next, we analyzed responses of the ROIs in terms of their multivariate response patterns. The MVPA gauges the uniqueness of a response pattern to the experimental conditions in each ROI. In Figure 4, we present MVPA classification accuracies for discriminating between (a) plausible versus implausible triangle stimuli, (b) plausible versus implausible cube stimuli, (c) plausible triangle versus plausible cube stimuli, and lastly, (d) implausible triangle versus implausible cube stimuli. The classification accuracies of all ROIs, for all comparisons, were tested

**Figure 3.** Univariate fMRI responses, presented in terms of GLM beta weights, for plausible and implausible triangle and cube stimuli, across ROIs. Error bars represent  $\pm 1 SEM$  ( $n = 19$ ).



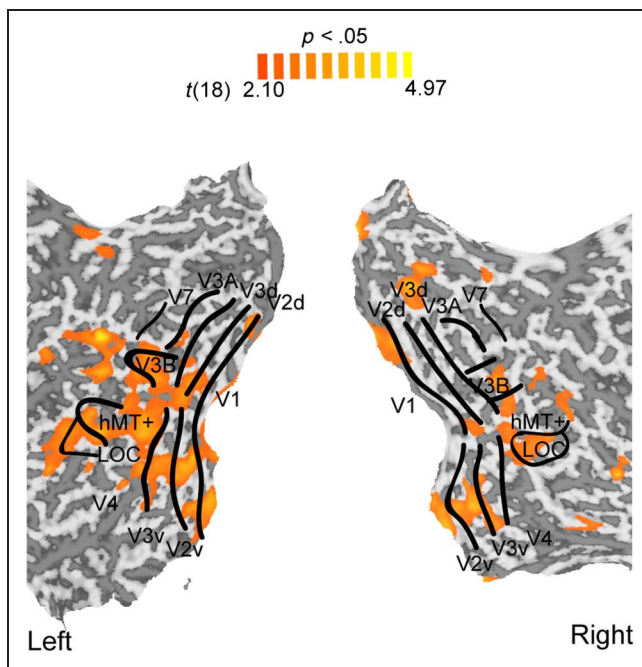
with paired  $t$  tests against a shuffled baseline of 0.53 (see Methods), with corrections made for multiple comparisons holding false discovery rate  $q < 0.05$ . Accuracies for discriminating between the plausible and implausible

triangles were above baseline in early visual areas (V1, V2, V3), in intermediate and extrastriate dorsal areas (V3A, V3B, V7), and in a higher-order ventral region (LOC; Figure 4A). By contrast, SVM accuracies for the cube



**Figure 4.** MVPA classification accuracies for discriminating between response patterns for (A) plausible and implausible triangle stimuli, (B) plausible and implausible cube stimuli, (C) plausible triangle and plausible cube stimuli, and lastly (D) implausible triangle and implausible cube stimuli. The dotted line indicates the shuffled baseline of 0.53. Above baseline accuracies are indicated by asterisks (false discovery rate  $< 0.05$ ). Error bars represent  $\pm 1 SEM$  ( $n = 19$ ).





**Figure 5.** Searchlight map generated by moving a 6-mm sphere across cortex. The representative map shown here shows regions where classification accuracies for discriminating between plausible and implausible stimuli were significantly different from chance (statistics as indicated by the orange gradient). Results are superimposed onto the representative flat maps of a single participant. Sulci are coded in dark gray.

stimuli were above baseline only in V2, V3, V3A, V3B, and V7 (Figure 4B). In terms of the SVM analyses involving comparisons of stimulus classes (i.e., triangles vs. cubes), classification accuracies for discriminating between the plausible triangle and plausible cube as well as between the implausible triangle and implausible cube were above baseline across all the ROIs (Figure 4C, D).

In a supplementary analysis, we collapsed the two object types and trained and tested the SVM for discriminating between plausible versus implausible conditions more broadly. We found that accuracies for discriminating between plausible objects versus implausible objects were above baseline in many of the same areas, including V3A, V7, and LOC, in addition to V1, V2, V3, V3B, and V4 (Supplementary Figure S1<sup>1</sup>). The implication of additional areas here is likely due to the fact that test sensitivity has been artificially increased by pooling the two objects' data (thereby decreasing variance).

We further tested the generality of multivariate representations by “training” the SVM classifier to distinguish between plausible versus implausible variations of one object class (e.g., triangles) and “testing” using the other object class (e.g., cubes). We found that, under this scenario, the classification accuracy of V2, V3, and LOC were above baseline for cross-training/discriminations in both directions (i.e., training on plausible vs. implausible triangles and testing on plausible vs. implausible

cubes; training on plausible vs. implausible cubes and testing on plausible vs. implausible triangles), suggesting a key role of these regions for decoding object plausibility (Supplementary Figure S2).

Finally, in addition to the ROI-based MVPA analyses, we conducted a searchlight analysis by moving a spherical ROI of 6-mm radius across the brain, testing pattern discriminability for plausible versus implausible objects, collapsed across the two object classes. The results indicated that the relevant clusters are well captured by our choice of ROIs (Figure 5).

## DISCUSSION

We investigated whether object context, defined here in terms of the object's congruence with standard 3-D Euclidean geometry (i.e., physical plausibility), affects behavioral sensitivities and neural responses to disparity-defined depth. In Experiment 1, we showed that, for both judging a target's depth position from noise (SNR task) and discriminating fine disparity differences between coherent stimuli (feature task), thresholds were lower (i.e., depth judgments were better) for implausible objects (particularly the implausible triangle stimuli) than for the corresponding plausible counterparts. In Experiment 2, we selected the task that elicited the more robust effects (SNR task) and used fMRI to ask how the observed behavioral modulations are reflected in the brain. Although we found that the univariate responses across cortex were largely comparable regardless of object plausibility, comparisons of the multivariate response patterns were more revealing. MVPAs showed that early and intermediate retinotopic areas (V1, V2, V3, V3A, V3B) and higher order extrastriate areas V7 and LOC were well able to differentiate between response patterns for plausible and implausible triangle stimuli. All of these regions, except for V1 and LOC, were similarly able to discriminate between patterns for plausible and implausible cubes.

Our behavioral data add to the growing body of literature in other visual domains demonstrating surprising contextual modulations of object-level information on behavioral judgments of basic visual features such as luminance, orientation, color, and contrast (e.g., Marlow & Anderson, 2015; Marlow, Todorović, & Anderson, 2015; Olkkonen et al., 2008; Schwartz, Hsu, & Dayan, 2007; Levin & Banaji, 2006). In 3-D relevant work, Bülthoff, Bülthoff, and Sinha (1998) presented stereoscopic biological motion walkers and asked participants to report whether three indicated points laying on the walker structure were on the same plane. They found that dots on the same limb were more often perceived to be coplanar as compared with dots on different limbs suggesting that the familiarity of object affects depth retrieval.

In line with this work, we found modulations of behavioral depth judgments based on object plausibility (Experiment 1, and replicated in-bore in Experiment 2). Specifically, thresholds for judging implausible triangles

were lower than those for plausible triangles in both the SNR and feature tasks (although to a lesser extent in the latter task). That is, paradoxically, unfamiliarity (i.e., implausibility) translates into greater ease for discerning depth position from noise and discriminating fine feature differences. This suggests that object recognition or form analysis plays a robust role in the encoding of stereoscopic information.

At first glance, our findings appear somewhat at odds with those of previous studies that adopted priming paradigms and tests of agnostic patients, showing object implausibility is associated with worsened memory (Soldan, Mangels, & Cooper, 2006; Schacter, Cooper, Tharan, & Rubens, 1991) or even perceptual performances (Freud et al., 2017; Freud, Ganel, & Avidan, 2015). However, we note that it has been shown that participants tend to expect plausible objects even after being primed by implausible objects (Soldan et al., 2006; Schacter et al., 1991), rendering priming data difficult to interpret. Critically, in all of these studies (Freud et al., 2017; Freud, Ganel, et al., 2015; Soldan et al., 2006; Schacter et al., 1991), stimuli were 2-D drawings implying depth form, rather than stereo-defined depth forms, restraining our ability to extrapolate from their conclusions to stereovision mechanisms.

Still, our findings, despite being robust and replicable, are somewhat counterintuitive. Based on our data, we believe that daily exposure to plausible objects results in strengthened representations that are easier to retrieve, without engaging in detailed and computationally expensive analyses of object features. To better understand this, we need to first consider the mechanisms engaged during the completion of our two tasks: noise filtering and signal (disparity) readout (Doshier & Lu, 2005), both of which is needed for the SNR task and the latter of which is needed for the feature task. In principle, object identity can act on either of these mechanisms; that is, either (1) noise filtering or (2) signal readout may be better for unnatural, novel objects as compared with everyday relevant objects. In light of the consistent findings across both of our tasks, including the feature task that does not require external noise filtering, we consider it more likely that object plausibility affects readout mechanisms. Specifically, we posit that object familiarity acts to downweigh the importance of disparity computations for retrieving the object's representation. That is, object representations of familiar objects are more readily retrievable from higher order mechanisms, and therefore, detailed disparity computations are not as crucial. Meanwhile, unfamiliar objects do not have readily retrievable representations; therefore, they may require more extensive computations of disparity to retrieve curvature and positional detail. Critically, this would entail an interaction between disparity readout mechanisms and object representation mechanisms.

In Experiment 2, the implication of both early and extrastriate areas for our stimuli broadly aligns with findings

from earlier neurophysiological readings and recent neuroimaging studies: V1 (Parker, 2007; Cumming & Parker, 1997), V2 (Thomas, Cumming, & Parker, 2002; von der Heydt, Zhou, & Friedman, 2000), V3 (Nasr, Polimeni, & Tootell, 2016; Preston et al., 2008; Tsao et al., 2003), V3A (Ban & Welchman, 2015; Anzai, Chowdhury, & DeAngelis, 2011; Preston et al., 2008; Chandrasekaran et al., 2007; Tsao et al., 2003; Backus, Fleet, Parker, & Heeger, 2001), V3B (Orban et al., 2006; Brouwer, van Ee, & Schwarzbach, 2005), and V7 (Georgieva et al., 2009; Preston et al., 2008; Chandrasekaran et al., 2007; Backus et al., 2001). In particular, V3A and V7 have been shown to be differentially responsive to stereo-defined shapes (Gilaie-Dotan, Ullman, Kushnir, & Malach, 2002; Mendola, Dale, Fischl, Liu, & Tootell, 1999), curvatures, and 3-D convex objects (Georgieva, Peeters, Kolster, Todd, & Orban, 2009). That hMT+ can discriminate between stereo-defined shapes (Figure 4C) is not surprising as previous macaque work has revealed that disparity-sensitive neurons in MT respond to stimuli/task demands very similar to those in our present study (i.e., coarse depth position; Chowdhury & DeAngelis, 2008). Additionally, hMT+ has been shown to be engaged in the analysis of object shapes, showing particularly enhanced responses when objects are rendered in stereo (Kourtzi, Bühlhoff, Erb, & Grodd, 2002). Finally, LOC has been shown to respond to correlated RDS and is able to discriminate the sign of binocular disparity (Preston et al., 2008; Welchman, 2011).

Critically, we observed modulations of fMRI responses to the depth stimuli as a function of physical plausibility. Our results indicated that patterns for the plausible and implausible triangle stimuli could be reliably distinguished by V1, V2, V3, V3A, V3B, V7, and LOC. Interestingly, response patterns for plausible and implausible cube stimuli could be similarly distinguished by V2, V3, V3A, V3B, and V7, but not by LOC, although LOC appears to be prominently able to decode plausibility under a cross-training scenario (i.e., when the SVM was trained with the alternative object data). Considering the corresponding behavioral observations of an apparent lack of distinguishability between the cubes, it is appealing to speculate as to whether LOC may play a key role in establishing plausibility-related contextual modulations of depth judgments. Indeed, the data here are interesting for the reason that object plausibility seems to be prominently encoded by the visual system under task conditions that do not require object identity to be evaluated (i.e., during judgments of depth). These data offer some reconciliation for our behavioral observations that the object's identity (plausibility) appears to be very much relevant for determining coarse target position (SNR task) and fine depth differences (feature task). We note here that we are not assuming explicit judgments of object plausibility on the part of the observer (or perhaps not even of the brain, in the case of the fMRI data). Indeed, the observer (and our brain) could be tagging an implausible object simply as "ambiguous" if its geometry cannot be readily reconciled.

However, as we did not test observers under a more explicit task question using these same stimuli (e.g., requiring explicit judgments of object plausibility), we cannot rule out the possibility that the more “ambiguous” stimulus is in fact not the implausible object, but rather the plausible one—specifically the Necker cube, which may elicit bistable interpretations. Here in the current study, stimuli were presented for a brief 500 msec and were defined in terms of stereoscopic disparity. Previous work investigating bistable perception using 2-D stimuli reported that “continuous, prolonged viewing of an ambiguous stimulus” is required for initiating perceptual fluctuations (Leopold, Wilke, Maier, & Logothetis, 2002). For work involving stereoscopic stimuli, Erkelens (2012) reported that disparity-defined Necker cubes were able to elicit bistable interpretations, but notably, during much longer (3-min) stimulus presentations. Importantly, this work has reported perceptual switches to occur on the scale of 7.5–7.7 sec (Erkelens, 2012). This suggests a perceptual change within our 500 msec is unlikely. Nonetheless, as we did not explicitly probe the bistability of percepts of our stimuli, we are unable to definitively rule out a potential role for this factor for the present observations. Nonetheless, our data appear to indicate that the plausibility manipulation of the objects matters. Although this alone is not a surprise in light of previous demonstrations by others (Freud, Ganel, et al., 2015; Freud, Hadad, Avidan, & Ganel, 2015; Freud, Ganel, & Avidan, 2013) of the brain’s sensitivity to object plausibility, the intriguing part of our data lay in the fact that the plausibility manipulation cannot be ignored by the observer: Even when stereo content and magnitude are equated, depth retrieval seems to be modulated by the configuration of this information (i.e., plausible vs. implausible; Experiment 1). Our fMRI data further show the potential neural correlates of these observed behavioral modulations.

It is further interesting to put our findings more broadly in the context of neuroimaging work in object perception and, specifically, work involving the use of plausible and implausible objects (Freud, Hadad, et al., 2015; Freud, Rosenthal, et al., 2015; Freud et al., 2013). In one study, Freud et al. (2013) found adaptation effects during same/different judgments of plausible and implausible objects in LOC, mid fusiform gyrus, and posterior fusiform. Importantly, they found a significant correlation between the MR adaptation effect and behavioral RTs for plausible objects in LOC only. This work suggests a substantial role of LOC for discerning stimulus plausibility. In a follow-up behavioral study, participants were instructed to report whether two red circles on a plausible or implausible object were inside the boundaries of the object or not. Results indicated higher accuracies for distinguishing plausible and implausible line drawings of objects when the stimuli were presented at long stimulus durations (986 msec) as compared with short durations (85 msec; Freud, Hadad, et al., 2015), further implying that differentiation of object plausibility occurs in higher-order stages.

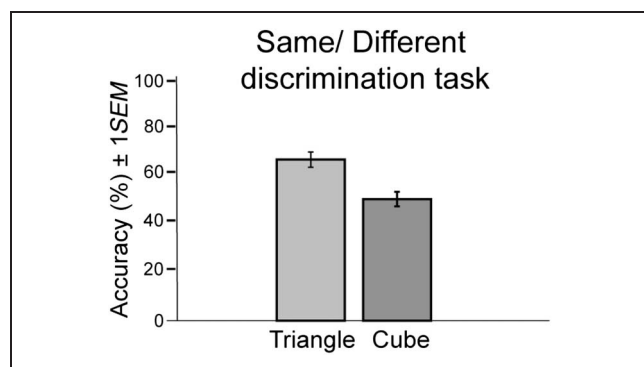
Here, we put forward a different question. Although we were not interested in object-level representations of plausible versus implausible objects per se, we asked how object plausibility, when not explicitly probed, is able to modulate retrieval of stereoscopic depth information. Our finding is surprising, as it is not immediately apparent as to why object identity should matter at all when the visual system is asked to assess a target’s depth position. More broadly, it has been shown that object-directed action, requiring engagement of premotor and inferior parietal cortex, facilitates object recognition (a ventral function) relative to animal stimuli and semantic stimuli without accompany action (Mahon et al., 2007; Helbig, Graf, & Kiefer, 2006). Data acquired from schizophrenic patients also seem to suggest that dorsal stream involvement is critical for successful object recognition (Sehatpour et al., 2010; Doniger, Foxe, Murray, Higgins, & Javitt, 2002). More relevant, other work has shown extensive cross-talk between shape-related regions and lower visual areas. For example, multiunit recording work reported reductions in activity of V1/V2 neurons with increases in collinearity (Cardin, Friston, & Zeki, 2010). In a separate line of work, Murray et al. (2002) demonstrated that human V1 BOLD responses decreased, whereas responses in higher visual areas increased during the grouping of stimuli into meaningful representations. This is congruent with work from an earlier study that found reduced BOLD responses in LOC together with a significant increase in V1 responses as participants viewed a series of progressively scrambled images (Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001). Taken together, the above suggests that object identity, or at least coherency, could potentially modulate earlier visual encoding, likely through feedback mechanisms arising from higher ventral cortex. We interpret our data to suggest an extensive intertwining of object processing mechanisms, known to involve extrastriate ventral areas along inferior temporal cortex and LOC, and disparity mechanisms in intermediate and dorsal extrastriate cortex. Indeed, there are well-documented connections between ventral cortex and dorsal visual areas (e.g., V3A to temporal-occipital area TEO [Webster, Bachevalier, & Ungerleider, 1994], anterior intraparietal area connections to inferior temporal cortex [Borra et al., 2007]).

Since the lower level differences between the two variations of each stimulus class were controlled for, including ensuring the beam width, surface area, and overall disparity equivalent across the two variations of the objects, and stimulus difficulty among the different stimuli were otherwise equated by sampling test values around subject-specific thresholds, we believe that the differences in neural response patterns are reflecting the presence or absence of structural violations in the stimuli (i.e., plausibility) rather than reflecting lower order task-irrelevant differences or higher order discrepancies in terms of attentional engagement. However, if object plausibility has such robust effects on disparity computations, why did not we observe the same robust

differences in behavioral and neural responses for the cube variations? In the follow-up control task, we asked participants to judge whether a pair of consecutively presented disparity-defined objects were identical or not. We found that accuracies for discriminating between same and different plausible and implausible cubes were consistently at chance (Figure 6). This lack of discriminability, paired with the lack of behavioral threshold differences between these same cubes, further impart the importance of object recognition and form-related mechanisms in disparity-based depth calculations. Interestingly, LOC is sensitive to global shapes but not to locally disconnected edges or surfaces (Vinberg & Grill-Spector, 2008; Grill-Spector & Malach, 2004; Grill-Spector, Kourtzi, & Kanwisher, 2001). Crucially, implausible objects have valid local cues but not coherent global structures for assisting differentiation of object type. The lack of differences (both behaviorally and in terms of MR patterned responses in LOC) in response to the two cubes further suggests that our observations cannot be merely explained by differences in local cues, as this information is as much, if not more pronounced, between the two cube stimuli, as compared with the two triangle stimuli.

### Stimulus Generalizability

In light of the above, in a second follow-up experiment (Supplementary Experiment 2), we replicated the behavioral paradigms while retaining the original triangle stimuli (as used in Experiments 1 and 2), in addition to two other novel object classes drawn from Freud, Rosenthal, et al. (2015). Critically, we found again significantly lower thresholds for implausible objects as compared with their plausible counterparts for both the SNR and feature tasks. This finding shows strong generalizability of plausibility-based modulations of disparity perception



**Figure 6.** Behavioral performance on a control, plausibility discrimination task (Supplementary Experiment 1). Participants completed a total of 80 trials. Half of the trials were of the same object variation (i.e., both plausible), and the remaining half were different (i.e., one plausible and one implausible). Performance is presented here in terms of discrimination accuracies (%). Error bars represent  $\pm 1$  SEM ( $n = 10$ ).

across different exemplars and adds strength to the effects demonstrated in the two main experiments.

Still, data from both Experiment 1 and this supplementary experiment suggest that certain exemplars seem to work better than others in terms of yielding modulations of depth judgments. It is therefore important to acknowledge potential roles of factors beyond stimulus plausibility that may contribute to depth judgments. Indeed, data from our plausibility discrimination task (Supplementary Experiment 1 control task) suggest that stimulus complexity or shape may be two such factors. It is further important to note that, by definition, manipulation of object plausibility is conflated here with global versus local incongruities (at least in terms of Euclidean geometry), as well as familiarity (indeed, a plausible object should be more familiar than an implausible one), both of which may be important for depth retrieval and worth teasing out in future work.

Moving forward, it may also be interesting to investigate the temporal dynamics of responses to stereo-defined 3-D shapes, which may shed light on how potential object disparity interactions may be achieved. Although fMRI lacks the temporal resolution to probe this question using the present data set, we also had not measured meaningful behavioral RTs (in all cases, responses were permitted only following the offset of the stimulus). For 2-D (motion- and luminance-defined objects), Jiang et al. (2008) have reported LOC as a locus for assembling object features. Magnetoencephalographic data showed that cortical engagement to their stimuli proceeded from early visual areas, then to the LOC, followed by ventral temporal areas, and lastly parietal cortex and early visual areas in parallel, all within a time frame of 500 msec. Whether a broadly similar temporal order of engagement occurs for stereoscopically defined 3-D objects will require further investigation but would be particularly fascinating to probe in the context of differences in temporal responses between plausible and implausible objects.

Finally, as mentioned, LOC appears to be essential for the perception of object plausibility (Freud, Ganel, et al., 2015; Freud, Hadad, et al., 2015; Freud et al., 2013). Although it is may be reasonable to ask whether a more extensive network of regions, including prefrontal or frontal-parietal regions are tapped into during the viewing of objects with differing plausibility, their implication, even under scenarios where the task requires explicit attention to object identity, are not apparent. Freud et al. (Freud, Ganel, et al., 2015; Freud, Hadad, et al., 2015) focused on the object-selective network that consisted of the LOC, inferior temporal sulcus, transverse occipital sulcus, intraparietal sulcus, and posterior fusiform, but did not report responses in prefrontal or parietal regions. Returning to our present data, our SVM results, and in particular the searchlight analysis that should be able to detect any relevant regions not captured by the selected ROIs, also did not indicate additional involvement of these areas.

Using behavioral and fMRI paradigms, we showed here that both behavioral judgments of depth position and

corresponding neural responses are modulated by the physical plausibility of complex 3-D objects. We conjecture that our observations of lower disparity sensitivity toward plausible objects may be explained by down-weighted disparity computations (readouts) that scale as a function of object familiarity during object recognition, which is then consequently reflected in modulations in accuracies for depth position judgments. The current study challenges our current understanding of the disparity processing network in the primate brain, signaling a need to recognize the potential for dynamically changing mechanisms that closely interact within the larger object recognition problem. Such work may come in the form of more rigorous manipulations of object familiarity and observations of whether such manipulations result in systematic changes in depth judgments.

## Note

1. Supplementary material for this paper can be retrieved from <https://changelab.hk/2019jocn>.

## Acknowledgments

This work was supported by an Early Career Scheme grant (Research Grants Council, 27612119) and the Foreign Researcher Invitation Program, National Institute of Information and Communications Technology, Japan, to D. C. and JSPS KAKENHI (17H04790 and 17K20021) and ERATO (JPMJER1801) to H. B. N. W. and D. C. contributed to stimulus and design development. All authors were involved in data collection, analyses, and the preparation of this manuscript.

Reprint requests should be sent to Dorita H. F. Chang, Department of Psychology, Jockey Club Tower, Centennial Campus, the University of Hong Kong, Hong Kong, or via e-mail: [changd@hku.hk](mailto:changd@hku.hk).

## REFERENCES

- Anzai, A., Chowdhury, S. A., & DeAngelis, G. C. (2011). Coding of stereoscopic depth information in visual areas V3 and V3A. *Journal of Neuroscience*, *31*, 10270–10282.
- Autodesk. (2018). *Autodesk inventor professional [Computer software]*. Los Angeles: Author.
- Backus, B. T., Fleet, D. J., Parker, A. J., & Heeger, D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, *86*, 2054–2068.
- Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, *20*, 8188–8198.
- Ban, H., & Welchman, A. E. (2015). fMRI analysis-by-synthesis reveals a dorsal hierarchy that extracts surface slant. *Journal of Neuroscience*, *35*, 9823–9835.
- Blake, A., & Bühlhoff, H. (1990). Does the brain know the physics of specular reflection? *Nature*, *343*, 165–168.
- Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., et al. (2007). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cerebral Cortex*, *18*, 1094–1111.
- Brainard, D. H. (1997). The Psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Brouwer, G. J., van Ee, R., & Schwarzbach, J. (2005). Activation in visual cortex correlates with the awareness of stereoscopic depth. *Journal of Neuroscience*, *25*, 10403–10413.
- Bülhoff, I., Bülhoff, H., & Sinha, P. (1998). Top-down influences on stereoscopic depth-perception. *Nature Neuroscience*, *1*, 254–257.
- Cardin, V., Friston, K. J., & Zeki, S. (2010). Top-down modulations in the visual form pathway revealed with dynamic causal modeling. *Cerebral Cortex*, *21*, 550–562.
- Chandrasekaran, C., Canon, V., Dahmen, J. C., Kourtzi, Z., & Welchman, A. E. (2007). Neural correlates of disparity-defined shape discrimination in the human brain. *Journal of Neurophysiology*, *97*, 1553–1565.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*, 27.
- Chang, D. H. F., Kourtzi, Z., & Welchman, A. E. (2013). Mechanisms for extracting a signal from noise as revealed through the specificity and generality of task training. *Journal of Neuroscience*, *33*, 10962–10971.
- Chang, D. H. F., Mevorach, C., Kourtzi, Z., & Welchman, A. E. (2014). Training transfers the limits on perception from parietal to ventral cortex. *Current Biology*, *24*, 2445–2450.
- Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*, 478–484.
- Chowdhury, S. A., & DeAngelis, G. C. (2008). Fine discrimination training alters the causal contribution of macaque area MT to depth perception. *Neuron*, *60*, 367–377.
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage*, *40*, 353–366.
- Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, *389*, 280–283.
- De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., & Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *Neuroimage*, *43*, 44–58.
- Doniger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., & Javitt, D. C. (2002). Impaired visual object recognition and dorsal/ventral stream interaction in schizophrenia. *Archives of General Psychiatry*, *59*, 1011–1020.
- Dosher, B. A., & Lu, Z. L. (2005). Perceptual learning in clear displays optimizes perceptual expertise: Learning the limiting process. *Proceedings of the National Academy of Sciences, U.S.A.*, *102*, 5286–5290.
- Erkelens, C. J. (2012). Contribution of disparity to the perception of 3D shape as revealed by bistability of stereoscopic Necker cubes. *Seeing and Perceiving*, *25*, 561–576.
- Freud, E., Ganel, T., & Avidan, G. (2013). Representation of possible and impossible objects in the human visual cortex: Evidence from fMRI adaptation. *Neuroimage*, *64*, 685–692.
- Freud, E., Ganel, T., & Avidan, G. (2015). Impossible expectations: fMRI adaptation in the lateral occipital complex (LOC) is modulated by the statistical regularities of 3D structural information. *Neuroimage*, *122*, 188–194.
- Freud, E., Ganel, T., Shelef, I., Hammer, M. D., Avidan, G., & Behrmann, M. (2017). Three-dimensional representations of objects in dorsal cortex are dissociable from those in ventral cortex. *Cerebral Cortex*, *27*, 422–434.
- Freud, E., Hadad, B. S., Avidan, G., & Ganel, T. (2015). Evidence for similar early but not late representation of possible and impossible objects. *Frontiers in Psychology*, *6*, 94.
- Freud, E., Rosenthal, G., Ganel, T., & Avidan, G. (2015). Sensitivity to object impossibility in the human visual cortex:

- Evidence from functional connectivity. *Journal of Cognitive Neuroscience*, *27*, 1029–1043.
- Georgieva, S., Peeters, R., Kolster, H., Todd, J. T., & Orban, G. A. (2009). The processing of three-dimensional shape from disparity in the human brain. *Journal of Neuroscience*, *29*, 727–742.
- Gilaie-Dotan, S., Ullman, S., Kushnir, T., & Malach, R. (2002). Shape-selective stereo processing in human object-related visual areas. *Human Brain Mapping*, *15*, 67–79.
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage*, *9*, 416–429.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677.
- Helbig, H. B., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, *174*, 221–228.
- Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*, 7195–7205.
- Jiang, Y., Boehler, C. N., Nönnig, N., Düssel, E., Hopf, J. M., Heinze, H. J., et al. (2008). Binding 3-D object perception in the human visual cortex. *Journal of Cognitive Neuroscience*, *20*, 553–562.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Cambridge, MA: MIT Press.
- Kourtzi, Z., Bühlhoff, H. H., Erb, M., & Grodd, W. (2002). Object-selective responses in the human motion area MT/MST. *Nature Neuroscience*, *5*, 17–18.
- Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: Monkey and human fMRI studies. *Neuron*, *37*, 333–346.
- Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*, 605–609.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex*, *11*, 287–297.
- Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces: The role of race categories. *Journal of Experimental Psychology: General*, *135*, 501–512.
- Mahon, B. Z., Milleville, S. C., Negri, G. A., Rumiati, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, *55*, 507–520.
- Marlow, P. J., & Anderson, B. L. (2015). Material properties derived from three-dimensional shape representations. *Vision Research*, *115*, 199–208.
- Marlow, P. J., Todorović, D., & Anderson, B. L. (2015). Coupled computations of three-dimensional shape and material. *Current Biology*, *6*, R221–R222.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 8560–8572.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 15164–15169.
- Nasr, S., Polimeni, J. R., & Tootell, R. B. (2016). Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. *Journal of Neuroscience*, *36*, 1841–1857.
- Norman, J. F., Todd, J. T., & Orban, G. A. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science*, *15*, 565–570.
- Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2008). Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision*, *8*, 13.
- Orban, G. A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J. T., et al. (2006). Mapping the parietal cortex of human and non-human primates. *Neuropsychologia*, *44*, 2647–2667.
- Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., et al. (2003). Similarities and differences in motion processing between the human and macaque brain: Evidence from fMRI. *Neuropsychologia*, *41*, 1757–1768.
- Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, *8*, 379–391.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Preston, T. J., Li, S., Kourtzi, Z., & Welchman, A. E. (2008). Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. *Journal of Neuroscience*, *28*, 11315–11327.
- Rossetti, Y., & Pisella, L. (2002). Several “vision for action” systems: A guide to dissociating and integrating dorsal and ventral functions (tutorial). In W. Prinz, & B. Hommel (Eds.), *Attention and performance: Common mechanisms in perception and action* (pp. 62–119). Oxford: Oxford University Press.
- Schacter, D. L., Cooper, L. A., Tharan, M., & Rubens, A. B. (1991). Preserved priming of novel objects in patients with memory disorders. *Journal of Cognitive Neuroscience*, *3*, 117–130.
- Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, *8*, 522–535.
- Sehatpour, P., Dias, E. C., Butler, P. D., Revheim, N., Guilfoyle, D. N., Foxe, J. J., et al. (2010). Impaired visual object processing across an occipital-frontal-hippocampal brain network in schizophrenia: An integrated neuroimaging study. *Archives of General Psychiatry*, *67*, 772–782.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.
- Shiozaki, H. M., Tanabe, S., Doi, T., & Fujita, I. (2012). Neural activity in cortical area V4 underlies fine disparity discrimination. *Journal of Neuroscience*, *32*, 3830–3841.
- Soldan, A., Mangels, J. A., & Cooper, L. A. (2006). Evaluating models of object-decision priming: Evidence from event-related potential repetition effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 230–248.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-Dimensional proportional system: An approach to cerebral imaging*. New York: Thieme.
- Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience*, *5*, 472–478.
- Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 663–678.
- Tsao, D. Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., et al. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron*, *3*, 555–568.
- Uka, T., & DeAngelis, G. C. (2006). Linking neural representation to function in stereoscopic depth perception:

- Roles of the middle temporal area in coarse versus fine disparity discrimination. *Journal of Neuroscience*, *26*, 6791–6802.
- Uka, T., Tanabe, S., Watanabe, M., & Fujita, I. (2005). Neural correlates of fine depth discrimination in monkey inferior temporal cortex. *Journal of Neuroscience*, *25*, 10796–10802.
- Vinberg, J., & Grill-Spector, K. (2008). Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *Journal of Neurophysiology*, *99*, 1380–1393.
- von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, *15*, 1955–1967.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120.
- Watt, S. J., & Bradshaw, M. F. (2002). Binocular information in the control of prehensile movements in multiple-object scenes. *Spatial Vision*, *15*, 141–155.
- Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cerebral Cortex*, *4*, 470–483.
- Welchman, A. E. (2011). Decoding the cortical representation of depth. In J. Trommershauser, K. Kording, & M. S. Landy (Eds.), *Sensory cue integration* (pp. 345–367). New York: Oxford University Press.
- Welchman, A. E. (2016). The human brain in depth: How we see in 3D. *Annual Review of Vision Science*, *2*, 345–376.
- Witzel, C., Valkova, H., Hansen, T., & Gegenfurtner, K. R. (2011). Object knowledge modulates colour appearance. *i-Perception*, *2*, 13–49.