University of Nevada, Reno

Size coding of real-world objects in human dorsal cortex

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Psychology

by

Desiree E. Holler

Dr. Jacqueline C. Snow/Thesis Advisor

December, 2017

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

DESIREE EVE HOLLER

Entitled

Size Coding Of Real-World Objects In Human Dorsal Cortex

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

Jacqueline C. Snow, Advisor

Jeffrey Hutsler, Committee Member

Fang Jiang, Committee Member

Benjamin Young, Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

December, 2017

ABSTRACT

Neural object representations were discovered in human posterior parietal cortex (PPC) almost a decade ago [1] yet the functional contribution of this area to object perception remains poorly understood [2]. Classically, visual object coding in human ventral and dorsal cortex has been studied using images [3-7], but two-dimensional (2D) images do not convey exact egocentric distance or size information, and neither 2D nor three-dimensional (3D) stereoscopic images of objects afford grasping and interaction. Unlike ventral cortex, object representations in PPC are especially sensitive to 3D shape [8, 9], and are proximal to hand orientation and pre-shaping areas [10-12] and to egocentric spatial maps of the visual field [13], making these representations uniquely positioned to code the size and shape of graspable objects. Here, we isolated the contribution of dorsal cortex to size and shape processing by studying object recognition in patients with visual form agnosia following bilateral damage to shape-processing areas of ventral cortex. The patients were asked to identify objects, whose size was congruent or incongruent with typical real-world size (RWS), presented in different display formats (real objects, 2D and 3D images). While recognition of 2D and 3D images was extremely poor, real object recognition was surprisingly preserved, but only when physical size reflected RWS. Analogous display format and size manipulations had no effect on recognition of basic geometric shapes that lacked RWS associations. These findings reveal, for the first time, size-coding of real objects in dorsal cortex and highlight the importance of using real-world stimuli to understand perception.

Table of Contents

List of Figures

INTRODUCTION

Patients D.F. and J.W. have unique bilateral lesions of ventral occipito-temporal cortex resulting in visual form agnosia, a severe impairment in object recognition with comparative sparing of early visual processes and stereoscopic acuity [14]. D.F.'s brain lesion involves predominantly the lateral occipital (LO) cortex, bilaterally [15, 16]. LOC is known to play a critical role in the visual processing of object shape [1, 6, 17], and, consistent with her severe recognition deficit, D.F. evinces no neural responses to 2D greyscale images of objects, as measured using fMRI [15, 16]. J.W. has bilateral lesions to the ventral occipital lobes (posterior to LOC) –areas that provide critical inputs to LOC, and severe deficits in shape recognition especially when color cues are absent [18, 19] (more detailed imaging is precluded because of J.W.'s pacemaker; see **Methods**). Although D.F. has reduced cortical thickness along the posterior intraparietal sulcus [15, 16], dorsal cortex is largely intact in both patients. Given that single-cell recording studies in monkeys [13, 20-24] and neuroimaging studies in humans [1, 9, 25-32] have revealed non-action-based object-selective representations in PPC within the dorsal processing stream, here we examine the role of PPC in object recognition in patients with extensive ventral lesions.

In a series of experiments, we measured the ability of D.F. and J.W. to recognize everyday objects presented in different Display Formats: real

tangible objects versus high-resolution 2D and/or 3D stereoscopic computerized images of the same items. The images were matched closely to their real-world counterparts for size, orientation, background and illumination.

Critically, we also manipulated the Size of the stimuli in each display format relative to the typical RWS. All stimuli were achromatic to prevent the use of color to bootstrap object processing [16, 18]. We predicted that if object representations in PPC are uniquely sensitive to the size and shape of graspable objects, then agnosia patients should show a recognition advantage for real objects (i.e., a 'real object advantage', ROA) [33-36] compared to 2D and 3D images of the same exemplars. The ROA should be disrupted, however, when the physical size of real objects deviates from typical real-world-size. Size manipulations should have little influence on identification of images (which do not afford grasping), or of basic geometric shapes that lack strong RWS associations

MATERIALS AND METHODS

Participants

Patient D.F.

D.F. is a right-handed native English speaking female who was born in 1954 (age 61 at time of testing). Following toxic exposure to carbon monoxide at the age of 34, D.F. suffers from a profound visual form agnosia characterized by a severe impairment in object recognition that cannot be reduced to a simple visual sensory deficit. An anatomo-functional examination of D.F.'s brain by Bridge et al. [15] using high-resolution MRI revealed a substantial reduction in cortical thickness across occipital cortex that was most severe in the lateral occipital complex (LOC) bilaterally. Cortical thickness was also reduced in the posterior IPS bilaterally, and the anterior portion of V1 in the LH, resulting in a partial quadrantanopia in the upper right visual field from ~30deg eccentricity. White matter connections between LOC and other visual areas were also reduced. D.F.'s recognition is poorest for silhouettes and black and white line-drawings, but (as in other patients with similar lesions) her performance improves with the addition of color and surface texture cues [35, 51]. In line with the damage to D.F.'s LOC and its connections, she shows no object-selective response to black-and-white images of objects as measured by fMRI [15, 16].

The earliest formal tests of D.F.'s recognition abilities were conducted 3

months after the accident [51]. Although D.F. made some errors in naming everyday objects (e.g., a fork was misidentified as a knife) recognition was profoundly disrupted on standardized neuropsychological 'paper and pencil' tests of object perception. For example, D.F. had difficulty identifying line-drawings of objects, with her performance being equivalent to a mental age of 2 years 2 months. At 4 months post-accident, administration of the Peabody Picture Vocabulary Test, which involves pointing to one of four line-drawings in response to a verbal cue, also revealed severe deficits, and there was little improvement 4 months later [51]. D.F.'s impairment in object recognition has remained relatively stable in the years post injury. The beneficial effect of color is most apparent for natural objects whose surface color is highly diagnostic of object identity [35]. Importantly, several studies have reported that D.F. is better at recognizing real- world tangible objects than photographs of objects –a ROA [33-36].

Although D.F. can use color to facilitate recognition in some cases [15, 16], fMRI responses to colored (versus phase-scrambled) images of objects around D.F.'s left LOC are small and fragmented (and none survive cluster-size correction), and no detectable object-selective responses are apparent in the RH [15]. James et al. [16] found no object-selective fMRI activation in D.F. in the vicinity of LOC to colored and greyscale object images, and critically, no object- selective fMRI response whatsoever to black and white line-drawings of objects. Images of colored and greyscale objects activate D.F.'s

ventromedial temporal cortex [16], in areas that also respond to colored Mondrian patterns (versus black and white patterns) in D.F. [15], and areas in healthy observers that respond selectively to color and surface texture [52]. Given D.F.'s ability to use color (and possibly surface texture) to facilitate her recognition performance, the stimuli used in the current experiment were either coated in white paint, or 3D printed using white plastic (see *Stimuli*). Detailed autobiographical information and further information about other aspects of D.F.'s early clinical neuropsychological profile are available in other published studies [51, 53].

Patient J.W.

JW is a right-handed native English speaking male who was born in 1961 and was 55-56 years over the course of this testing. Like D.F., JW suffers from a profound visual form agnosia, characterized by a severe impairment in object recognition that cannot be reduced to a simple visual sensory deficit, following a lesion resulting from an episode of anoxic encephalopathy associated with cardiac abnormality when he was 35 years old. CT scans obtained shortly after the insult revealed an extensive ischemic wrap around the occipital pole with widespread V2 and V1 injury and with no damage to higher-order visual cortex. We were unable to obtain more recent imaging data from JW, due to a non-fMRI compatible cardiac pacemaker. J.W.'s recognition is poor for silhouettes and black and white line-drawings, but, as for D.F., his performance improves with the addition of color and surface texture cues [18, 54]. Importantly, J.W.

also has intact stereo vision and course shape recognition (circle versus square) [19].

General Methods

Stimuli and Apparatus

The stimuli in all experiments were real-world objects and matched high-resolution 2D and/or 3D stereoscopic computerized images of the same objects. In all experiments, the real objects were presented on a black pedestal positioned in front of an LCD computer monitor (RGB: 0,0,0). The objects were presented at the same canonical orientation on each trial. The real objects used in each experiment are described below:

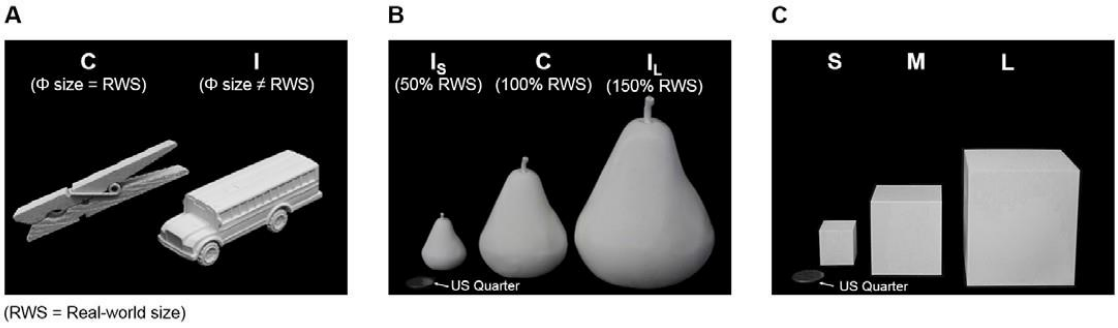

Figure 1: Stimuli used in Experiments 1-3.

(A) Stimuli in Experiment 1: Half were exemplars whose physical size (Φ) was congruent (**C**) with typical real-world size (RWS); the remainder were 'miniaturized' items whose size was incongruent (**I**) with RWS.

(B) Stimuli in Experiment 2: 3D-printed objects whose size was scaled to be smaller than (50%, **IS**), congruent with (100%, **C**), or larger than (150%, **IL**), typical RWS.

(C) Stimuli in Experiment 3: Basic geometric shapes without strong RWS associations.

Size-Congruent (C) vs. Size-Incongruent (I) Objects: The stimuli were 140 real objects: 70 were C (e.g., clothespin, spoon, key), and the remainder were I (e.g., bus, horse, table) (**Figure 1A**). Stimulus size, measured by mean number of pixels in the 2D photographs, was matched between the C and I stimulus sets ($t(69)=0.92$, $p=0.82$).

Size-Congruent (C) vs. Size-Incongruent Small (IS) and Large (IL) Objects: For Patient D.F. the stimuli were 21 everyday object items, each of which was 3D-printed in three different sizes, yielding a total of 63 stimuli (**Figure 1B**). For Patient J.W., the stimuli were the same 21 items used for D.F., plus an additional 49 items, for a total of 70 different exemplars. The 70 exemplars were printed in the three different sizes, yielding a total of 210 stimuli. The real objects were 3D- printed in white ABS plastic using either a Stratasys uPrint SE Plus 3D-printer, Afina H800+ or a DeltaMaker 2T. To create the objects, digital 3D-renderings of each item were generated by 3D- scanning the real-world objects using an Arctec Spider 3D Scanner. The remaining exemplars were sourced from online 3D image files. The stimuli were 3D printed so that the volume of the object matched that of the real-world exemplar (C stimuli: 100%). For size-incongruent (I) stimuli, the digital files were 3D printed so that their volume was either 50% of RWS (IS), or 150% of RWS (IL). Stimulus

volume was measured using CURA software.

Basic Geometric Shapes (**Figure 1C**)*:* For patient D.F., the stimuli were four basic geometric shapes that were 3D-printed in three different sizes, for a total of 12 stimuli. For Patient J.W. there were eight basic shapes, for a total of 24 stimuli.

To create matching 2D images of the real objects, the stimuli were photographed (mounted on the pedestal) using a Canon Rebel T2i DSLR camera with constant F-stop and shutter speed.

The camera (mounted on a tripod) was positioned so that the view of the object in the image matched the viewpoint of the real object when looking straight-ahead from the chin rest. The resulting images were cropped using Adobe Photoshop (to remove background) and re-sized so that the depicted objects matched exactly the size of the real objects. The images were displayed on the computer screen at approximately the same vertical and horizontal position as the real objects. The 2D images were displayed on a 27” ACER (G276HL) LCD monitor (60-Hz refresh rate) with a screen resolution of 1920 x 1080 pixels, controlled by an Intel Core I5 1.5-3210M GHz laptop computer (8 GB RAM). We created 3D stereo images of the objects by photographing each stimulus with a forward-facing camera (described above) positioned at 57 cm from the screen, from a horizontal distance of 3.2 cm to the left and right of

midline. The 3D stereo images were viewed binocularly through active shutter glasses (3D Vision 2, NVIDIA, USA) that present the two offset images separately to the left and right eye. The 3D imagesf were displayed on an ASUS (VG278HE) LCD monitor (120 Hz) with a screen resolution of 1920 x 1080 pixels using an Intel Core I7-4770 3.4 GHz computer (16 GB RAM). Stimulus presentation time was controlled in all experiments using computer-controlled PLATO liquid crystal occlusion glasses (Translucent Technologies, Toronto, Canada) that alternated between opaque (closed) and transparent (open) states. A chin rest was used to prevent head motion in 'head-fixed' viewing conditions. A second LCD monitor displayed to the experimenter the identity and position of the real object on the upcoming trial. Image presentation, timing of PLATO glasses, and the order of trials, was controlled using MATLAB (Mathworks, USA) and Psychtoolbox (Brainard, 1997). Verbal response times were recorded using a Logitech HD 720p webcam controlled by MATLAB.

General Procedure

At the start of each trial, the PLATO glasses opened (transparent state), and the participant named the object. After each response (or if > 60 sec elapsed with no response), the experimenter closed the glasses (opaque state) via a keypress. During the ITI, the experimenter either placed the next real object manually onto the display platform, or displayed a computerized image of one of the objects on the computer monitor. The inter-trial interval was 10 sec in

image blocks, and ~10 sec in real object blocks. White-noise was played during the ITI to mask information about the upcoming trial. Viewing distance from the subject to the stimulus was ~57 cm in both the real object and image blocks. The patients were instructed to name each object as quickly and accurately as possible, and that their responses would be recorded and timed. No feedback was provided.

Size-Congruent (C) vs. Size-Incongruent (I) Objects: Stimuli were presented to the patients in different Viewing Conditions (**Figure 2A**). In the '*binocular free*' condition, the real objects were viewed with full binocular vision without restricting lateral head movements, thereby providing both stereo disparity and motion parallax depth cues. In the '*binocular fixed*' condition, the objects were presented with binocular vision but the head was positioned in a chin rest, thereby providing stereo disparity but not motion parallax depth cues. In the '*monocular free*' condition, the non- dominant eye was occluded (by closing one lens of the PLATO glasses) and the head was free to move, providing motion parallax but not stereo depth cues. In the '*monocular fixed*' condition, the stimuli were viewed monocularly and head position was fixed. On *2D image trials*, the patients viewed each exemplar with head fixed in a chin rest. On *3D stereoscopic image trials* (for patient J.W.) the stimuli were viewed binocularly with head position fixed.

For patient D.F., the 140 exemplars were divided into five sets of 28 items, half of which were C, and the remainder I. The experiment was divided into five

separate testing sessions. Each session comprised five consecutive blocks of trials, in which each set of items was presented in one of the five viewing conditions. There were 140 trials per testing session (28 items per set x 5 viewing conditions), for a total of 700 trials in the experiment. The five testing sessions were conducted over two consecutive days (~12 hours total testing time), with rest breaks between sessions. Order of real object viewing conditions was counterbalanced across sessions using a Balanced Latin Square design. The Image condition was always presented last within each session. This design permits a strong test of the presence of a ROA because recognition of the images can presumably only *benefit* from having viewed previously the very same items as real world objects. The order of trials was randomized within each block. Testing sessions were conducted consecutively with rest breaks between sessions. For patient J.W., the 140 exemplars were divided into two sets of 70 items, half of which were C and the remainder I. The experiment was divided into two separate testing sessions. The first session was comprised of four consecutive blocks of trials, in which the two sets of 70 items were presented as real objects and as 3D stereoscopic images, and then vice versa, using an ABBA design. The second testing session was comprised of 2 consecutive blocks of trials in which the two sets of 70 items stimuli were presented as 2D images. There were 70 trials per testing block, for a total of 420 trials in the experiment. The testing sessions were conducted over two consecutive days (~8 hours total testing time), with rest breaks between sessions.

Size-Congruent (C) vs. Size-Incongruent Small (I_S) and Large (I_L) Objects: We attributed the recognition advantage for size-congruent objects in Experiment 1 to the preserved sensitivity to physical object size in the patients. This advantage might, however, arise from underlying differences between the stimuli in our C versus I sets: for example, the larger objects might be encountered less frequently or might be more visually complex due to feature compression. To rule out these alternative explanations, we conducted a second Experiment in which the same exemplars were presented across conditions, with size parametrically varied. This is a stringent test of size effects on recognition because the mismatch between actual versus expected size for the stimuli in Experiment 2 was orders of magnitude smaller than for the stimuli in Experiment 1. All stimuli were presented binocularly with fixed-head position.

For D.F., the 63 real objects (and 63 matching 2D images) were divided into three equal sets of 21 different exemplars. Within each set, one third of the exemplars were C, one third were I_S, and the remainder were I_L. Each set was displayed once in each display format. D.F. viewed the three different sets of exemplars as real objects first, in three consecutive blocks of trials. She then viewed the same exemplars as 2D images in three consecutive blocks of image trials, for a total of 126 trials. The order of presentation of items in each block was randomized. D.F. completed the experiment in one testing session

that took ~4 hours to complete, including rest breaks. For J.W, the 210 real objects (and 210 images) were divided into six equal sets of 35 different exemplars. Within each set, one third of the exemplars were C, one third were IS, and the remainder were IL. Each set was displayed once in each display format. Three sets of stimuli were presented to J.W. across two separate testing sessions, with the real objects presented first in each session, followed by 2D images. In each session, the sets were viewed in consecutive blocks of trials, for a total of 210 trials (420 across both sessions). The order of presentation of items in each block was randomized. J.W. took ~5 hours to complete each session, including rest breaks.

Basic Geometric Shapes:

All stimuli were presented binocularly with fixed-head position. For Patient D.F. the 12 basic geometric shapes were each presented as real objects first, and then as images, for a total of 24 trials. For Patient J.W., the 24 stimuli were presented across two separate testing sessions. In each session, patient J.W. viewed the real objects first, followed by the corresponding 2D image trials. The experiment took D.F. ~1 hr to complete, and J.W. ~ 2 hrs to complete. The order of stimuli was randomized within each testing session.

Data Analysis

For both patients, reaction times (RTs) were long and highly variable in all experiments and so our analyses focused on naming accuracy (% correct) and the proportions of errors on incorrect trials. Naming responses were analyzed by examining the digitized recordings of each trial using Audacity software. RT was measured as the point of initial inflection of the digitized sound waveform when a verbal word response was produced. Incorrect responses were defined as trials in which the object was named incorrectly (examined further in the Error analysis). Trials in which no response was produced for 60 seconds from the time of stimulus onset were defined as 'no response' trials.

For the analysis of errors, we examined the types of erroneous responses that the patients produced when the target objects were misidentified. Small Errors were defined as responses in which the target was misidentified as another object that is typically small in size (i.e., a fork, pen, whisk, or toothbrush). Conversely, Large Errors were defined as responses in which the target was misidentified as another object that (in everyday life) is typically much larger (i.e., a dresser, deer, airplane, or house). We reasoned that if the patients are more sensitive to the identity of objects whose physical size reflects RWS, then there should be a higher proportion of Small (versus Large) Error

responses for C than for I stimuli. Analyses on the accuracy and error data were performed using parametric (t-tests), and non-parametric tests (Cochran's Q, Wilcoxon, Chi-squared) in which objects were treated as the random factor in the analysis [36]. All statistical tests were two-tailed, except for the Fisher's Exact test used for the analysis of errors, which were one-tailed.

RESULTS

Recognition of Size-Congruent (C) vs. Size-Incongruent (I) Objects

In Experiment 1, we evaluated in different Display Formats (real objects, 2D or 3D) the patients' recognition of objects whose Size was either consistent with (i.e., Size Congruent: C), or orders of magnitude smaller than (i.e., Size Incongruent: I), typical RWS (**Figure 1A**). The real objects were presented in different Viewing Conditions to quantify the relative contribution of stereo disparity (binocular versus monocular) and/or motion parallax (head position free to move versus fixed in a chin rest) to perception (**Figure 2A**). In J.W., we also measured perception of 3D stereoscopic images. Stimulus onset was controlled using liquid-crystal occlusion spectacles (and active shutter glasses for the 3D displays).

Naming Accuracy

Figure 2B displays the percentage (%) of correct responses made by each patient, separately for each Display Format, Size and Viewing Condition. As

expected, both patients performed below ceiling in all conditions, consistent with a generalized impairment in shape recognition. Importantly, however, both patients evinced a clear ROA, reflected by superior recognition of real objects versus 2D images when viewed binocularly with fixed head position (D.F., p=0.002, J.W., p=0.018, Wilcoxon). For J.W., recognition of real objects was also better than 3D images of the same items (p=0.003, Wilcoxon), which was equal to 2D images (p=0.317, Wilcoxon). Critically, the patients' ability to recognize real objects depended on the physical size of the stimuli. Overall, accuracy was higher for C versus I objects (D.F., C=25.1%, I=5.2%, $\chi 2=53.14$, $p<0.001$; J.W., C=21.9%, I=8.1%; $\chi 2=15.705$, $p<0.001$). However, whereas gradual elimination of depth cues from binocular disparity and/or motion parallax interfered with recognition of C objects (D.F., Cochran's Q (4) = 18.952, $p = 0.001$; J.W., Cochran's Q (2) =10.00 $p = 0.007$), this was not so for I objects, for which performance remained close to floor across all viewing conditions (D.F., Cochran Q (4) = 0.632, p=0.959; J.W., Cochran Q (2) = 4.00, p=0.135). Similarly, within each viewing condition, the patients showed significant recognition advantages for C (vs I) stimuli when depth cues were present (D.F., real objects binocular free, binocular fixed, monocular free, all p-values ≤ 0.001, $\chi 2 \geq 11.706$; J.W., real objects binocular fixed p=0.002, $\chi 2=9.786$, 3D image p=.014, $\chi 2=6.048$). When depth cues were absent, there was no effect of Size for 2D images in (p=0.217, $\chi 2=1.522$). D.F. showed a small advantage for C vs. I objects even when depth cues were absent (real objects monocular head fixed, p=0.026, $\chi 2=4.965$; 2D images, p=0.026,

$\chi 2=4.965$), possibly because of the greater number of times D.F. (versus J.W.) saw the same objects in the other viewing conditions with depth cues available.

To quantify the effect of size-congruence on recognition accuracy, we calculated a size facilitation index for each patient, by subtracting % correct on I trials from C trials, separately for each viewing condition (**Figure 2C**). The individual facilitation indices do not lend themselves to statistical analysis. Critically, however, parametric contrasts of the average facilitation indices for D.F. and J.W. for real objects vs. 2D images (viewed with binocular fixed head position) confirmed that size-congruence facilitated recognition of real objects ($t(1)= 14.101$, $p=0.045$, $d=28.202$), but not 2D images ($t(1) = 3.728$, $p=0.167$, $d=7.456$), and that the index for real objects was significantly greater than 2D images ($t(1)=13.518$, $p=0.047$, $d=4.31$).

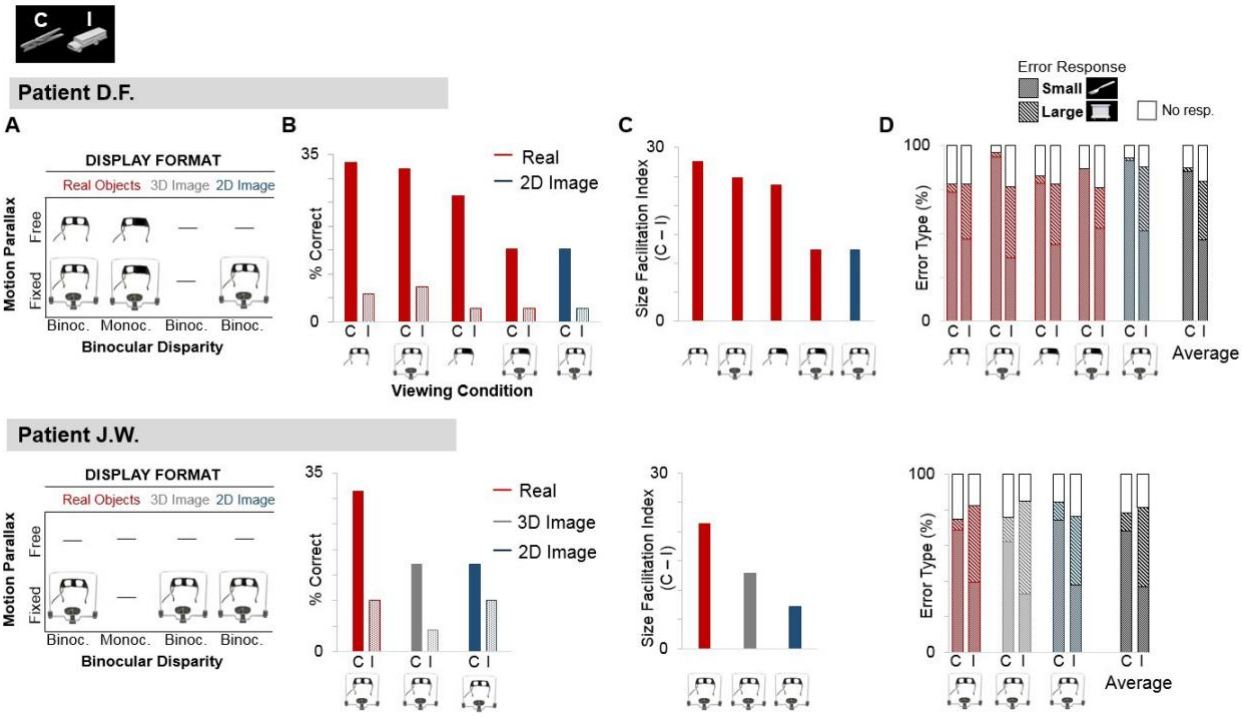

Figure 2: Recognition of Size-Congruent (C) vs. Size-Incongruent (I) Real Objects and Images

(A) The patients (D.F., *upper panel*; J.W., *lower panel*) named objects presented in different Display Formats (real objects, 3D images, 2D images). Stimuli were presented in different Viewing Conditions: with or without depth cues from motion parallax (head position free vs. fixed in chin rest) and binocularly disparity (binocular vs. monocular vision).

(B) % correct responses, shown separately for stimuli in each Display Format, Size (C, I) and Viewing Condition.

(C) Facilitatory effect of size-congruence for stimuli in each Display Format and Viewing Condition.

(D) Error analysis: % trials where target was misidentified as another 'small' or 'large' object (or no response), separately

for each Display Format and Viewing Condition. Data averaged across conditions shown far right (black bars).

Naming Errors

In addition to analyzing the data from *correct* trials, we examined the types of errors that the patients produced when the target objects were misidentified. **Figure 2D** displays the % of trials in which each patient named incorrectly C and I target objects as other objects that were either 'small' or 'large'. For example, a *clothespin* might be misidentified as another *small object* (i.e., a fork), or a *large object* (i.e., a *dresser*). If the patients are more sensitive to the identity of objects whose physical size is appropriate, then they should produce a higher proportion of Small (versus Large) Error responses for C than for I stimuli for which error type should be of equal likelihood (i.e., indicative of true guessing). Accordingly, both patients were more likely overall to produce Small Errors for C objects (D.F., probability of Small Errors=0.97, Large Errors=0.03, J.W., probability of Small Errors=0.87, Large Errors=0.13), whereas for I objects they were equally likely to produce Small and Large Errors (D.F., probability of Small Errors=0.58, Large Errors=0.42; J.W., probability of Small Errors=0.45, Large Errors=0.55), reflecting a statistically significant difference in error proportions (D.F., 0.39 ($\chi 2=101.948$, $p<0.001$); J.W., 0.42 ($\chi 2=53.178$, $p<0.000$)). Similar differences in the proportion of Small versus Large Errors for C vs. I stimuli were apparent in each viewing condition (all p-values <0.001), suggesting that patients evince sensitivity to size cues

for recognition and are constrained in the error responses accordingly.

Recognition of Size-Congruent (C) vs. Size-Incongruent Small (IS) and Large (IL) Objects

For Experiment 2, we ensured that the same everyday objects were presented by scanning and 3D-printing each exemplar in *three* different sizes: 50% of RWS (incongruent small; IS), equal to (100%) RWS (congruent; C), or larger objects that were 150% of their RWS (incongruent large; IL) (**Figure 1B**), with each exemplar serving as its own control for familiarity. If RWS is important for recognition, then incongruence costs in the patients' performance should be apparent for objects that are physically smaller *or larger* than their typical RWS (and size effects should be minimal for images).

Naming Accuracy

Figure 3A presents the patients' % correct naming responses for stimuli in each display Format (Real Objects, Images) and Size condition (C, IS, IL). Overall, D.F.'s performance was comparable for the real object and image displays (26.7% average percent correct, Cochran's Q=0.00, p=1.00; no ROA), which is unsurprising given that she saw a relatively small stimulus set (n=21) and viewed all items as real objects before 2D images. J.W., who

viewed more than three times as many exemplars (n=70), was better at recognizing real objects (12.86% correct) compared to 2D images (5.71% correct, Cochran's Q (1) =9.783, p=0.002). As in Experiment 1, there were striking format-dependent effects of Size on recognition. For real objects, recognition was

significantly better for C versus both IS and IL stimuli (IS: D.F., p=0.025, J.W., p=0.021; IL: D.F., p=0.014, J.W., p=0.008, Wilcoxon). Conversely, for 2D images, recognition was equally poor for C versus IS and IL stimuli (IS: D.F., p=.059; J.W., p=0.655; IL: D.F., p=0.257, J.W., p=0.317, Wilcoxon). Next, we calculated a size facilitation index by subtracting the patient's mean % correct on IS and IL trials from that of congruent trials (C – (IS + IL /2)) (**Figure 3B**). Parametric contrasts of the average facilitation indices for real objects versus 2D images indicated that the size facilitation effect for real objects was significantly greater than for 2D images (t(1)=14.86, p=0.042, d=0.66), although neither index was significantly greater than zero (real objects: t(1)=2.277, p=0.263, d=4.554; images: t(1)=1.24, p=0.432, d=2.48)) due to the large difference in absolute indices between the two patients.

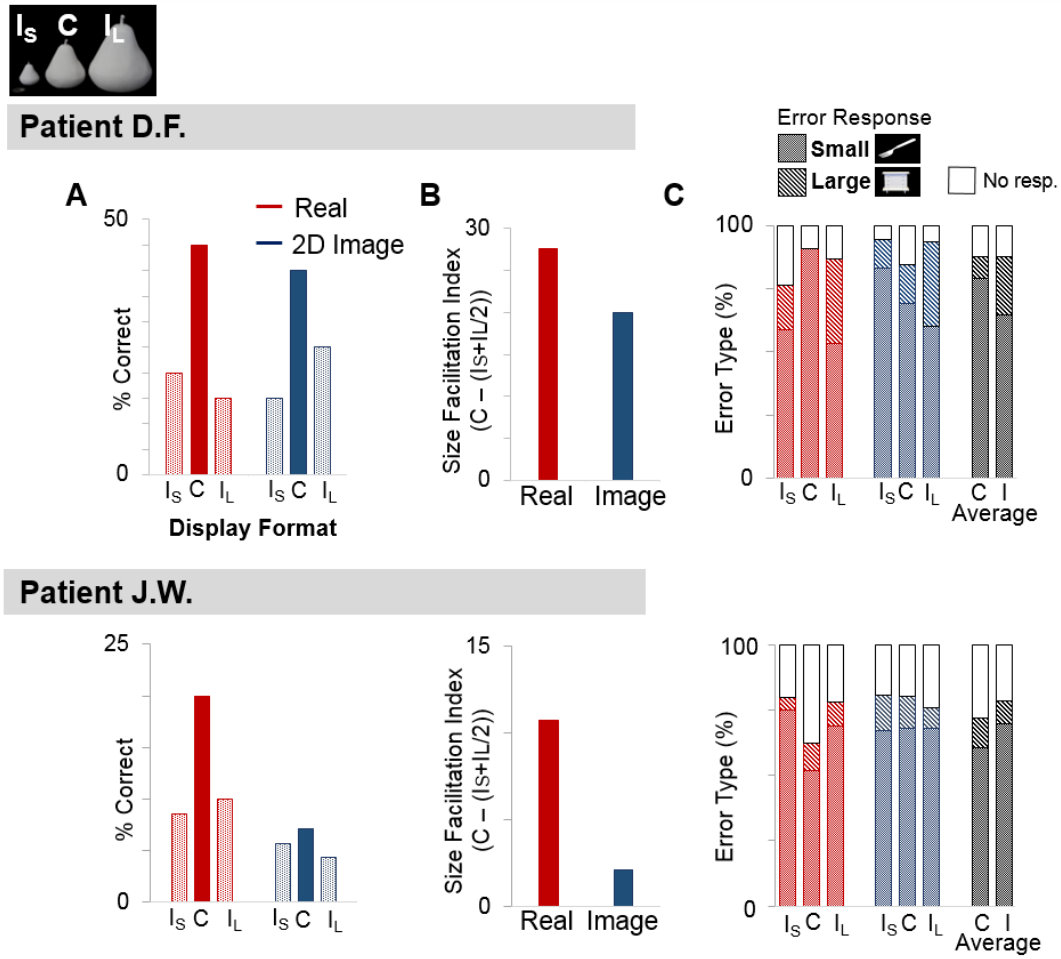

Figure 3: Recognition of Size-Congruent (C) vs. Size-Incongruent Small (IS) and Large (IL) Objects

(A) % correct responses by D.F. (*upper panel*) and J.W. (*lower panel*) for real objects and 2D images, separately for stimuli in each Size condition. All stimuli viewed binocularly with fixed head position.

(B) Facilitatory effect of size-congruence for real objects vs. 2D images.

(C) % trials where target was misidentified as a 'small' or 'large' object (or no response) separately for stimuli in each Display Format. Data averaged across Display Formats shown far right (black bars).

Naming Errors

Figure 3C illustrates the % of trials in which the patients made Small vs. Large Errors, separately for stimuli in each Format and Size condition. Because the stimuli in Experiment 2 were 'Small' objects, we did not expect the patients to produce high proportions of 'Large' Errors. Accordingly, D.F., showed a marginal difference in the probability of error responses for C versus I objects for real objects (D.F., C objects: p(Small Error)=1.0, p(Large Error)=0.0; I objects: p(Small Error)=0.69, p(Large Error)=0.31), $\chi 2=3.956$, p=0.047), and her pattern of errors was similar for both the IS and IL objects. For 2D images, D.F. showed no difference in the proportion of errors across the Size conditions (C objects: p(Small Error)=.89, p(Large Error)=.11; I objects: p(Small Error)=.89, p(Large Error)=.11, $\chi 2=0.001$, p=0.973). J.W. showed no difference in the proportion of Small versus Large Errors across Size conditions for real objects, or for images (J.W., Real C objects: p(Small Error)=.83, p(Large Error)=.17; Real I objects: p(Small Error)=.91, p(Large Error)=.09, $\chi 2=1.795$, p=0.180; Image C objects: p(Small Error)=.85, p(Large Error)=.15; Image I objects: p(Small Error)=.87, p(Large Error)=.13, $\chi 2=0.078$, p=0.780).

Recognition of Basic Geometric Shapes

Experiment 3 examined whether size-related effects on recognition were apparent for Basic Geometric Shapes, such as cylinders or spheres that do

not have a RWS association (**Figure 1C**).

Naming Accuracy

The % correct naming responses for D.F. and J.W. in the Basic Geometric Shapes task are illustrated in **Figure 4**. Recognition performance did not differ between the real object and image displays for D.F. (Real 66.67%, Image 66.67%, Cochran Q=0.00, p=1.00) or for J.W. (Real 66.67%, Image 74.40%, Cochran Q (5) =3.462, p=0.629). Unlike Experiment 2, there were no effects of stimulus size on recognition in either patient (D.F., all p-values> 0.167, J.W., all p- values> 0.317, Wilcoxon).

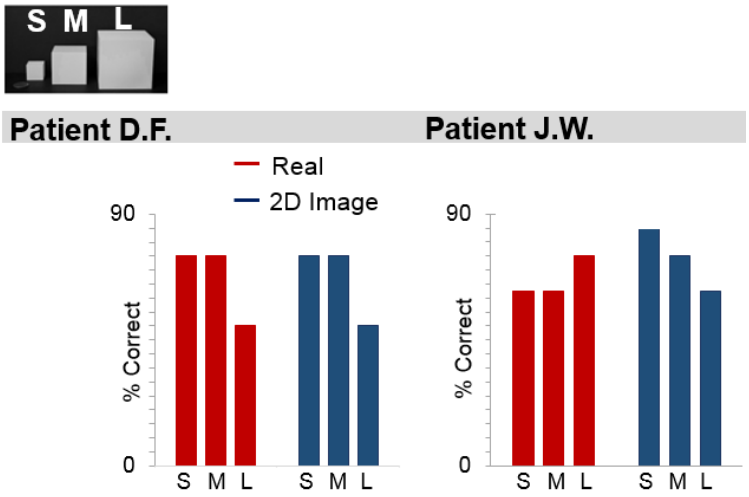

Figure 4: Recognition of Basic Geometric Shapes
% correct responses made by D.F. (*left panel*) and J.W. (*right panel*) in naming basic shapes displayed as real objects or 2D images, separately for small (S), medium (M), and large (L) sizes.

DISCUSSION

A key framework for understanding visual perception has been the division of visual cortical functions into two distinct processing pathways, a ventral stream that supports object perception, and a dorsal stream that supports spatial vision and visually-guided action [37, 38].

Unsurprisingly, damage to ventral stream areas (even just a small unilateral infarct [39]) along lateral occipital and ventral temporal cortex that are sensitive to object shape impairs object recognition –a condition known as visual form agnosia [14], and results in bilateral reductions in neural responses to images of objects [39]. However, the notion that perceptual processes are subserved solely by the ventral stream has come under increasing scrutiny following the discovery of object-selective representations in PPC, within the dorsal pathway in both primates [13, 20-24, 40, 41] and humans [1, 9, 25-31]. Isolating the functional contribution of dorsal object areas to perception has been hampered, however, by possible 'contamination' of signals from ventral cortex with which it is tightly interconnected.

Here, we isolated the functional significance of neural object representations in dorsal cortex by studying shape recognition in two patients with severe visual form agnosia following well- documented bilateral lesions to object processing

areas of either the lateral occipital area (D.F.) [15], or ventral occipital cortex (J.W.) [18], and concomitant severe deficits in image recognition. Critically, dorsal cortex in both patients is largely intact. The within-subject experimental manipulations permitted the patients to serve as their own controls. Although the patients showed severe deficits in recognizing 2D and 3D images of objects, they showed remarkably better recognition of real-world exemplars of the same items, particularly when depth cues were available (i.e., a real object advantage [33-36]). Importantly, the real object advantage depended crucially on the physical size of the stimulus: real objects viewed in depth were recognized most accurately when their physical size was consistent with their RWS. Recognition of size- incongruent real objects was severely impaired, whether the target objects were smaller or larger than their RWS. Conversely, size had little, if any, effect on recognition of 2D or 3D stereo images, or basic geometric shapes, even though the retinal extent of the images was configured to match their real-world counterparts. This pattern of performance was replicated clearly across both patients. Sensitivity to the identity of real-world sized objects was also apparent in the patients' error responses, particularly in Experiment 1 in which many objects were tested, some of whose physical size provided little information about RWS.

In conclusion, these neuropsychological data provide the first causal evidence for the role of human dorsal cortex in size-coding of real-world objects. Although healthy observers can show performance decrements when the

expected size of object images deviate dramatically from other objects in a scene [42, 43], here we show that in patients with visual agnosia (for whom contributions from ventral-stream object-processing areas are disrupted and recognition of object images is severely impaired), there is a remarkable sparing of recognition of real-world objects – but that this ability breaks down when the object's physical size deviates, even minimally, from typical real-world size. While a medial-to-lateral organization of expected 'real-world size' of 2D images is apparent in object coding in occipito-temporal cortex [44], concomitant documentation of similar computations in dorsal cortex has been surprisingly absent in these investigations.

However, topographic representations of object size have recently been identified in bilateral human parietal cortex using ultra-high-field (7T) functional MRI [45]. Furthermore, levels of GABA (the primary inhibitory neurotransmitter in the human brain) in parietal cortex can selectively influence size perception [46]. Together, the converging neuropsychological findings presented here raise the intriguing possibility that the unique effects that real-world objects have on behavior and brain responses [47-50] reflect a fundamental involvement of dorsal cortex in naturalistic vision, and underscore the importance of using richer stimuli to understand perception and neural coding in the human brain.

FUNDING

This research was supported by grants to J.C. Snow from the NEI of the NIH (R01EY026701). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

1. Konen, C.S., and Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. Nat. Neurosci. *11*, 224-231.
2. Freud, E., Plaut, D.C., and Behrmann, M. (2016). 'What' Is Happening in the Dorsal Visual Pathway. Trends Cogn. Sci. *20*, 773-784.
3. Kourtzi, Z., Erb, M., Grodd, W., and Bülthoff, H.H. (2003). Representation of the Perceived 3-D Object Shape in the Human Lateral Occipital Complex. Cereb. Cortex *13*, 911-920.
4. Kourtzi, Z., and Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. Science *293*, 1506-1509.
5. Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron *24*, 187-203.
6. Malach, R., Reppas, J., Benson, R., Kwong, K., Jiang, H., Kennedy, W., Ledden, P., Brady, T., Rosen, B., and Tootell, R. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. *92*, 8135- 8139.
7. Georgieva, S., Peeters, R., Kolster, H., Todd, J.T., and Orban, G.A.

(2009). The Processing of Three-Dimensional Shape from Disparity in the Human Brain. J. Neurosci. *29*, 727-742.

8. Freud, E., Ganel, T., Shelef, I., Hammer, M.D., Avidan, G., and Behrmann, M. (2015). Three-Dimensional Representations of Objects in Dorsal Cortex are Dissociable from Those in Ventral Cortex. Cereb. Cortex *27*, 422-434.
9. Van Dromme, I.C., Premereur, E., Verhoef, B.E., Vanduffel, W., and Janssen, P. (2016). Posterior Parietal Cortex Drives Inferotemporal Activations During Three-Dimensional Object Vision. PLoS Biol. *14*, e1002445.
10. Jeannerod, M. (1986). The formation of finger grip during prehension. A cortically mediated visuomotor pattern. Behav. Brain Res. *19*, 99-116.
11. Jeannerod, M., Decety, J., and Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. Neuropsychologia *32*, 369-380.
12. Gallivan, J.P., and Culham, J.C. (2015). Neural coding within human brain areas involved in actions. Curr. Opin. Neurobiol. *33*, 141-149.
13. Sereno, A.B., and Maunsell, J.H. (1998). Shape selectivity in primate lateral intraparietal cortex. Nature *395*, 500-503.
14. Farah, M.J. (2004). Visual agnosia, (MIT press).
15. Bridge, H., Thomas, O.M., Minini, L., Cavina-Pratesi, C., Milner, A.D., and Parker, A.J. (2013). Structural and functional changes across the

visual cortex of a patient with visual form agnosia. J. Neurosci. *33*, 12779-12791.

16. James, T.W., Culham, J., Humphrey, G.K., Milner, A.D., and Goodale, M.A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. Brain *126*, 2463-2475.

17. Grill-Spector, K. (2003). The neural basis of object perception. Curr. Opin. Neurobiol. *13*, 159-166.

18. Mapelli, D., and Behrmann, M. (1997). The role of color in object recognition: Evidence from visual agnosia. Neurocase *3*, 237-247.

19. Rosenthal, O., and Behrmann, M. (2006). Acquiring long-term representations of visual classes following extensive extrastriate damage. Neuropsychologia *44*, 799-815.

20. Murata, A., Gallese, V., Luppino, G., Kaseda, M., and Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. J. Neurophysiol. *83*, 2580-2601.

21. Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y., and Tsutsui, K. (1998). Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. Philos. Trans. R Soc. Lond. B Biol. Sci. *353*, 1363-1373.

22. Sereno, M.E., Trinath, T., Augath, M., and Logothetis, N.K. (2002). Three-dimensional shape representation in monkey cortex. Neuron *33*, 635-652.

23. Lehky, S.R., and Sereno, A.B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. J. Neurophysiol. *97*, 307-319.

24. Sakata, H., Tsutsui, K.-I., and Taira, M. (2005). Toward an understanding of the neural processing for 3D shape perception. Neuropsychologia *43*, 151-161.

25. Jeong, S.K. (2016). Behaviorally Relevant Abstract Object Identity Representation in the Human Parietal Cortex. J. Neurosci. *36*, 1607-1619.

26. Zachariou, V., Klatzky, R., and Behrmann, M. (2014). Ventral and dorsal visual stream contributions to the perception of object shape and object location. J. Cogn. Neurosci. *26*, 189-209.

27. Zachariou, V., Nikas, C.V., Safiullah, Z.N., Behrmann, M., Klatzky, R., and Ungerleider,

L.G. (2015). Common Dorsal Stream Substrates for the Mapping of Surface Texture to Object Parts and Visual Spatial Processing. J. Cogn. Neurosci. *27*, 2442-2461.

28. Bracci, S., and Op de Beeck, H. (2016). Dissociations and Associations between Shape and Category Representations in the Two Visual Pathways. J. Neurosci. *36*, 432-444.

29. Freud, E., Culham, J.C., Plaut, D.C., and Behrmann, M. (2017). The large-scale organization of shape processing in the ventral and dorsal pathways. eLife *6*, e27576.

30. Xu, Y. (2009). Distinctive Neural Mechanisms Supporting Visual Object Individuation and Identification. J. Cogn. Neurosci. *21*, 511-518.

31. Freedman, D.J., and Assad, J.A. (2009). Distinct Encoding of Spatial and Nonspatial Visual Information in Parietal Cortex. J. Neurosci. *29*, 5671-5680.

32. Bettencourt, K.C., and Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. Nat. Neurosci. *19*, 150.

33. Farah, M.J. (1990). Visual agnosia, (MIT press).

34. Ratcliff, G., and Newcombe, F. (1982). Object recognition: Some deductions from the clinical evidence. Normality and Pathology in Cognitive Functions, 147-171.

35. Humphrey, G.K., Goodale, M.A., Jakobson, L.S., and Servos, P. (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. Perception *23*, 1457-1481.

36. Chainay, H., and Humphreys, G.W. (2001). The real-object advantage in agnosia: Evidence for a role of surface and depth information in object recognition. Cogn. Neuropsychol. *18*, 175-191.

37. Ungerleider, L., Mishkin, M. (1982). Two cortical visual systems. Analysis of visual behavior (Ingle, D.J. et al., eds), 549-586.

38. Goodale, M.A., and Milner, A.D. (1992). Separate visual pathways for perception and action. Trends Neurosci. *15*, 20-25.

39. Konen, C.S., Behrmann, M., Nishimura, M., and Kastner, S. (2011). The functional neuroanatomy of object agnosia: a case study. Neuron *71*, 49-60.

40. Srivastava, S., Orban, G.A., De Mazière, P.A., and Janssen, P. (2009). A Distinct Representation of Three-Dimensional Shape in Macaque Anterior Intraparietal Area: Fast, Metric, and Coarse. J. Neurosci. *29*, 10613-10626.

41. Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W., and Orban, G.A. (2005). Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. J. Neurosci. *25*, 4294-4306.

42. Eckstein, M.P., Koehler, K., Welbourne, L.E., and Akbas, E. (2017). Humans, but Not Deep Neural Networks, Often Miss Giant Targets in Scenes. Curr. Biol. *27*, 2827- 2832.e2823.

43. Konkle, T., and Oliva, A. (2012). A familiar-size Stroop effect: real-world size is an automatic property of object representation. J. Exp. Psychol. Hum. Percept. Perform. *38*, 561-569.

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

45. Harvey, B.M., Fracasso, A., Petridou, N., and Dumoulin, S.O.

(2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. Proc. Natl. Acad. Sci. *112*, 13525-13530.

46. Song, C., Sandberg, K., Andersen, L.M., Blicher, J.U., and Rees, G. (2017). Human Occipital and Parietal GABA Selectively Influence Visual Perception of Orientation and Size. J. Neurosci. *37*, 8929-8937.

47. Snow, J.C., Pettypiece, C.E., McAdam, T.D., McLean, A.D., Stroman, P.W., Goodale, M.A., and Culham, J.C. (2011). Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. Sci. Rep. *1*, 130.

48. Snow, J.C., Skiba, R.M., Coleman, T.L., and Berryhill, M.E. (2014). Real-world objects are more memorable than photographs of objects. Front. Hum. Neurosci. *8*, 837.

49. Squires, S.D., Macdonald, S.N., Culham, J.C., and Snow, J.C. (2015). Priming tool actions: Are real objects more effective primes than pictures? Exp. Brain Res.

50. Gomez, M., Skiba, R., and Snow, J. (in press). Graspable objects grab attention more than images. Psychol. Sci.

51. Milner, A., Perrett, D., Johnston, R., Benson, P., Jordan, T., Heeley, D., Bettucci, D., Mortara, F., Mutani, R., and Terazzi, E. (1991). Perception and action in 'visual form agnosia'. Brain *114*, 405-428.

52. Cant, J.S., and Goodale, M.A. (2007). Attention to form or surface

properties modulates different regions of human occipitotemporal cortex. Cereb. Cortex *17*, 713-731.

53. Goodale, M.A., Milner, A.D., Jakobson, L.S., and Carey, D.P. (1991). A neurological dissociation between perceiving objects and grasping them. Nature *349*, 154-156.

54. Marotta, J.J., Behrmann, M., and Goodale, M.A. (1997). The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. Exp. Brain Res. *116*, 113-121.