



# HHS Public Access

## Author manuscript

*Cortex*. Author manuscript; available in PMC 2020 October 01.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Published in final edited form as:

*Cortex*. 2019 October ; 119: 555–568. doi:10.1016/j.cortex.2019.02.030.

## Real-world size coding of solid objects, but not 2-D or 3-D images, in visual agnosia patients with bilateral ventral lesions

**Desiree E. Holler<sup>1</sup>, Marlene Behrmann<sup>2</sup>, Jacqueline C. Snow<sup>1,\*</sup>**

<sup>1</sup>Department of Psychology, University of Nevada, Reno, USA

<sup>2</sup>Department of Psychology, Carnegie Mellon University, Pittsburgh, USA

### Abstract

Patients with visual agnosia show severe deficits in recognizing two-dimensional (2-D) images of objects, despite the fact that early visual processes such as figure-ground segmentation, and stereopsis, are largely intact. Strikingly, however, these patients can nevertheless show a preservation in their ability to recognize real-world objects –a phenomenon known as the ‘real-object advantage’ (ROA) in agnosia. To uncover the mechanisms that support the ROA, patients were asked to identify objects whose size was congruent or incongruent with typical real-world size, presented in different display formats (real objects, 2-D and 3-D images). While recognition of images was extremely poor, real object recognition was surprisingly preserved, but only when physical size matched real-world size. Analogous display format and size manipulations did not influence the recognition of common geometric shapes that lacked real-world size associations. These neuropsychological data provide evidence for a surprising preservation of size-coding of real-world-sized tangible objects in patients for whom ventral contributions to image processing are severely disrupted. We propose that object-size information is largely mediated by dorsal visual cortex and that this information, together with detailed representation of object shape which is also subserved by dorsal cortex, serve as the basis of the ROA.

### Keywords

Visual form agnosia; object recognition; real-world size; real-world objects; two-dimensional images

---

**Correspondence:** Jacqueline Snow: Jacqueline.c.snow@gmail.com, Department of Psychology, University of Nevada, Reno, NV 89557.

CRedit author statement

**Desiree E. Holler:** Software, Validation, Formal Analysis, Investigation, Provision, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization. **Marlene Behrmann:** Provision, Writing – Original Draft, Writing – Review & Editing. **Jacqueline C. Snow:** Software, Validation, Formal Analysis, Investigation, Provision, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Supervision, Project Administration, Funding Acquisition.

**AUTHOR CONTRIBUTIONS**

J.C.S., M.B. and D.E.H. contributed to the concept and design of the study. D.E.H. and J.C.S. were responsible for patient data collection and analyses. J.C.S., M.B. and D.E.H. wrote the manuscript.

\*Lead Contact

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**DECLARATIONS**

Declarations of interest: none

## 1. INTRODUCTION

Patients with visual agnosia have severe impairments in object recognition, despite having intact early visual function (Farah, 1990). Visual agnosia is typically associated with bilateral damage to lateral and ventral occipitotemporal cortex, although even small unilateral right hemisphere infarcts confined to the lateral occipital mantle are sufficient to disrupt recognition and produce bilateral reductions in neural responses to images of objects (Bridge et al., 2013, Konen, Behrmann, Nishimura, & Kastner, 2011). The deficits shown by agnosia patients in object recognition are paralleled by findings from neuroimaging studies conducted with neurologically healthy observers, which show that ventral-stream areas along lateral occipital and ventral temporal cortex are sensitive to object shape (Grill-Spector et al., 1999, Kourtzi & Kanwisher, 2000, Malach et al., 1995) and respond largely invariantly to changes in retinal size, viewpoint, and pictorial format of object images (Grill-Spector et al., 1999, Kourtzi & Kanwisher, 2000, Vuilleumier, Henson, Driver, & Dolan, 2002).

Although patients with visual agnosia show severe deficits in recognizing two-dimensional (2-D) line drawings or photographs of objects, they can nevertheless show striking improvements in recognition when the items are presented as real-world solid objects (Chainay & Humphreys, 2001, Farah, 1990, Hiraoka, Suzuki, Hirayama, & Mori, 2009, Humphrey, Goodale, Jakobson, & Servos, 1994, Ratcliff & Newcombe, 1982, Riddoch & Humphreys, 1987, Servos, Goodale, & Humphrey, 1993, Turnbull, Driver, & McCarthy, 2004). This phenomenon, which has been termed the 'Real Object Advantage' (ROA) in agnosia (Chainay & Humphreys, 2001), has been documented in numerous case reports over the past 40 years, yet the underlying mechanism has received surprisingly little investigation.

The fact that agnosia patients show a ROA suggests that they may be able to use visual cues other than lines and edges to bootstrap recognition. There are a number of possible factors that might favor recognition of real objects over 2D images. For example, real objects may provide richer information about surface color and texture than is true for 2D images (Farah, 1990, Riddoch & Humphreys, 1987), and agnosia patients might make use of these cues to facilitate identification (Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010, Humphrey et al., 1994, James, Culham, Humphrey, Milner, & Goodale, 2003). However, agnosia patients with achromatopsia (who are unable to perceive color) also show a ROA (Chainay & Humphreys, 2001, Riddoch & Humphreys, 1987), as do those who are asked to identify objects under monochromatic viewing conditions (Humphrey et al., 1994). Another possibility is that recognition of real objects is favored due to the presence of depth information in the stimulus. When real objects are viewed with laterally-displaced eyes (as in humans) the resulting retinal disparities in the left- and right-eye images are combined to yield stereoscopic information about the 3-D depth structure of the object (Julesz, 2006). Stereoscopic cues become less informative with increasing viewing distance (Banks, Read, Allison, & Watt, 2012). Information about the 3-D shape of real objects (but not 2-D images) can also be obtained by moving the head to inspect the item from different viewing angles to reveal occluded surfaces, and patients with agnosia often show large head movements during the recognition process (Humphrey et al., 1994). Lateral translations of the head also elicit differences in the relative motion on the retinal image of objects and their

parts that lie at different distances. These relative motion signals, known as motion parallax, also provide information about depth and 3-D shape (Kim, Angelaki, & DeAngelis, 2016).

The only two studies to date that have investigated the causal basis for the ROA have examined the contribution of depth information. Humphrey et al. (1994) tested whether depth information from disparity modulates the ROA in a patient (D.F.) with severe visual agnosia resulting from bilateral lateral-occipital lesions. D.F. was asked to identify objects that were viewed monocularly versus binocularly (as well as in full color versus monochromatic), with head position fixed in a chin rest. While the elimination of surface color resulted in a pronounced decline in recognition performance, accuracy also declined when monochromatic real objects were viewed monocularly vs. binocularly (although performance in all real object conditions was still better than for black and white photographs). Later, Chainay and Humphreys (2001) examined the relative importance of depth information from stereoscopic disparity and motion parallax on object recognition in another patient (H.J.A.) with severe visual agnosia and achromatopsia resulting from bilateral occipito-temporal lesions that extended medially into the fusiform and lingual gyri. H.J.A. was asked to identify objects positioned at 'close' or 'far' egocentric positions (i.e., stereoscopic disparity), with head position either free to move vs. fixed in a chin rest (i.e., motion parallax). The ROA was evident when either stereoscopic disparity or motion parallax cues were available, but the patient's performance dropped sharply to the level of images when both of these depth cues were removed (Chainay & Humphreys, 2001). The authors concluded that agnosia patients show an ROA because real objects are typically viewed under conditions in which depth information from binocular disparity and motion parallax can be used to help differentiate the various surfaces and parts of 3-D objects (Chainay & Humphreys, 2001).

Importantly, however, these results leave open the question of whether agnosia patients use depth information to facilitate 3-D shape processing *per se*, or whether they are sensitive to object properties *other than* shape that are also conveyed by depth cues. For example, knowing the position of a solid object in depth provides information about its real-world physical size. Because planar images of objects do not convey real-world size information (since only the distance to the monitor or picture is known by the observer, but not the distance to the depicted object), this could favor recognition of real objects. In patients with agnosia, for whom normal edge-based coding mechanisms in the ventral pathway are impaired, physical size may serve as a more powerful clue to identity than 3-D shape cues. Previous patient studies (Chainay & Humphreys, 2001, Humphrey et al., 1994) cannot disentangle the relative contribution of 3-D shape versus size to the ROA because the stimuli were all objects whose physical size matched the typical real-world-size. Alternatively, depth cues may be used to code the size and shape of real objects according to the affordances they offer for manual interaction (Chainay & Humphreys, 2001, Gibson, 1979, Humphreys, 2013). Distance, orientation, and size, are important cues that guide actions with objects, and the dorsal visual stream is dedicated to processing spatial position and the planning and execution of goal-directed actions (Goodale & Milner, 1992, Kastner & Ungerleider, 2001). Neurophysiological studies in non-human primates have revealed neural populations in dorsal cortex that code for surface depth and orientation, as well as object size (Murata, Gallese, Kaseda, & Sakata, 1996, Sakata, Taira, Murata, & Mine, 1995, Taira,

Mine, Georgopoulos, Murata, & Sakata, 1990). The ROA in agnosia patients could therefore reflect a contribution to recognition from action-related visual processes that are largely unaffected by the lesion.

Here, across a series of experiments, we studied the causal contribution of real-world size and display format on shape recognition in two patients with profound visual agnosia following bilateral ventral stream lesions (Figure 1). We measured the patients' object recognition in different display formats: real tangible objects and high-resolution 2-D and/or 3-D stereoscopic computerized images of the same items. The images were matched closely to their real-world counterparts for retinal size, orientation, background and illumination, and all stimuli were achromatic to prevent the use of color to bootstrap object processing (James et al., 2003, Mapelli & Behrmann, 1997). Critically, we manipulated the size of the stimuli in each display format relative to the typical real-world size. We predicted that if agnosia patients are sensitive to real-world size cues in the stimulus, then they should show a recognition advantage for real objects versus 2D images (when depth cues are available) when the physical size of the object matches its expected real-world size. Recognition of real objects should be impaired, however, when physical size is inconsistent with real-world size. Manipulations of size should have little, if any, influence on recognition of 2-D images of objects because these stimuli do not convey real-world size information. Size manipulations should also have no influence on the recognition of geometric shapes that lack strong real-world size associations. While stereoscopic images are visually similar to solid objects because they convey information about the 3-D structure, apparent distance and size of an object, they are not solids and they do not afford manual interaction.

## 2. MATERIALS AND METHODS

### 2.1 Participants

Patients D.F. and J.W. provided informed consent to participate in the experiments. All procedures were carried out in accordance with the guidelines approved by the University of Nevada, Reno Institutional Review Board.

**Patient D.F.:** D.F. is a right-handed native English speaking female who was age 61 at time of testing. D.F. suffered toxic exposure to carbon monoxide at the age of 34, leading to a profound visual agnosia characterized by a severe impairment in object recognition that cannot be reduced to a simple visual sensory deficit (Milner et al., 1991). D.F.'s lesion (Figure 1, *upper panel*) involves predominantly bilateral lateral occipital (LO) cortex (Bridge et al., 2013, James et al., 2003). An anatomo-functional examination of D.F.'s brain by Bridge et al. (2013) 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. D.F.'s stereoscopic vision is strikingly unimpaired, and she is able to make conscious judgments of absolute depth (Milner et al., 1991, Read, Phillipson, Serrano-Pedraza, Milner, & Parker, 2010). D.F.'s recognition is poorest for silhouettes and black and white line-drawings, but improves with the addition of color and surface texture cues (Humphrey et al., 1994, Milner et al., 1991), as mentioned above. 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 (Bridge et al., 2013, James et al., 2003).

Importantly, D.F. is better at recognizing real-world tangible objects than photographs of objects – a ROA (Humphrey et al., 1994, Milner et al., 1991).

**Patient J.W.:** J.W. is a right-handed native English speaking male who was 55–56 years of age over the course of testing. J.W. suffers from a profound visual 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. J.W. has bilateral lesions to the ventral occipital lobes, areas that provide critical inputs to LOC (Mapelli & Behrman, 1997, Rosenthal & Behrman, 2006). 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 (Figure 2, *lower panel*). We were unable to obtain more recent imaging data from JW, due to the presence of an automated implantable cardioverter defibrillator (AICD). J.W.’s recognition is poor for silhouettes and black and white line-drawings, but his performance improves with the addition of color and surface texture cues (Mapelli & Behrman, 1997, Marotta, Behrman, & Goodale, 1997). J.W. has intact stereo vision and coarse shape recognition (circle versus square) (Rosenthal & Behrman, 2006).

## 2.2 General Methods

We report all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

**2.2.1 Stimuli and Apparatus**—The stimuli in all experiments were real-world objects and matched high-resolution 2-D and/or 3-D stereoscopic computerized images of the same objects. 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. To create the 2-D images, the real objects were photographed using a Canon Rebel T2i DSLR camera with constant F-stop and shutter speed. The camera 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 2-D images were displayed on a 27" ACER (G276HL) LCD monitor (60-Hz refresh rate) with a screen resolution of 1920 × 1080 pixels, controlled by an Intel Core I5 1.5–3210M GHz laptop computer (8 GB RAM). We created 3-D 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 stereo images were viewed binocularly through active shutter glasses (3-D Vision 2, NVIDIA, USA) that present the two offset images separately to the left and right eye. The 3-D images were displayed on an ASUS (VG278HE) LCD monitor (120 Hz) with a screen resolution of 1920 × 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. 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. The MATLAB code we developed for initiating and recording the audio patient responses on each trial and information about the stimuli used in each experiment are available from <http://hdl.handle.net/11714/4902>.

Both patients were tested with three different stimulus sets:

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

#### **2.2.1.2 Size-Congruent (C) vs. Size-Incongruent Small ( $I_S$ ) and Large ( $I_L$ ) Objects:**

For D.F., the stimuli were 21 everyday objects, each of which was 3-D-printed in three different sizes, yielding a total of 63 stimuli (Figure 2B). For J.W., who was able to complete more trials than D.F. within the available testing time, the stimuli were 70 different exemplars printed in three sizes, for a total of 210 stimuli. The real objects were 3-D-printed in white ABS plastic using either a Stratasys uPrint SE Plus 3-D-printer Afina H800+, or a DeltaMaker 2T. An Arctec Spider 3-D Scanner was used to create digital 3-D renderings of each object. For C (100%) stimuli, the digital renderings were 3-D printed so that volume matched the real-world exemplar. Size-incongruent (I) stimuli were 3-D printed so that volume was 50% of real-world size ( $I_S$ ), or 150% of real-world size ( $I_L$ ). Stimulus volume was measured using CURA software.

**2.2.1.3 Geometric Shapes:** The stimuli were four (D.F.) or 8 (J.W.) common geometric shapes that were 3-D-printed in three different sizes, for a total of 12 (D.F.) or 24 (J.W.) stimuli (Figure 2C).

**2.2.2 General Procedure**—On 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 on the monitor. The ITI was 10 sec in image blocks, and  $\sim 10$  sec in real object blocks, with white-noise played throughout. Viewing distance was  $\sim 57$  cm in all conditions. Patients were instructed to name each object as quickly and accurately as possible, and advised that responses would be timed and recorded. No feedback was provided. Order of trials was randomized within each block.

**2.2.2.1 Size-Congruent (C) vs. Size-Incongruent Small ( $I_S$ ) Objects:** Stimuli were presented to the patients under 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 stereo disparity and motion parallax depth cues. In the '*binocular fixed*' condition, objects were presented with binocular vision but the head was positioned in a chin rest, thereby providing stereo disparity but not motion parallax 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 *2-D image, and 3-D image trials* (J.W.), the patients viewed each exemplar binocularly with head fixed in a chin rest.

For D.F. the 140 exemplars were divided into five sets of 28 items, half C, half  $I_S$ . 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 session (28 items per set x 5 viewing conditions), for a total of 700 trials. The sessions were conducted over two consecutive days (~12 hours total testing time). 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 to permit a strong test of the presence of a ROA. For J.W., the 140 exemplars were divided into two sets of 70 items, half C and half  $I_S$ . The experiment was divided into two testing sessions. The first session was comprised of four blocks of trials, in which the two sets of 70 items were presented as real objects and as 3-D stereoscopic images, using an ABBA design. The second testing session was comprised of two blocks of trials in which the two sets of 70 items stimuli were presented as 2-D images. There were 70 trials per block, for a total of 420 trials. Testing sessions were conducted over two consecutive days (~8 hours testing). In summary, in Experiment 1, D.F. saw the stimuli *five times* as real objects before viewing the same exemplars as 2-D images, whereas J.W. saw the stimuli once as real objects, and once as 3-D images, before the 2-D image condition.

### **2.2.2.2 Size-Congruent (C) vs. Size-Incongruent Small ( $I_S$ ) and Large ( $I_L$ )**

**Objects:** For D.F., the 63 real objects (and 63 matching 2-D 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 sets of exemplars first as real objects, across three consecutive blocks. She then viewed the same exemplars as 2-D images in three consecutive blocks, for a total of 126 trials. D.F. completed the experiment in ~4 hrs. 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  $I_S$ , and the remainder were  $I_L$ . 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, followed by 2-D images. In each session, the sets were viewed in consecutive blocks of trials, for a total of 210 trials per session. J.W. took ~5 hours to complete each session. All stimuli were presented binocularly with fixed-head position.

**2.2.2.3 Geometric Shapes:** For patient D.F., the 12 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 2-D image trials. The experiment took D.F. ~1 hr to complete, and J.W. ~ 2 hrs to complete. All stimuli were presented binocularly with fixed-head position. Both patients were informed that the stimuli were common geometric shapes.

**2.2.3 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. Digitized recordings of naming responses on each trial were analyzed 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. For the analysis of errors we examined the types of responses that were 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., fork, pen, toothbrush). ‘Large Errors’ were defined as responses in which the target was misidentified as another object that is typically much larger (i.e., dresser, airplane, house). If the patients are more sensitive to the identity of objects whose physical size reflects real-world size, 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 (Chainay & Humphreys, 2001). All statistical tests were two-tailed, except for the Fisher’s Exact tests in analysis of errors, which were one-tailed.

The terms of our ethics approval do not permit public archiving data from single-case designs used here, as data publically archived has to be completely anonymized, and due to the study design (which differed for both cases in the study) the patient data cannot be fully anonymized. Readers seeking access the raw data should contact the lead author or the University of Nevada, Reno Institutional Review Board. Access will be granted only to named individuals in accordance with ethical procedures governing the reuse of sensitive data. No part of the study procedures or analyses was pre-registered in a time-stamped, institutional registry prior to the research being conducted.

### 3. RESULTS

#### 3.1 Recognition of Size-Congruent (C) vs. Size-Incongruent Small (I<sub>S</sub>) Objects

In Experiment 1, we evaluated in different Display Formats (real objects, 2-D or 3-D images) the patients’ recognition of objects whose Size was either consistent with, or orders of magnitude smaller than, typical real-world size (Figure 2A).

**3.1.1 Naming Accuracy**—Both patients performed below ceiling in all conditions, consistent with a generalized impairment in shape recognition but evinced a clear ROA, reflected by superior recognition of real objects versus 2-D images when viewed binocularly

with fixed head position (D.F.,  $p=0.002$ , J.W.,  $p=0.018$ , Wilcoxon) (Figure 3A–B). For J.W., recognition of real objects was also better than 3-D images of the same items ( $p=0.003$ , Wilcoxon), which was equal to 2-D 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  $I_S$  objects (D.F., C=25.1%,  $I_S$ =5.2%,  $\chi^2=53.14$ ,  $p<0.001$ ; J.W., C=21.9%,  $I_S$ =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  $I_S$  objects, for which performance remained severely impaired 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  $I_S$ ) stimuli when depth cues were present (D.F., real objects binocular free, binocular fixed, monocular free, all  $p$ -values  $\leq 0.001$ ,  $\chi^2\geq11.706$ ; J.W., real objects binocular fixed  $p=0.002$ ,  $\chi^2=9.786$ , 3-D image  $p=.014$ ,  $\chi^2=6.048$ ). When depth cues were absent, there was no effect of Size for 2-D images in J.W. ( $p=0.217$ ,  $\chi^2=1.522$ ). Although D.F. showed a small advantage for C vs.  $I_S$  objects when stereo and parallax depth cues were absent (real objects monocular head fixed,  $p=0.026$ ,  $\chi^2=4.965$ ; 2-D images,  $p=0.026$ ,  $\chi^2=4.965$ ), her recognition of size-congruent stimuli under these viewing conditions was nevertheless significantly *worse* than when the same objects were viewed with binocular depth cues available (binocular free vs. monocular fixed:  $p=0.002$ ; binocular free vs. image:  $p=0.002$ ; binocular fixed vs. monocular fixed:  $p=0.002$ ; binocular fixed vs. image:  $p=0.003$ , Wilcoxon), and intermediate recognition of the objects under monocular head fixed conditions (monocular free vs. monocular fixed:  $p=0.102$ ; monocular free vs. image:  $p=0.131$ , Wilcoxon).

Next, to quantify the effect of size-congruence on recognition accuracy, we calculated a size facilitation index for each patient by subtracting % correct on  $I_S$  trials from C trials, separately for each viewing condition (Figure 3C). 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. 2-D 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 2-D images ( $t(1) = 3.728$ ,  $p=0.167$ ,  $d=7.456$ ), and that the index for real objects was significantly greater than 2-D images ( $t(1)=13.518$ ,  $p=0.047$ ,  $d=4.31$ ).

**3.1.2 Naming Errors**—Next, we examined the types of errors that the patients produced when objects were misidentified (Figure 3D). For example, a *clothespeg* 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  $I_S$  stimuli for which error type should be of equal likelihood (i.e., indicative of 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  $I_S$  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<sub>S</sub> stimuli were apparent in each viewing condition (all p-values <0.001), suggesting that patients are sensitive to size cues for recognition and are constrained in the error responses accordingly.

### 3.2 Recognition of Size-Congruent (C) vs. Size-Incongruent Small (I<sub>S</sub>) and Large (I<sub>L</sub>) Objects

For Experiment 2, the same objects were presented in each condition, thereby ruling out explanations based on familiarity or feature compression (Figure 2B). To do this, we 3-D scanned and 3-D-printed each exemplar in *three* different sizes: 50% of real-world size (incongruent small; I<sub>S</sub>), equal to (100%) real-world size (congruent; C), or larger objects that were 150% of their real-world size (incongruent large; I<sub>L</sub>). Each exemplar therefore served as its own control for object-related effects, such as familiarity. If real-world size 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 size (and size effects should be reduced for images).

**3.2.1 Naming Accuracy**—As in Experiment 1, there were striking format-dependent effects of Size on recognition in both patients (Figure 4A). For patient D.F. (who viewed 21 different exemplars), although overall performance was comparable for both the real object and image displays (26.7% correct, Cochran's Q=0.00,  $p=1.00$ ), for real objects, recognition was significantly better for C versus both I<sub>S</sub> and I<sub>L</sub> stimuli (I<sub>S</sub>:  $p=0.025$ , I<sub>L</sub>:  $p=0.014$ , Wilcoxon), whereas for 2-D images, there was no difference in recognition for C versus I<sub>S</sub> and I<sub>L</sub> stimuli (I<sub>S</sub>:  $p=.059$ ; I<sub>L</sub>:  $p=0.257$ , Wilcoxon). For J.W., who viewed more than three times as many exemplars as D.F. ( $n=70$ ), the effects of display format and size were even more pronounced. Overall, J.W. was better at recognizing real objects (12.86% correct) compared to 2-D images (5.71% correct, Cochran's Q (1)=9.783,  $p=0.002$ ; a strong ROA). Similar to D.F., for real objects, J.W.'s recognition performance was significantly better for C versus both I<sub>S</sub> and I<sub>L</sub> stimuli (I<sub>S</sub>:  $p=0.021$ ; I<sub>L</sub>:  $p=0.008$ , Wilcoxon), whereas for 2-D images, recognition was equally poor for C versus I<sub>S</sub> and I<sub>L</sub> stimuli (I<sub>S</sub>:  $p=0.655$ ; I<sub>L</sub>:  $p=0.317$ , Wilcoxon).

Next, we calculated a size facilitation index by subtracting the patient's mean % correct on I<sub>S</sub> and I<sub>L</sub> trials from that of congruent trials (C – (I<sub>S</sub> + I<sub>L</sub> / 2)) (Figure 4B). Parametric contrasts of the average facilitation indices for real objects versus 2-D images indicated that the size facilitation effect for real objects was significantly greater than for 2-D 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.

**3.2.2 Naming Errors**—Because the stimuli in Experiment 2 were all '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:  $p(\text{Small Error})=1.0$ ,  $p(\text{Large Error})=0.0$ ; I objects:  $p(\text{Small Error})=0.69$ ,  $p(\text{Large Error})=0.31$ ,  $\chi^2=3.956$ ,  $p=0.047$ ), and her pattern of errors was similar for both the  $I_S$  and  $I_L$  objects (Figure 4C, *upper panel*). For 2-D images, D.F. showed no difference in the proportion of errors across the Size conditions (C objects:  $p(\text{Small Error})=.89$ ,  $p(\text{Large Error})=.11$ ; I objects:  $p(\text{Small Error})=.89$ ,  $p(\text{Large Error})=.11$ ,  $\chi^2=0.001$ ,  $p=0.973$ ). 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:  $p(\text{Small Error})=.83$ ,  $p(\text{Large Error})=.17$ ; Real I objects:  $p(\text{Small Error})=.91$ ,  $p(\text{Large Error})=.09$ ,  $\chi^2=1.795$ ,  $p=0.180$ ; Image C objects:  $p(\text{Small Error})=.85$ ,  $p(\text{Large Error})=.15$ ; Image I objects:  $p(\text{Small Error})=.87$ ,  $p(\text{Large Error})=.13$ ,  $\chi^2=0.078$ ,  $p=0.780$ ) (Figure 4C, *lower panel*).

### 3.3 Recognition of Geometric Shapes

Experiment 3 examined whether size effects on recognition were apparent for common geometric shapes, such as cubes or spheres, that lack strong real-world size associations (Figure 2C).

**3.3.1 Naming Accuracy**—There was no difference in recognition performance between the real object and image displays, for either 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,  $p=0.629$ ) (Figure 5). 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).

## 4. DISCUSSION

Here, we studied shape recognition in two patients with severe visual agnosia following well-documented bilateral lesions to ventral pathway shape-selective areas (D.F.) or their afferents (J.W.). We were particularly interested in whether the patients, each of whom has great difficulty recognizing black-and-white object images, would show a relative sparing in their ability to recognize monochromatic real-world objects (i.e., a ‘Real Object Advantage’ (ROA)) (Chainay & Humphreys, 2001, Farah, 1990, Hiraoka et al., 2009, Humphrey et al., 1994, Ratcliff & Newcombe, 1982, Turnbull et al., 2004), and whether the ROA would be modulated by object size. To address this question, we compared the patients’ ability to recognize objects that were presented in different display formats (real objects vs. 2-D or 3-D stereoscopic images of the same items), and whose size was congruent or incongruent with typical real-world size. The 2-D and 3-D images were matched closely to their real-world counterparts for retinal size, viewpoint, background and illumination. The within-subject experimental manipulations permitted the patients to serve as their own controls.

Consistent with the patients’ lesion anatomy, and with previous case studies of agnosia, both D.F. and J.W. were extremely poor at recognizing 2-D and 3-D monochromatic images of objects. Strikingly, however, both patients showed a relative sparing in their ability to recognize monochromatic real-world exemplars of the same items. This recognition advantage for real objects (versus images) was most apparent when depth cues were available (Chainay & Humphreys, 2001, Humphrey et al., 1994). Unlike previous studies, however, we found that the ROA depended critically upon the physical size of the object.

Real objects viewed in depth were recognized most accurately when their physical size matched the expected real world size. In contrast, recognition of size-incongruent real objects was severely impaired in all viewing conditions, whether or not depth cues were available. These size-dependent effects on shape recognition are unlikely to reflect changes in visual complexity associated with feature compression because size manipulations interfered with recognition whether the objects were smaller or larger than their typical real-world size, and whether physical size deviated from real-world size on a large or small scale. The effect of physical size on recognition of real objects is also unlikely to reflect differences in the patients' familiarity with small vs. large objects, because size modulated recognition even when object identity remained constant across parametric size manipulations. 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 real-world size.

Whereas the patients' recognition of real objects was severely affected by mismatches between physical versus real-world size, equivalent size manipulations had little, if any, effect on recognition of 2-D or 3-D stereoscopic images of objects. This result is especially surprising given that the retinal extent of the images was configured to match their real-world counterparts, and the patients viewed the same objects as real-world exemplars before viewing them as images. Similarly, size manipulations had no influence on the patients' ability to recognize geometric shapes that lack specific real world size associations. Together, these data reveal a selective sparing of perception for real-world-sized tangible objects in patients who otherwise suffer from profound visual agnosia.

Although both D.F. and J.W. showed a remarkably similar pattern of performance across experiments, the patients nevertheless showed subtle differences in the 2-D image conditions. First, whereas J.W.'s recognition of 2-D (and 3-D) images was unaffected by visual size in all experiments, D.F. evinced a small yet significant effect of size on recognition of 2-D images in Experiment 1. Importantly, however, D.F.'s performance for size-congruent 2-D images was nevertheless markedly worse than for real objects viewed with either free- or fixed-head position. Second, whereas both patients showed a strong overall advantage for real objects versus images in Experiment 1, only J.W. showed a strong ROA in Experiment 2, whereas D.F.'s pattern of recognition was similar for the 2-D images and the real objects. Importantly, however, any 'carryover' effects from previously viewing the real objects before the image condition in patient D.F. in Experiments 1 and 2 did not facilitate her recognition of size-incongruent images. In summary, although we cannot rule out the possibility that these differences in performance in the 2-D image conditions reflect differences in underlying lesion anatomy, the patterns can be explained by aspects of the study design, which differed across the patients, for practical reasons, with respect to the number of exposures to the real objects before the image condition (Experiment 1), and overall stimulus set sizes (Experiment 2; see Methods).

Interestingly, patient J.W., showed marked impairments in his recognition of both 2-D and 3-D stereoscopic images of objects (Experiment 1). Although we were unable to collect convergent data from patient D.F. with stereoscopic displays, the data from J.W. are intriguing because they demonstrate for the first time that the ROA is not observed in the

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

context of 3-D images. Compared to 2-D planar images, stereoscopic images of objects have a more similar 3-D structure to real objects and they provide richer information about size and distance. Nevertheless, there are subtle cue conflicts (i.e., vergence) that reveal that the stereoscopic images are not solids. It follows that the poor recognition of 3-D stereoscopic images in patient J.W. could reflect a discounting of size information in stimuli for which cue conflicts have been detected (i.e., knowing that it is an image for which retinal size is arbitrary). The available data suggest that the ROA is an effect that is limited to solid objects. Whether ‘solidity’ is important for weighting the reliability of size information, or whether it triggers processes related to action (i.e., action planning), remains an open question for future research.

Together, these findings demonstrate a clear dissociation between size processing of real objects and 2-D (as well as 3-D) objects in patients with bilateral ventral stream lesions. Specifically, size congruency influences recognition of real objects but has little effect on the recognition of images. The results make a compelling case that the ROA is driven primarily by information about physical size, which is present in real objects, but not in images. Given the low overall level of performance in both patients, it is possible that when cues about shape and form cannot be used to guide recognition, real object size cues may provide critical top-down information that could help to facilitate selection of the correct response from among several possible alternatives (which could not otherwise be discriminated based on shape and form alone). Moreover, whereas some of the previous findings regarding the influence of depth information on object recognition in agnosia might be a result of confounded size information, it may well be the case that cues other than size, such as depth but also surface properties such as color and texture, may also play some role in the ROA (Humphrey et al., 1994). In the following section we consider in more detail the possible underlying neural mechanisms for the ROA and size-specific recognition performance in the patients.

#### 4.1 Potential mechanisms underlying the ROA

These neuropsychological data raise important questions about the underlying neural mechanisms of the ROA, and about the nature of object-related processing in human cortex. Both patients in the current study have extensive bilateral occipitotemporal lesions, together with well-documented, severe and lasting impairments in recognizing images of objects. Although there is an extension of D.F.’s lesion into parietal cortex (Bridge et al., 2013), dorsal cortex is largely structurally intact in both patients. Given that there was extensive impairment to ventral cortex in both D.F. and J.W., we suggest that the ROA is likely a result of computations in dorsal cortex.

The findings complement and extend previous neuropsychological and fMRI studies on the nature of object processing in dorsal cortex in healthy adults (Freud et al., 2017) and in patients with visual agnosia (Freud et al., 2015). With regard to the patients, Freud et al. (2015a) presented five agnosia patients with 2-D shape images that conveyed 3-D structure via pictorial depth cues of occlusion and surface shading. The authors (Freud et al., 2015) found that although the patients were unable to determine explicitly whether the geometric shapes had a possible vs. impossible geometric depth structure, they nevertheless showed

*implicit* sensitivity to depth of the images as revealed by above-chance performance on same/different shape classification and depth comparison tasks. Furthermore, two of the agnosic individuals evinced the same BOLD signature as controls in response to possible vs. impossible objects but only in dorsal and not in ventral cortical regions. Here, using real object displays, we show that patients with severe agnosia can process size-appropriate objects beyond the level of implicit sensitivity, in many cases to the level of conscious perception. Moreover, by including both familiar objects and geometric shapes in our experiments, we reveal an effect of physical size on recognition that is unique to meaningful objects (but not geometric shapes). Differences in the patients' performance for meaningful objects versus shapes are reminiscent of fMRI findings in healthy observers, showing that whereas viewing images of tools (versus baseline scrambled images) elicited bilateral activation in left pre- and post-central gyri and medial frontal gyrus, viewing unfamiliar basic shapes (versus scrambled shapes) did not activate dorsal areas (Creem-Regehr & Lee, 2005).

In summary, we have suggested that the ROA may be a product of dorsal cortex computations and the data we have offered provide novel and direct clinical neuropsychological evidence that dorsal cortex may well contribute to perception independently from ventral cortex (Konen & Kastner, 2008). These findings provide critical evidence for a functional role for dorsal cortex in visual object perception that extends beyond the implicit extraction of geometric depth information (Durand, Peeters, Norman, Todd, & Orban, 2009, Freud et al., 2015, Georgieva, Peeters, Kolster, Todd, & Orban, 2009, Orban, 2011, Verhoef, Michelet, Vogels, & Janssen, 2015), even when the task requires no explicit visuo-motor hand action towards the stimulus.

**4.1.1 Size coding in dorsal cortex**—These data suggest that it is not simply that the patients compute abstract size information about *any* object (i.e., 'the object in front of me is 3 cm tall'); rather, recognition is relatively spared for familiar objects when physical size matches the typical real-world size. One possible interpretation of these findings is that the object size information is derived by the spared dorsal cortex in the two agnosic patients. In line with the idea that dorsal cortex in humans is involved in size coding, a recent study found that levels of GABA (the primary inhibitory neurotransmitter in the human brain) selectively influence size perception (Song, Sandberg, Andersen, Blicher, & Rees, 2017).

An obvious question concerns the source or mechanism by which the size information comes to be coded in dorsal cortex. Given that the sparing of recognition performance in the patients was specific to familiar objects of appropriate size, object areas in dorsal cortex might compute detailed information about the real-world size of familiar graspable objects. Alternatively, size information may be computed in areas that are distinct from those that represent object shape in dorsal cortex. Physical size is an important visual cue for planning and executing goal-directed actions with objects (Faillenot, Decety, & Jeannerod, 1999, Gallivan, McLean, Valyear, Pettypiece, & Culham, 2011, Murata, Gallese, Lupino, Kaseda, & Sakata, 2000), as well as for determining affordances when there is no explicit plan to grasp (Gomez, Skiba, & Snow, 2017, Tucker & Ellis, 2001). Parietal cortex is tiled with areas known to be involved in hand orientation and pre-shaping (Castiello, 2005, Gallivan & Culham, 2015, Jeannerod, 1986, Jeannerod, Decety, & Michel, 1994), and some represent

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

physical size. For example, Murata et al. (2000) recorded from cells in the anterior intraparietal area (AIP) while monkeys either passively viewed or grasped 3-D graspable objects of three different sizes (small, medium and large). More than half of the visually-responsive neurons responded during object fixation ('object-type' neurons), and most were tuned for both the shape and size of the object. In the human brain, dorsal premotor cortex (Fabbri, Stubbs, Cusack, & Culham, 2016) (as well as the posterior medial temporal gyrus) has been shown to represent the size of graspable solid geometric shapes, as indicated by representational similarity analysis of fMRI data. Using ultra-high-field (7T) functional MRI, topographic representations of size have also recently been identified within bilateral human parietal cortex immediately caudal to the postcentral sulcus (Harvey, Fracasso, Petridou, & Dumoulin, 2015).

Finally, the improved recognition of size-congruent real objects in patients with bilateral ventral cortex damage relates to a relevant discussion in the literature as to whether or not visuomotor control and grasping (supported by the dorsal stream) is based on information about object size (Ganel, Chajut, & Algom, 2008), or about object location (Schot, Brenner, & Smeets, 2017, Smeets & Brenner, 2008). (Ganel et al., 2008) identified a striking dissociation between perception and action in neurologically healthy observers: Weber's law, a fundamental principle of visual perception, operated in the context of perceptual estimations of size, but was violated during online control of visually-guided grasping actions. Similar dissociations between perceiving objects and grasping them have also been identified in other behavioral investigations in healthy observers (Ganel, Tanzer, & Goodale, 2008) and in neuropsychological patients (Goodale, Milner, Jakobson, & Carey, 1991). Together, these studies suggest that different representations of object size are used for action and for perception (Ganel et al., 2008; Ganel, Chajut, Tanzer, & Algom, 2008). Others, however, have demonstrated that the digits of the grasping hand are directed independently to different locations in space (Schot et al., 2017, Smeets & Brenner, 2001), and have argued that grasping is computed based on an object's position, rather than on its size (Smeets & Brenner, 2008). Although we did not look directly at visuomotor control, since the patients in our study were never required to interact manually with the stimuli, our findings of improved performance with size congruency (supported by dorsal-stream processing) suggest that the size of (real) objects, rather than merely their location, is effectively coded in the dorsal stream. Our findings further advance previous work in this domain by showing that perception (not just action) may be based on physical size cues when ventral-stream contributions to perception are minimized or eliminated, and the stimuli are meaningful objects that afford goal-directed action.

**4.1.2 Does the ROA reflect independence or interactions between dorsal and ventral cortex?**—Although dorsal cortex may carry out size and shape analysis, it is unclear whether or how the ROA can arise from dorsal computations of size and associations with shape, without access to memory and semantics. Potentially, dorsally-derived size information could become available to ventral cortex via cortical connections, as the two visual pathways are anatomically (Yeatman et al., 2014) and functionally (Freud, Rosenthal, Ganel, & Avidan, 2015, Kiefer, Sim, Helbig, & Graf, 2011, Sim, Helbig, Graf, & Kiefer, 2015) interconnected.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

One possibility is that interaction between the dorsal- and ventral-streams is facilitated when dorsal shape representations are coupled with information about egocentric distance and physical size. In the patients, converging shape and physical size information may help to reroute information flow around (otherwise damaged) high-fidelity ventral shape representations, to get to areas such as the anterior temporal lobe (ATL) or inferior frontal cortex to facilitate naming.

The ROA could also arise from dorsal stream computation of size and shape which, when coupled with motor programming or implicit grasp preparation, generates interconnections between the dorsal and ventral visual streams or other areas linked to semantics and naming. It is well-known that the dorsal stream can size-code real objects during visually-guided actions (Goodale, Jakobson, & Keillor, 1994, Goodale et al, 1991). Both of the patients in the current study became agnosic decades ago and have presumably learned to use subtle strategies to help recognize objects; one such strategy may have been to exploit information inherent in their visually-guided actions. Although explanations based on coupling of visual shape information with motor plans might account for the ROA, what about the finding of selective sparing of perception for real-world-sized objects? For example, the patients might have made motoric accommodations with the hand to real object size that could be detected and then used the motor information to capitalize on dorsal size information. Careful inspection of the patients' behavior during the experiments, however, revealed no discernable hand movements, and presumably, these actions would need to be highly specific to support recognition. Nevertheless, it could be that the patients are able to utilize more subtle motor-related strategies to facilitate recognition that have become, in some cases, internalized and not visible to the experimenter. For example, the patients may be uniquely sensitive to patterns of eye movements or other oculomotor cues, or implicit grasp preparation for 'familiar' three-dimensional graspable objects. This explanation could account for the results of Experiment 1 because all the incongruent objects were items that could not be grasped and picked up in real life (bus, horse, table) whereas all the congruent objects were largely graspable items (clothes-peg, spoon, key). In other words, D.F. and J.W. may have been able to bootstrap recognition of size-congruent objects because these are ones that they regularly pick up and have well-developed motor routines. And it is possible that this is why, when the objects were congruent (and graspable in the real world), that the naming errors were much more likely to be for other graspable objects for which similar but incorrect oculomotor routines and grasping plans could be elicited –while this wasn't the case for the incongruent objects which would never have been associated with such routines and plans. Although this explanation could be extended to account for the results of Experiment 2 in which the real-world size of the stimuli was varied on a finer scale than Experiment 1 (because the correctly-sized graspable object would be more likely to generate associated oculomotor and implicit grasping plans), it would require a remarkable sensitivity to subtle variations in real-world size and shape.

An alternative explanation, however, is that the dorsal size information is computed with detailed object information, without requiring access to, or coupling with, ventral cortex. Wolk, Coslett and Glosser (2005) reported a case study of a patient (A.D.) with a right occipito-temporal lesion and visual agnosia. A.D. was impaired in the visual recognition of living, but not non-living, entities. Critically, the patient showed superior recognition of line-

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

drawings of non-living objects whose visual appearance conveyed important information about how they would be used. For example, objects such as scissors or hammer, that were associated with a high ‘manipulability index’, were named more accurately than items such as ashtray or table, with a lower manipulability index whose form did not predict the mode of interaction. The authors concluded that given the close relationship between an object’s form and the way in which it is manipulated, visual form information activates sensorimotor knowledge specific to the object (Wolk et al., 2005). They further argued that in patients with ventral stream damage, the preserved dorsal stream may provide information that is sufficient to facilitate recognition, without requiring access to ventral-stream stores (Wolk et al., 2005).

The notion that dorsal cortex may be able to perform the relevant computations of size and shape independently from ventral cortex is further supported by recent data from human neuroimaging. Freud, Culham, Plaut, and Behrmann (2017) demonstrated the large-scale topography of sensitivity to objects in both ventral and dorsal cortex under conditions where visuomotor transformations were not implicated. Observers viewed intact images of objects as well as parametrically scrambled or distorted images of the same objects and in every voxel, the slope of the beta weights across the levels of scrambling was derived as an index of shape sensitivity. Very similar signatures of shape sensitivity were obtained in dorsal and ventral cortex with a gradient of sensitivity increasing from posterior to more anterior regions and then decreasing again in the most anterior ventral and dorsal regions. Moreover, the shape sensitivity indices were well correlated with behavioral recognition accuracy evaluated outside the scanner. The detailed object representations in dorsal cortex, together with the dorsally-derived size information, may give rise to the ROA.

Finally, as we have argued above, our results suggest that object size is processed differently in the dorsal and ventral visual streams. The differences in recognition of objects and images evinced by the patients could reflect the use of oculomotor distance cues for the real objects, but less so for images (as evident in Experiment 1 by the progressive decline in recognition performance for solids, but not images, with the systematic elimination of depth cues). It could be, therefore, that object size in the dorsal stream is computed based on oculomotor distance information, while size in the ventral stream is computed based on pictorial cues. Such size-selectivity in dorsal cortex deviates markedly from object processing within the ventral pathway, which shows a posterior-to-anterior gradient of increasing invariance in object format, position, and size (Grill-Spector et al., 1999, Kourtzi & Kanwisher, 2000, Vuilleumier et al., 2002, Xu, 2018a, 2018b).

Given the above considerations about possible mechanisms for the size-selective ROA in agnosia patients, fruitful avenues for future research in this domain will be to compare gaze patterns during object identification for stimuli of different sizes, and in different display formats. Other studies will no doubt examine whether eye movements in patients with agnosia differ for items that have, versus do not have, well-defined motor routines (such as living versus non-living entities) (Wolk et al., 2005). Double dissociations between recognition of size-congruent versus size-incongruent objects in different formats could also be investigated in patients with dorsal versus ventral-stream lesions. Ultimately, detailed functional neuroimaging studies in patients with bilateral damage to ventral-stream object

processing areas, such as LOC, will shed critical insights into the neural basis of size and shape coding of real-world objects and images within the dorsal and ventral processing streams.

#### 4.2 Summary

In summary, although healthy observers can show performance decrements when the expected size of object images deviates dramatically from other objects in a scene (Eckstein, Koehler, Welbourne, & Akbas, 2017, Konkle & Oliva, 2012), here we show that in patients with severe visual agnosia following bilateral ventral stream lesions, there can be a remarkable sparing of recognition of real-world objects. Importantly, this real object advantage in recognition breaks down when the object's physical size deviates, even minimally, from typical real-world size. These neuropsychological findings provide convergent evidence that real-world objects have unique effects on attention (Gomez et al., 2017), memory (Snow, Skiba, Coleman, & Berryhill, 2014), decision-making (Romero, Compton, Yang, & Snow, 2017) and brain responses (Snow et al., 2011), and underscore the importance of using richer ecologically-relevant stimuli to understand perception and neural coding in the human brain, especially in uncovering object-related information in dorsal cortex.

### ACKNOWLEDGMENTS

This work was supported by grants to J.C. Snow from the National Eye Institute of the National Institutes of Health (NIH) under award number R01EY026701. The research was also supported by the National Science Foundation (NSF) [grant number 1632849]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NSF. We thank Melvyn Goodale for helpful comments on the manuscript.

### REFERENCES

Banks MS, Read JC, Allison RS, & Watt SJ (2012). Stereoscopy and the human visual system. *SMPTE Motion Imaging J*, 121(4), 24–43. doi:10.5594/j18173 [PubMed: 23144596]

Bridge H, Thomas OM, Minini L, Cavina-Pratesi C, Milner AD, & Parker AJ (2013). Structural and functional changes across the visual cortex of a patient with visual form agnosia. *J Neurosci*, 33(31), 12779–12791. doi:10.1523/jneurosci.4853-12.2013 [PubMed: 23904613]

Castiello U (2005). The neuroscience of grasping. *Nat Rev Neurosci*, 6(9), 726–736. doi:10.1038/nrn1744 [PubMed: 16100518]

Cavina-Pratesi C, Kentridge RW, Heywood CA, & Milner AD (2010). Separate processing of texture and form in the ventral stream: evidence from fMRI and visual agnosia. *Cereb Cortex*, (New York, NY), 20(2), 433–446. doi:10.1093/cercor/bhp111

Chainay H, & Humphreys GW (2001). The real-object advantage in agnosia: Evidence for a role of surface and depth information in object recognition. *Cogn Neuropsychol*, 18(2), 175–191. doi:10.1080/02643290042000062 [PubMed: 20945210]

Creem-Regehr SH, & Lee JN (2005). Neural representations of graspable objects: are tools special? *Brain Res Cogn Brain Res*, 22(3), 457–469. doi:10.1016/j.cogbrainres.2004.10.006 [PubMed: 15722215]

Durand J-B, Peeters R, Norman JF, Todd JT, & Orban GA (2009). Parietal regions processing visual 3D shape extracted from disparity. *Neuroimage*, 46(4), 1114–1126. doi:10.1016/j.neuroimage.2009.03.023 [PubMed: 19303937]

Eckstein MP, Koehler K, Welbourne LE, & Akbas E (2017). Humans, but not deep neural networks, often miss giant targets in scenes. *Curr Biol*, 27(18), 2827–2832.e2823. doi:10.1016/j.cub.2017.07.068 [PubMed: 28889976]

Fabbri S, Stubbs KM, Cusack R, & Culham JC (2016). Disentangling Representations of Object and Grasp Properties in the Human Brain. *J Neurosci*, 36(29), 7648–7662. doi:10.1523/jneurosci.0313-16.2016 [PubMed: 27445143]

Faillenot I, Decety J, & Jeannerod M (1999). Human brain activity related to the perception of spatial features of objects. *Neuroimage*, 10(2), 114–124. doi:10.1006/nimg.1999.0449 [PubMed: 10498441]

Farah MJ (1990). Visual agnosia: MIT press.

Freud E, Culham JC, Plaut DC, & Behrmann M (2017). The large-scale organization of shape processing in the ventral and dorsal pathways. *eLife*, 6, e27576. doi:10.7554/eLife.27576 [PubMed: 28980938]

Freud E, Ganel T, Shelef I, Hammer MD, Avidan G, & Behrmann M (2015). Three-dimensional representations of objects in dorsal cortex are dissociable from those in ventral cortex. *Cereb Cortex*. doi:10.1093/cercor/bhv229

Freud E, Rosenthal G, Ganel T, & Avidan G (2015). Sensitivity to object impossibility in the human visual cortex: evidence from functional connectivity. *J Cogn Neurosci*, 27(5), 1029–1043. doi:10.1162/jocn\_a\_00753 [PubMed: 25390203]

Gallivan JP, & Culham JC (2015). Neural coding within human brain areas involved in actions. *Curr Opin Neurobiol*, 33, 141–149. doi:10.1016/j.conb.2015.03.012 [PubMed: 25876179]

Gallivan JP, McLean DA, Valyear KF, Pettypiece CE, & Culham JC (2011). Decoding action intentions from preparatory brain activity in human parieto-frontal networks. *J Cogn Neurosci*, 31(26), 9599–9610. doi:10.1523/jneurosci.0080-11.2011

Ganel T, Chajut E, & Algom D (2008). Visual coding for action violates fundamental psychophysical principles. *Curr Biol*, 18(14), R599–601. doi:10.1016/j.cub.2008.04.052 [PubMed: 18644333]

Ganel T, Chajut E, Tanzer M, & Algom D (2008). Response: When does grasping escape Weber's law? *Curr Biol*, 18(23), R1090–R1091.

Ganel T, Tanzer M, & Goodale MA (2008). A double dissociation between action and perception in the context of visual illusions: opposite effects of real and illusory size. *Psychol Sci*, 19(3), 221–225. doi:10.1111/j.1467-9280.2008.02071.x [PubMed: 18315792]

Georgieva S, Peeters R, Kolster H, Todd JT, & Orban GA (2009). The processing of three-dimensional shape from disparity in the human brain. *J Neurosci*, 29(3), 727–742. doi:10.1523/jneurosci.4753-08.2009 [PubMed: 19158299]

Gibson JJ (1979). The Ecological Approach to Visual Perception. Boston, MA: Houghton Mifflin.

Gomez MA, Skiba RM, & Snow JC (2017). Graspable objects grab attention more than images do. *Psychol Sci*, 956797617730599. doi:10.1177/0956797617730599

Goodale MA, Jakobson LS, & Keillor JM (1994). Differences in the visual control of pantomimed and natural grasping movements. *Neuropsychologia*, 32(10), 1159–1178. [PubMed: 7845558]

Goodale MA, & Milner AD (1992). Separate visual pathways for perception and action. *Trends Neurosci*, 15(1), 20–25. [PubMed: 1374953]

Goodale MA, Milner AD, Jakobson LS, & Carey DP (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, 349(6305), 154–156. doi:10.1038/349154a0 [PubMed: 1986306]

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, & Malach R (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, 24(1), 187–203. [PubMed: 10677037]

Harvey BM, Fracasso A, Petridou N, & Dumoulin SO (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proc Natl Acad Sci U S A*, 112(44), 13525–13530. doi:10.1073/pnas.1515414112 [PubMed: 26483452]

Hiraoka K, Suzuki K, Hirayama K, & Mori E (2009). Visual agnosia for line drawings and silhouettes without apparent impairment of real-object recognition: a case report. *Behav Neurol*, 21(3), 187–192. doi:10.3233/ben-2009-0244 [PubMed: 19996516]

Humphrey GK, Goodale MA, Jakobson LS, & Servos P (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. *Perception*, 23(12), 1457–1481. [PubMed: 7792135]

Humphreys GW (2013). Beyond serial stages for attentional selection: the critical role of action. *Action Science*, 229–251.

James TW, Culham J, Humphrey GK, Milner AD, & Goodale MA (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain*, 126(Pt 11), 2463–2475. doi:10.1093/brain/awg248 [PubMed: 14506065]

Jeannerod M (1986). The formation of finger grip during prehension. A cortically mediated visuomotor pattern. *Behav Brain Res*, 19(2), 99–116. doi:10.1016/0166-4328(86)90008-2 [PubMed: 3964409]

Jeannerod M, Decety J, & Michel F (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia*, 32(4), 369–380. [PubMed: 8047246]

Julesz B (2006). Foundations of cyclopean perception. Cambridge, MA: MIT Press.

Kastner S, & Ungerleider LG (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia*, 39(12), 1263–1276. [PubMed: 11566310]

Kiefer M, Sim EJ, Helbig H, & Graf M (2011). Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. *J Cogn Neurosci*, 23(8), 1864–1874. doi:10.1162/jocn.2010.21543 [PubMed: 20617882]

Kim HR, Angelaki DE, & DeAngelis GC (2016). The neural basis of depth perception from motion parallax. *Philos Trans R Soc Lond B Biol Sci*, 371(1697). doi:10.1098/rstb.2015.0256

Konen CS, Behrmann M, Nishimura M, & Kastner S (2011). The functional neuroanatomy of object agnosia: a case study. *Neuron*, 71(1), 49–60. doi:10.1016/j.neuron.2011.05.030 [PubMed: 21745637]

Konen CS, & Kastner S (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci*, 11(2), 224–231. doi:10.1038/nn2036 [PubMed: 18193041]

Konkle T, & 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(3), 561–569. doi:10.1037/a0028294 [PubMed: 22545601]

Kourtzi Z, & Kanwisher N (2000). Cortical regions involved in perceiving object shape. *J Neurosci*, 20(9), 3310–3318. [PubMed: 10777794]

Malach R, Reppas J, Benson R, Kwong K, Jiang H, Kennedy W, ... Tootell R (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A*, 92(18), 8135–8139. [PubMed: 7667258]

Mapelli D, & Behrmann M (1997). The role of color in object recognition: Evidence from visual agnosia. *Neurocase*, 3(4), 237–247.

Marotta JJ, Behrmann M, & Goodale MA (1997). The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. *Exp Brain Res*, 116(1), 113–121. [PubMed: 9305820]

Milner A, Perrett D, Johnston R, Benson P, Jordan T, Heeley D, ... Terazzi E (1991). Perception and action in 'visual form agnosia'. *Brain*, 114(1), 405–428. [PubMed: 2004249]

Murata A, Gallesse V, Kaseda M, & Sakata H (1996). Parietal neurons related to memory-guided hand manipulation. *J Neurophysiol*, 75(5), 2180–2186. doi:10.1152/jn.1996.75.5.2180 [PubMed: 8734616]

Murata A, Gallesse V, Luppino G, Kaseda M, & Sakata H (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol*, 83(5), 2580–2601. [PubMed: 10805659]

Orban GA (2011). The extraction of 3D shape in the visual system of human and nonhuman primates. *Annu Rev Neurosci*, 34(1), 361–388. doi:10.1146/annurev-neuro-061010-113819 [PubMed: 21456959]

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

Read JC, Phillipson GP, Serrano-Pedraza I, Milner AD, & Parker AJ (2010). Stereoscopic vision in the absence of the lateral occipital cortex. *PLoS ONE*, 5(9), e12608. doi:10.1371/journal.pone.0012608 [PubMed: 20830303]

Riddoch MJ, & Humphreys GW (1987). A case of integrative visual agnosia. *Brain*, 110 (Pt 6), 1431–1462. [PubMed: 3427396]

Philos Trans R Soc Lond B Biol Sci, 369(1644), 20130420. doi:10.1098/rstb.2013.0420

Romero CA, Compton MT, Yang Y, & Snow JC (2017). The real deal: Willingness-to-pay and satiety expectations are greater for real foods versus their images. *Cortex*. doi:10.1016/j.cortex.2017.11.010

Rosenthal O, & Behrmann M (2006). Acquiring long-term representations of visual classes following extensive extrastriate damage. *Neuropsychologia*, 44(5), 799–815. doi:10.1016/j.neuropsychologia.2005.07.010 [PubMed: 16137724]

Sakata H, Taira M, Murata A, & Mine S (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cereb Cortex* (New York, NY), 5(5), 429–438.

Schot WD, Brenner E, & Smeets JB (2017). Unusual prism adaptation reveals how grasping is controlled. *eLife*, 6. doi:10.7554/eLife.21440

Servos P, Goodale MA, & Humphrey GK (1993). The drawing of objects by a visual form agnosic: contribution of surface properties and memorial representations. *Neuropsychologia*, 31(3), 251–259. [PubMed: 8492878]

Sim EJ, Helbig HB, Graf M, & Kiefer M (2015). When action observation facilitates visual perception: activation in visuo-motor areas contributes to object recognition. *Cereb Cortex*, 25(9), 2907–2918. doi:10.1093/cercor/bhu087 [PubMed: 24794918]

Snow JC, Pettypiece CE, McAdam TD, McLean AD, Stroman PW, Goodale MA, & Culham JC (2011). Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. *Sci Rep*, 1, 130. doi:10.1038/srep00130 [PubMed: 22355647]

Snow JC, Skiba RM, Coleman TL, & Berryhill ME (2014). Real-world objects are more memorable than photographs of objects. *Front Hum Neurosci*, 8, 837. doi:10.3389/fnhum.2014.00837 [PubMed: 25368568]

Smeets JB, & Brenner E (2001). Independent movements of the digits in grasping. *Exp Brain Res*, 139(1), 92–100. [PubMed: 11482847]

Smeets JB, & Brenner E (2008). Grasping Weber's law. *Curr Biol*, 18(23), R10891090; author reply R1090–1081. doi:10.1016/j.cub.2008.10.008

Song C, Sandberg K, Andersen LM, Blicher JU, & Rees G (2017). Human occipital and parietal GABA selectively influence visual perception of orientation and size. *J Neurosci*, 37(37), 8929–8937. doi:10.1523/jneurosci.3945-16.2017 [PubMed: 28821653]

Taira M, Mine S, Georgopoulos AP, Murata A, & Sakata H (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp Brain Res*, 83(1), 29–36. [PubMed: 2073947]

Tucker M, & Ellis R (2001). The potentiation of grasp types during visual object categorization. *Vis Cogn*, 8, 769–800.

Turnbull OH, Driver J, & McCarthy RA (2004). 2D but not 3D: pictorial-depth deficits in a case of visual agnosia. *Cortex*, 40(4–5), 723–738. [PubMed: 15505981]

Verhoef BE, Michelet P, Vogels R, & Janssen P (2015). Choice-related activity in the anterior intraparietal area during 3-D structure categorization. *J Cogn Neurosci*, 27(6), 1104–1115. doi:10.1162/jocn\_a\_00773 [PubMed: 25514653]

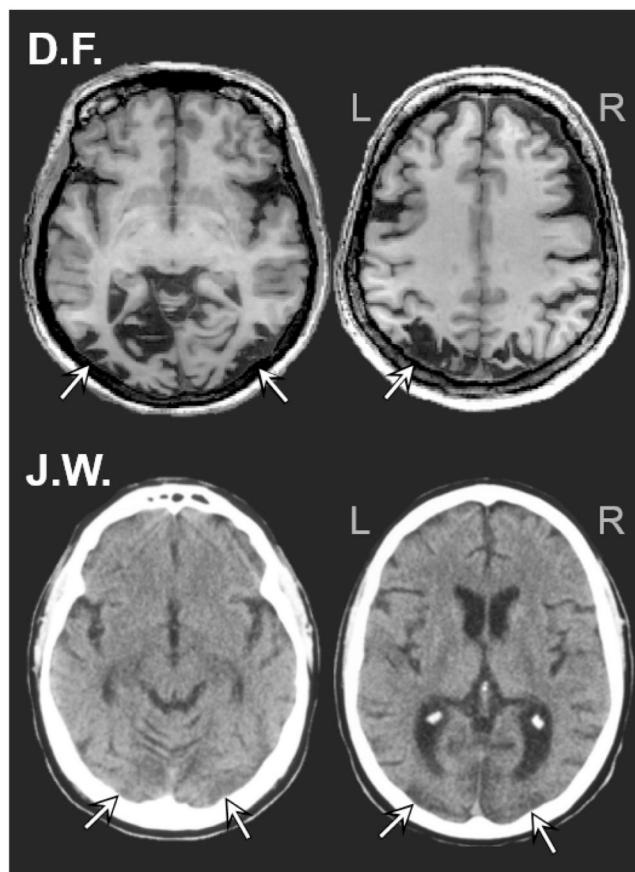
Vuilleumier P, Henson RN, Driver J, & Dolan RJ (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nat Neurosci*, 5(5), 491–499. doi:10.1038/nn839 [PubMed: 11967545]

Xu Y (2018a). The posterior parietal cortex in adaptive visual processing. *Trends Neurosci*. doi:10.1016/j.tins.2018.07.012

Xu Y (2018b). A tale of two visual systems: invariant and adaptive visual information representations in the primate brain. *Annu Rev Vis Sci*. doi:10.1146/annurev-vision-091517-033954

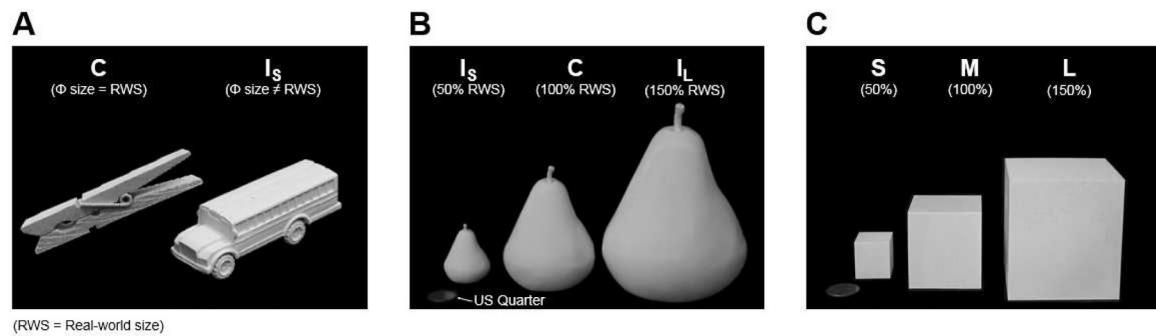
Wolk DA, Coslett HB, & Glosser G (2005). The role of sensory-motor information in object recognition: evidence from category-specific visual agnosia. *Brain Lang*, 94(2), 131–146. doi:10.1016/j.bandl.2004.10.015 [PubMed: 15896389]

Yeatman JD, Weiner KS, Pestilli F, Rokem A, Mezer A, & Wandell BA (2014). The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. *Proc Natl Acad Sci U S A*, 111(48), E5214–5223. doi:10.1073/pnas.1418503111 [PubMed: 25404310]



**Figure 1: Anatomical scans for patients D.F. and J.W.**

High resolution MRI scan of D.F.'s brain (*upper panel*) and a computed tomography (CT) scan of J.W.'s brain (*lower panel*). Axial slices are displayed in ascending order from left (ventral) to right (dorsal).

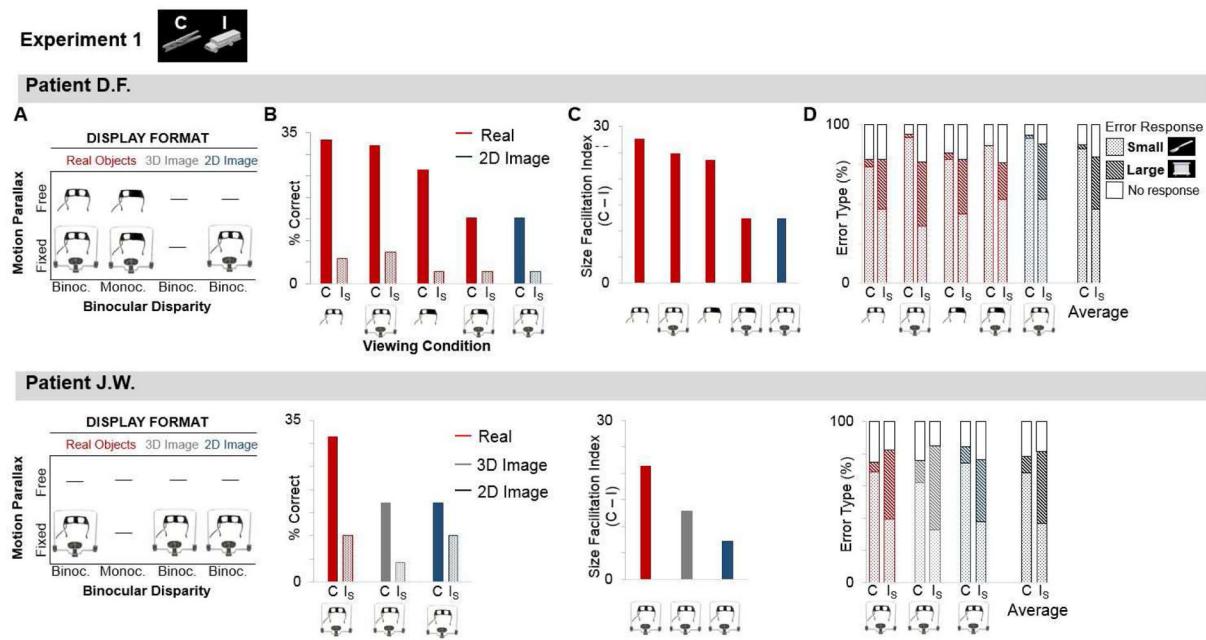


**Figure 2: Stimuli used in Experiments 1–3.**

(A) Stimuli in Experiment 1: Half were exemplars whose physical size ( $\Phi$ ) was congruent (C) with typical real-world size (RWS); the remainder were ‘miniaturized’ items whose size was incongruent small ( $I_S$ ) with real-world size.

(B) Stimuli in Experiment 2: 3-D-printed objects whose size was scaled to be smaller than (50%,  $I_S$ ), congruent with (100%, C), or larger than (150%,  $I_L$ ), typical real-world size.

(C) Stimuli in Experiment 3: Common geometric shapes without strong real-world size associations.



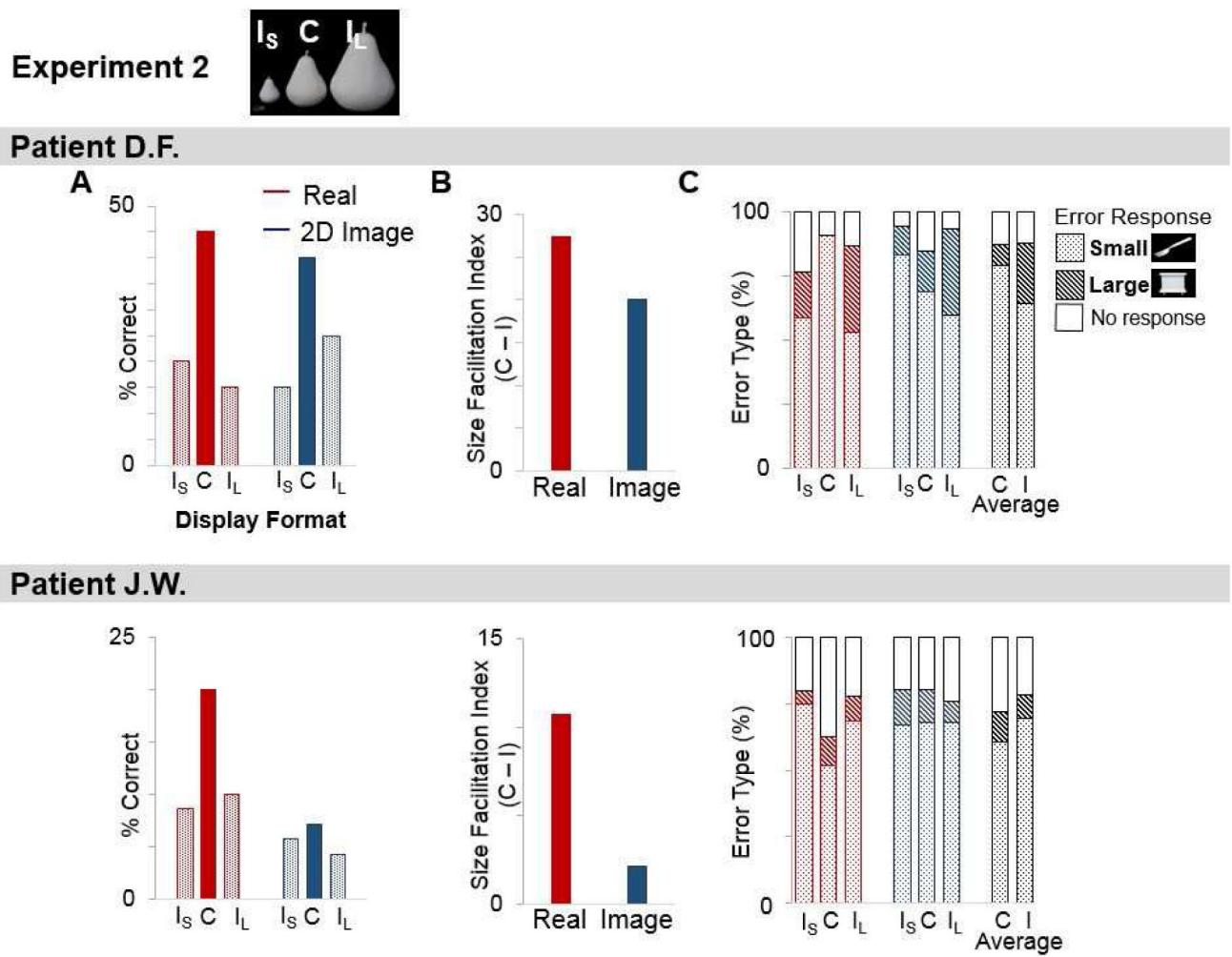
**Figure 3: Recognition of Size-Congruent (C) vs. Size-Incongruent (Is) Real Objects and Images**

(A) The patients (D.F., *upper panel*; J.W., *lower panel*) named objects presented in different Display Formats (real objects, 3-D images, 2-D 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, Is) 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).

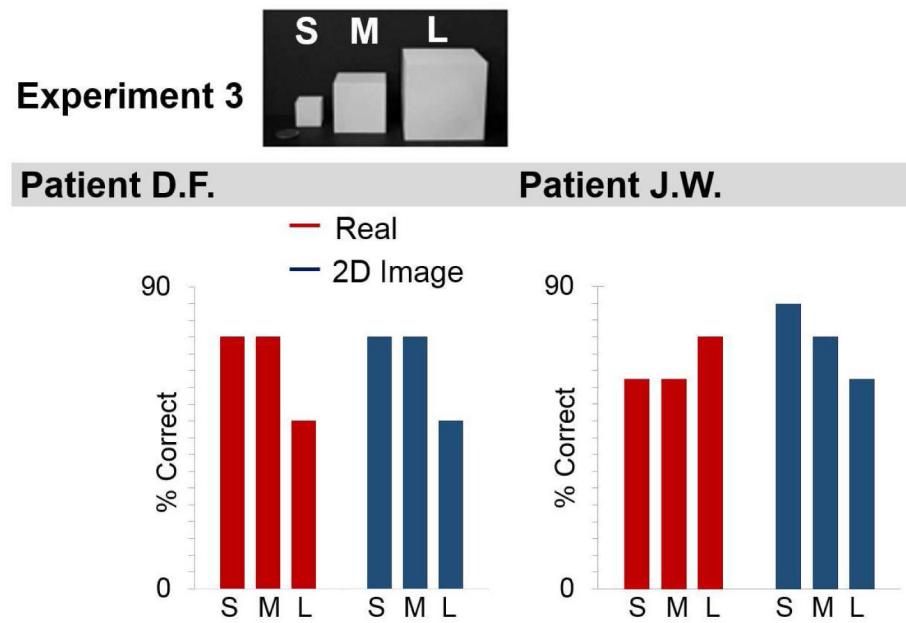


**Figure 4. Recognition of Size-Congruent (C) vs. Size-Incongruent Small ( $I_S$ ) and Large ( $I_L$ ) Objects**

(A) % correct responses by D.F. (*upper panel*) and J.W. (*lower panel*) for real objects and 2-D 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. 2-D 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).



**Figure 5: Recognition of Geometric Shapes**

% correct responses made by D.F. (*left panel*) and J.W. (*right panel*) in naming geometric shapes displayed as real objects or 2-D images, separately for small (S), medium (M), and large (L) sizes.