



Visual Awareness Is Limited by the Representational Architecture of the Visual System

Citation

Cohen, Michael A., Ken Nakayama, Talia Konkle, Mirta Stantić, and George A. Alvarez. 2015. Visual Awareness Is Limited by the Representational Architecture of the Visual System. *Journal of Cognitive Neuroscience* 27, no. 11: 2240–2252. doi:10.1162/jocn_a_00855.

Published Version

doi:10.1162/jocn_a_00855

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:30780170>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Visual Awareness Is Limited by the Representational Architecture of the Visual System

Michael A. Cohen^{1,2}, Ken Nakayama², Talia Konkle²,
Mirta Stantić², and George A. Alvarez²

Abstract

■ Visual perception and awareness have strict limitations. We suggest that one source of these limitations is the representational architecture of the visual system. Under this view, the extent to which items activate the same neural channels constrains the amount of information that can be processed by the visual system and ultimately reach awareness. Here, we measured how well stimuli from different categories (e.g., faces and cars) blocked one another from reaching awareness using two distinct paradigms that render stimuli invisible: visual

masking and continuous flash suppression. Next, we used fMRI to measure the similarity of the neural responses elicited by these categories across the entire visual hierarchy. Overall, we found strong brain–behavior correlations within the ventral pathway, weaker correlations in the dorsal pathway, and no correlations in early visual cortex (V1–V3). These results suggest that the organization of higher level visual cortex constrains visual awareness and the overall processing capacity of visual cognition. ■

INTRODUCTION

Visual awareness is surprisingly limited. These limitations have been demonstrated with a wide variety of psychophysical paradigms (for a review, see Kim & Blake, 2005). However, these dissimilar paradigms broadly fall into two distinct categories. First, there are perceptual manipulations that suppress sensory signals in the earlier parts of the visual system (e.g., visual masking or crowding; Kanai, Walsh, & Tseng, 2010). Second, there are attentional manipulations, which cause items to go unnoticed because of a lack of processing resources (e.g., change blindness or the attentional blink; Cohen, Alvarez, & Nakayama, 2011). Together, these different paradigms have been used to identify different neural regions associated with the limits of visual awareness. For example, paradigms that manipulate the strength of perceptual signals focus on early visual regions such as areas V1–V4 (Yuval-Greenberg & Heeger, 2013; Anderson, Dakin, Schwarzkopf, Rees, & Greenwood, 2012; Tong et al., 2006). Meanwhile, paradigms that manipulate attention have focused on the processing capacity of the frontoparietal network (Dehaene & Changeux, 2011; Lamme, 2010; Tononi & Koch, 2008). Although these neural regions undoubtedly play a role in determining what can be accessed by awareness, there are other possible sources that have yet to be thoroughly explored.

Here, we asked if visual awareness is limited by the representational architecture of the higher level visual system (i.e., beyond V1–V3). Previous neuroimaging

studies have identified several large-scale structures across higher level visual cortex involved in representing categories such as faces, bodies, and objects (Kanwisher, 2010) as well as superordinate categories such as animacy and real-world size (Konkle & Caramazza, 2013; Konkle & Oliva, 2012). It has recently been suggested that these structures are organized as an optimal solution to the problem of rapid and invariant object recognition (Grill-Spector & Weiner, 2014; Yamins et al., 2014). We suggest that, although this organization may be well suited for recognizing objects across a variety of changes in appearance, it actually imposes a limitation on the capacity of visual awareness. Under this view, these large-scale neural structures shape the underlying cognitive architecture of the visual system and form channels through which information is processed. Each channel has a finite processing capacity, and different stimuli elicit responses in these channels to varying degrees. When items elicit relatively high overlap among these channels, different bits of information will interfere with one another, leading to less information being available for conscious processing. When there is less activation overlap, these channels can operate alongside one another with minimal interference, increasing the amount of information that can be accessed by awareness. If this idea is correct, the degree to which information can reach awareness should correlate with the similarity of the neural responses involved in representing that information (Cohen, Konkle, Nakayama, & Alvarez, 2014; Cohen, Konkle, Rhee, Nakayama, & Alvarez, 2014).

We tested this hypothesis using stimuli that are difficult to categorize based on low-level features (e.g., luminance,

¹Massachusetts Institute of Technology, ²Harvard University

contrast) but are easy to distinguish at the category level (e.g., faces, cars). First, we used two distinct behavioral paradigms that render stimuli invisible: forward/backward masking (Breitmeyer & Ogmen, 2006) and continuous flash suppression (Lin & He, 2009; Tsuchiya & Koch, 2005). In the masking experiment, items from five categories—bodies, buildings, cars, chairs, and faces—served as both the target and the mask in all possible combinations. We used a staircase procedure to estimate the presentation durations necessary for each category pairing (e.g., detecting a car among faces) to result in equal behavioral performance across all conditions. In the continuous flash suppression experiment, we devised a novel variant of the paradigm to measure the time needed for one category to break suppression from another category (i.e., how long it takes for a car to break through suppression by faces). In both experiments, we found significant differences in how well these categories blocked one another from visual awareness.

To determine the extent to which these stimuli activate the same processing channels, we used fMRI to measure the similarity of the neural response patterns elicited by these categories across the visual hierarchy. In this case, voxels serve as a proxy for processing channels. The similarity of neural patterns across large swaths of voxels serves as our measure of channel overlap. We used representational similarity analysis to determine the relationship between the behavioral and neural measures (Kriegeskorte, Mur, & Bandettini, 2008). Across both experiments, we found strong brain–behavior correlations in ventral and lateral occipitotemporal cortex, weaker correlations in occipitoparietal cortex, and no correlations in early visual cortex (V1–V3). These results suggest that the organization of higher level visual cortex imposes a limit on the amount of information that can be accessed by visual awareness.

EXPERIMENT 1: MASKING

Methods

Participants

Twenty participants performed the behavioral experiment, and six participants performed the neuroimaging experiment.

Stimuli

Stimuli were images of bodies, buildings, cars, chairs, and faces with 30 exemplars in each category. Items were specifically selected to be as visually variable as possible and were controlled to maximize differences in low-level features across items within a category to target a higher level of representation. For example, the face set was composed of people who were of different ages, races, and genders, with variations in hairstyles and looking directions with respect to the camera. In addition, to

eliminate arbitrary visual differences between categories, all images were grayscaled and were normalized on their intensity histogram (i.e., contrast and luminance) and power at all spatial frequencies and orientations across the entire image using the SHINE toolbox (Willenbockel et al., 2010; Figure 1). These steps were taken to minimize the possibility of participants performing the task by relying on low-level features that might differentiate the stimulus categories (VanRullen, 2006).

The fMRI experiment was part of a broader project within our laboratory and included more object categories than were ultimately used for this study. As it is well known that the effectiveness of a mask increases as a function of spatial overlap relative to the target (Schiller, 1966), our aim was to explore masking and flash suppression mechanisms above and beyond the degree of spatial overlap. To this end, we selected the object categories that had the highest overlapping spatial footprint. To determine which categories to include, we overlaid the images within each category and calculated the total image area covered by at least 50% of the exemplars. The five categories with the greatest total area were selected for the experiment, such that there were 10 total category pairings, which could all be tested within the time frame of the experiment. Of the nine categories that were scanned, bodies, buildings, cars, chairs, and faces were included in this study, whereas cats, fish, hammers, and phones were excluded from all analyses.

Masking Task

Participants performed a visual masking task in which they had to detect a target item among several rapidly presented forward- and backward-masking items. A stream of 10 images was shown in immediate succession in the center of the display (Figure 2A). Participants were told that, on half of the trials, all images within the stream were all from one category (e.g., faces), whereas on the other half of the trials, one item from another category

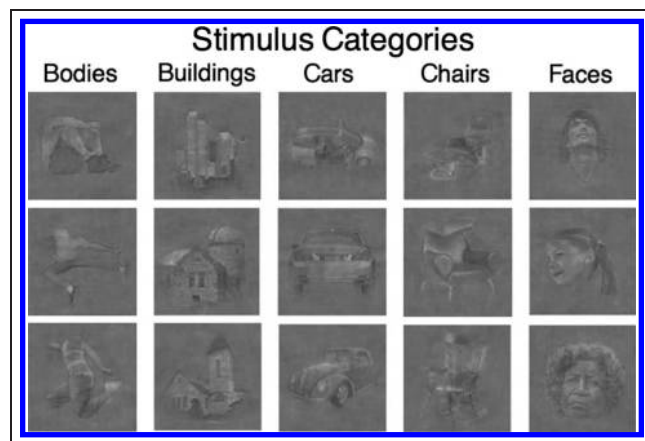


Figure 1. Example images from each of the five categories used in the behavioral and neuroimaging experiments.

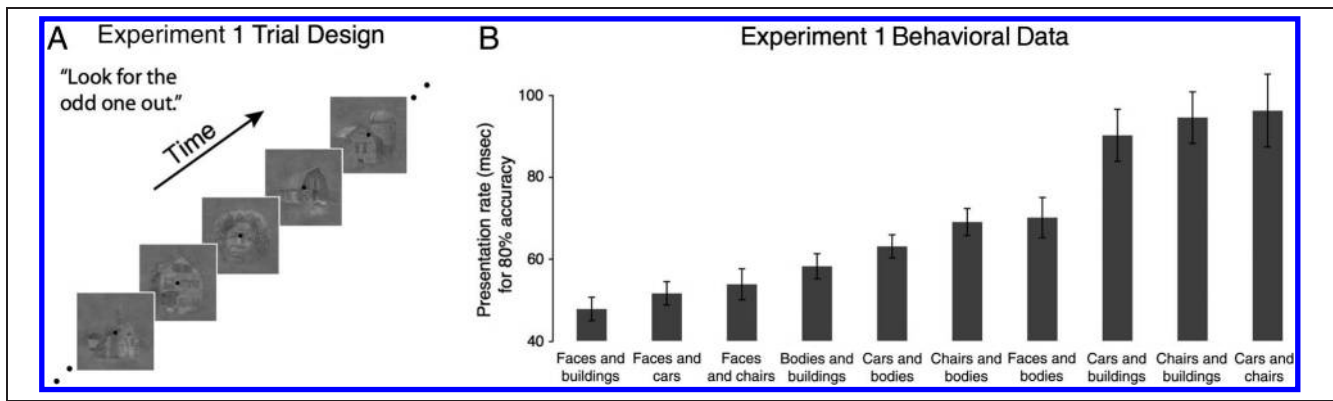


Figure 2. (A) Representation of a sample trial in which one target—a face—is shown among nine distracting masks (buildings), presented in rapid succession at the center of the display. At the beginning of each block, participants were told what two categories would be presented and were instructed to look for the odd one out, but which category was the odd-one-out varied randomly from trial to trial. (B) Results from behavioral Experiment 1. Estimated presentation durations needed to achieve equal performance (80% accuracy) for each category pairing are shown on the y axis. Each bar corresponds to a particular category pairing. Error bars reflect within-subject *SEM*.

would be embedded in the stream (e.g., a building). This oddball image from another category was the target. The task was to indicate whether there was a target.

On each trial, a red fixation dot ($\sim 0.1^\circ$) would appear in the center of the screen for 500 msec. Immediately afterward, the fixation dot turned black, and the 10 images would be shown with the fixation dot remaining on every image. When a target was present, it was either the fourth, fifth, sixth, or seventh item presented in the stream, with the target appearing in each temporal position equally often. At the end of a trial, participants gave their response by pressing a button on the keyboard. Visual feedback was immediately given for 500 msec. Participants had to press a key to proceed to the next trial.

There were 10 experimental blocks with 10 practice trials and 96 experimental trials. Each block was defined by its particular category pairing (e.g., buildings and faces, cars and chairs) such that all 10 possible category pairings were measured in each participant. The order in which the category pairings were shown was counterbalanced with a balanced Latin square design. Within a block, both categories served as the target and the masks equally often (e.g., buildings masking faces and faces masking buildings). A target appeared on half of the trials for both target-mask configurations. All trial types were randomly ordered within a block, so participants were unaware of what the target-mask configuration would be or if a target would be present from one trial to the next. We chose this design so that participants could not tune their attention to one particular category and suppress the other, forcing the two categories to interfere with one another in a more equitable fashion.

During the experiment, a staircase procedure was used to adaptively change the presentation duration of all items based on the accuracy of participants' response. The initial presentation rate of the practice trials was 112 msec per item, whereas the initial presentation rate of the experimental trials was 64 msec per item. Two in-

dependent staircases were interleaved within a block for each target-mask combination (e.g., face target-building mask vs. building target-face mask) and resulted in two estimates of the presentation rate that would yield 80% performance. The data from each block (i.e., the presentation rates and accuracy on each trial) were subsequently analyzed using QUEST, a Bayesian adaptive staircase procedure, to determine the final presentation duration estimates for both conditions within a block (Watson & Pelli, 1983). Those estimates were then averaged together to form a "category-pairing estimate."

Stimuli were presented on a 15.5-in. Nanao FlexScan T2-17TS monitor with a refresh rate of 120 Hz and a screen resolution of 800×600 and were created and controlled with MATLAB (The MathWorks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). It should be noted that the refresh rate of the monitor constrained the possible presentation times. At 120 Hz, each frame is shown for 8 msec; thus, as the staircase procedure changed the presentation duration in response to the participants' performance, changes were made in increments of 8 msec. Participants sat approximately 57 cm away from the display such that 1 cm on the screen subtended 1° of visual angle. All stimuli were square $5.9^\circ \times 5.9^\circ$ images.

Neuroimaging Procedure and Analyses

Six participants who did not perform the masking task were scanned using fMRI. Within the experimental runs, participants passively viewed the same items used in the behavioral experiment in a blocked design, with each block composed of images from a single category. Participants' only task was to perform a simple vigilance task to press a button when a red circle appeared around an item.

fMRI acquisition. Structural and functional imaging data were collected on a 3-T Siemens Trio scanner (Siemens,

Erlangen, Germany) at the Harvard University Center for Brain Sciences. Structural data were obtained in 176 axial slices with $1 \times 1 \times 1$ mm voxel resolution, repetition time = 2200 msec. Functional BOLD data were obtained using a gradient-echo echo-planar pulse sequence (33 axial slices parallel to the AC-PC line, matrix = 70×70 , field of view = 256×256 mm, voxel resolution = $3.1 \times 3.1 \times 3.1$ mm, gap thickness = 0.62 mm, repetition time = 2000 msec, echo time = 60 msec, flip angle = 90°). A 32-channel phased-array head coil was used. Stimuli were generated using the Psychophysics toolbox for MATLAB and displayed with an LCD projector onto a screen in the scanner that participants viewed via a mirror attached to the head coil.

Experimental runs. Participants viewed images of nine categories: bodies, buildings, cats, cars, chairs, faces, fish, hammers, and phones. All images of bodies, buildings, cars, chairs, and faces were the same images as those used in the behavioral masking experiment. Stimuli were presented in a rapid block design with each block corresponding to a particular category. Within a run, there were 90 total blocks of 4 sec each, with 10 blocks per category in a run. Within each block, six category exemplars were presented for 667 msec each. Blank periods occurred between blocks that lasted 2, 4, or 6 sec. The order in which the 90 blocks were presented and the number and duration of the blank periods were set using Optseq, which determines the optimal presentation of events for rapid-presentation fMRI (<http://surfer.nmr.mgh.harvard.edu/optseq/>). Participants were instructed to maintain fixation on a central cross and perform a vigilance task, pressing a button when a red circle appeared around an image. The red circle appeared in 40% of the blocks randomly on Image 2, 3, 4, or 5 of that block.

Meridian map runs. Participants were instructed to maintain fixation and were shown blocks of flickering black-and-white checkerboard wedge stimuli, oriented along either the vertical or horizontal meridian (Wandell, 1999; Sereno et al., 1995). The apex of each wedge was at fixation, and the base extended to 8° of visual angle in the periphery, with a width of 4.42° . The checkerboard pattern flickered at 8 Hz. The run consisted of four vertical and four horizontal meridian blocks. Each stimulus block was 12 sec with a 12-sec intervening blank period. The orientation of the stimuli (vertical vs. horizontal) alternated from one block to the other.

Localizer runs. Participants performed a 1-back repetition detection task with blocks of faces, bodies, scenes, objects, and scrambled objects. Stimuli in these runs were different from those in the experimental runs. Each run consisted of 10 stimulus blocks of 16 sec, with intervening 12-sec blank periods. Each category was presented twice per run, with the order of the stimulus blocks counterbalanced in a mirror reverse manner

(e.g., face, body, scene, object, scrambled, scrambled, objects, scene, body, face). Within a block, each item was presented for 1 sec followed by a 0.33-sec blank. In addition, these localizer runs contained an orthogonal motion manipulation: In half of the blocks, the items were presented statically at fixation. In the remaining half of the blocks, items moved from the center of the screen toward either one of the four quadrants or along the horizontal and vertical meridians at 2.05 deg/sec. Each category was presented in a moving and stationary block.

fMRI data analysis. All fMRI data were processed using Brain Voyager QX software (Brain Innovation, Maastricht, The Netherlands). Preprocessing steps included 3-D motion correction, slice scan-time correction, temporal high-pass filtering (128-Hz cutoff), spatial smoothing (4-mm FWHM kernel), and transformation into Talairach space. Statistical analyses were based on the general linear model (GLM). All GLM analyses included box-car regressors for each stimulus block convolved with a gamma function to approximate the idealized hemodynamic response. For each experimental protocol, separate GLMs were computed for each participant, yielding beta maps for each condition for each participant.

Defining neural sectors. Sectors were defined in each participant using the following procedure. Using the localizer runs, a set of visually active voxels was defined based on the contrast of [faces + bodies + scenes + objects + scrambled objects] versus rest (false discovery rate < 0.05, cluster threshold = 150 contiguous, $1 \times 1 \times 1$ voxels) within a gray matter mask. To divide these visually responsive voxels into sectors, the early visual sector included all active voxels within V1, V2, and V3, which were defined by hand on an inflated surface representation based on the horizontal-versus-vertical contrasts of the meridian mapping experiment. The occipitotemporal and occipitoparietal sectors were then defined as all remaining active voxels (outside early visual cortex), where the division between the dorsal and ventral streams was drawn by hand in each participant based on anatomical landmarks and the spatial profile of active voxels along the surface. Finally, the occipitotemporal sector was divided into ventral and lateral sectors by hand using anatomical landmarks, specifically, the occipitotemporal sulcus, which divides the ventral and lateral surfaces (Figure 3A).

Reliability analysis. To determine the reliability of the response patterns within a given sector, we calculated the group-level split-half reliability of each sector's representational structure. The first step of this process entailed splitting the data into odd and even runs. Then, we created two separate representational similarity matrices by correlating the patterns across all category pairings (e.g., correlation between faces and buildings, cars and chairs) for both the odd and even runs (Kriegeskorte

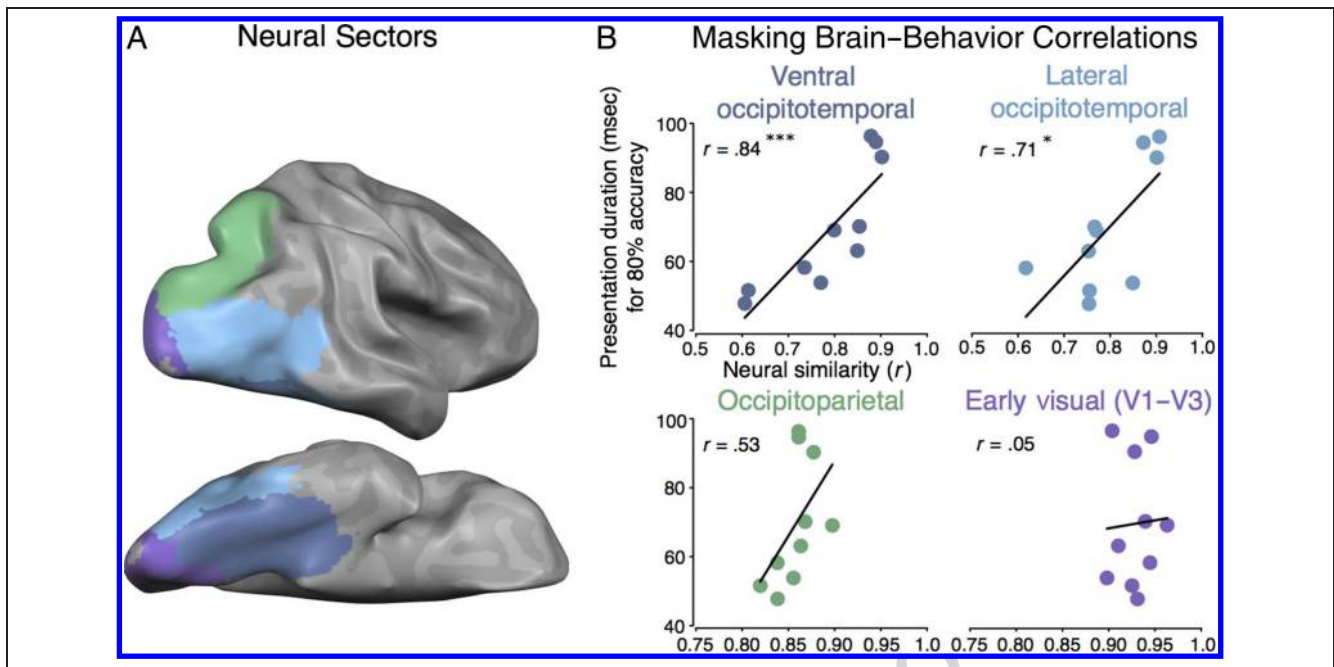


Figure 3. (A) Visualization of the four neural sectors from a representative participant. (B) Group level brain–behavior correlations in all sectors in Experiment 1. Neural similarity between category pairings plotted on the x axis. Estimated presentation durations for equal behavioral performance for all category pairings plotted on the y axis. Each dot corresponds to a single category pairing (e.g., faces and buildings). The values on both axes were calculated by averaging across all behavioral and neuroimaging participants respectively and were plotted accordingly. Note the change in the scales of the x axis between the two occipitotemporal and the occipitoparietal/early visual plots. * $p < .05$, *** $p < .001$.

et al., 2008). We then correlated these representational similarity matrices to determine the similarity in the representational structure between the two halves of the data. These correlation values were transformed using the Spearman–Brown split-half correction formula, which yields an estimate of each region’s full-test reliability (Brown, 1910; Spearman, 1910). This analysis revealed that, with as few as six participants, we were able to obtain highly reliable representational structures in each of the four sectors (ventral occipitotemporal: $r = .99$, lateral occipitotemporal: $r = .98$, occipitoparietal: $r = .82$, early visual cortex: $r = .88$).

Brain–Behavior Correlation Analyses

To determine if masking efficacy is predicted by neural similarity, we performed a representational similarity analysis in which the neural response patterns of all categories were correlated with one another using the Pearson correlation (r) in each sector (Kriegeskorte et al., 2008). This analysis yields a full neural similarity matrix for all category pairings, where the value in each cell of the matrix represents the correlation between a particular category pairing. We then asked which sector’s neural similarity structure best predicted the behavioral similarity structure as measured by the presentation duration needed to obtain 80% accuracy for each category pairing.

To assess the statistical significance of the brain–behavior correlations, we carried out two types of analyses. First, we

performed a permutation analysis on the group level data, which reflects fixed effects of both the behavioral and neural measures (Kriegeskorte et al., 2008). To determine if a given correlation was significant, the condition labels of the data of each individual fMRI participant ($n = 6$) and behavioral participant ($n = 20$) were shuffled and then averaged together to make new, group level structures. These structures were then correlated together, and this procedure was repeated 10,000 times, resulting in a distribution of correlation values. A given brain–behavior correlation was considered significant if it fell within the top 5% of values in this distribution.

In addition, we used linear mixed effects (LME) modeling, which estimates the brain–behavior correlations with random effects of both the behavioral and neural participants (Barr, Levy, Scheepers, & Tily, 2013; Winter, 2013). To do this, the Fisher z -transformed correlation values were modeled as a function of the neural sector, including random effects of behavioral and neuroimaging participants on both the intercept and slope term of the model. The models were implemented using R (R Development Core Team, 2008) and the R packages *lme4* (Bates & Maechler, 2009) and *languageR* (Baayen, 2009). To determine if the correlations observed in the four sectors were statistically significant, we performed likelihood ratio tests, comparing a model with brain region as a fixed effect to another model without it, but which was otherwise identical including the exact random effects structure. For significance testing, p values were estimated using the normal approximation to the

t statistics (Barr et al., 2013) and were considered significant if the *p* values were below the $\alpha = .05$ value.

Both of these analyses were used to assess the overall significance of the brain–behavior correlations. However, only the LME method was used to compare the strength of the correlations between sectors because this method allows us to test for a within-subject effect of brain region (e.g., ventral occipitotemporal vs. early visual cortex) while simultaneously generalizing across both behavioral and neuroimaging participants.

Bootstrap Analysis on Neural Similarity Range

To compare the range of neural similarity across sectors, we used the following procedure: We sampled with replacement the similarity values from each region and calculated the standard deviation of those samples. We then measured the difference in the standard deviation for each of those two samples. This was done 10,000 times to obtain a distribution of difference values. We then asked where the observed difference value fell within that distribution and considered it statistically significant when it fell within the top 5% of values in this distribution.

Results

In the masking experiment, there were significant differences in the estimated presentation durations needed to equate behavioral performance across the different category pairings ($F(1, 9) = 12.89, p < .001, \eta_p^2 = 0.40$; Figure 2B). An extreme example of this difference is that cars and chairs (95 msec per item) had to be presented almost twice as long as faces and buildings (48 msec per item) for equal performance.

What accounts for the variation in the masking efficacy between different category pairings? One possibility is that, despite our efforts to eliminate low-level differences between the categories, enough differences remained such that some category pairings had more similar low-level features than others. This would be consistent with numerous results showing that an effective mask will be similar to the target in terms of low-level features (Breitmeyer & Ogmen, 2006; Phillips & Wilson, 1984; Legge & Foley, 1980). Alternatively, the differences between the category pairings may reflect differing degrees of overlap in higher level processing channels. To distinguish between these possibilities, we correlated the behavioral results with the neural similarity values measured in each of the four sectors. If the differences in masking efficacy were because of low-level similarity, we would expect to find the strongest brain–behavior correlations in early visual areas (V1–V3), which contain neurons that respond to features such as orientation, spatial frequency, and feature combinations (e.g., Freeman, Ziemba, Heeger, Simoncelli, & Movshon, 2013). Alternatively, if overlap among higher level feature combinations determines these masking differences, we should see the strongest

correlations in the occipitotemporal cortex and, potentially, occipitoparietal cortex, where it has been suggested that complex visual shape information is also encoded (Ungerleider & Bell, 2011).

Using a group level permutation analysis, we found strong correlations in ventral ($r = .84, p < .001$) and lateral ($r = .71, p < .05$) occipitotemporal cortex and a marginally significant correlation in occipitoparietal cortex ($r = .53, p = .06$), consistent with the notion that the differences in the masking task were because of similarity in high-level neural response patterns (Figure 3B). Meanwhile, there was no correlation in early visual cortex ($r = .05, p = .44$), suggesting that lower level similarity between these images did not drive the masking differences.

A convergent pattern of results was obtained with the LME analysis. For this analysis, the masking results from every individual behavioral participant ($n = 20$) were correlated with the neural similarity values obtained in every neuroimaging participant ($n = 6$). This resulted in 120 participant-by-participant brain–behavior correlation values that were then Fisher *z* transformed. We then modeled these transformed correlation values as a function of the neural sector, including random effects of behavioral and neuroimaging participants on both the intercept and slope term of the model (Barr et al., 2013).

The results of this analysis found significant brain–behavior correlations in ventral occipitotemporal (parameter estimate = 0.74, $t = 9.76, p < .001$), lateral occipitotemporal (parameter estimate = 0.49, $t = 6.44, p < .001$), and occipitoparietal (parameter estimate = 0.32, $t = 2.54, p < .05$) cortices but not in early visual cortex (parameter estimate = 0.08, $t = 0.83, p = .40$). In addition, we also used the LME analysis to compare the strength of the correlations in the different sectors. To do this, we measured the differences in the slope term of the LME model for two given sectors to see if the slopes between two sectors were significantly different. This analysis revealed that the correlations in ventral occipitotemporal cortex were greater than those in the three other sectors (slope estimates $< -0.25, t < -2.30, p < .05$ in all cases), the correlations in lateral occipitotemporal cortex were significantly different from those in early visual cortex (slope estimates = $-.41, t = -3.00, p < .01$), and the correlations in occipitoparietal cortex were trending but not significantly different from early visual cortex (slope estimates = $-.24, t = -1.65, p = .09$). Meanwhile, the correlations in lateral occipitotemporal and occipitoparietal cortices were not significantly different (slope estimates = $-0.16, t < -1.41, p = .16$).

It is possible that there is a correlation between neural similarity in early visual cortex and the behavioral data but that the restricted range of neural similarity prevented us from detecting this relationship. By design, all of the categories were similar to one another in early visual cortex (range of *r* values is .88–.95, $SD = 0.21$). However, the range of neural similarity in occipitoparietal cortex was comparable (range of *r* values is .82–.90, $SD = 0.22$; bootstrap analysis: $p = .49$), yet this region showed

a significant brain–behavior correlation. This suggests that the lack of a correlation in early visual cortex is not solely because of the relative lack of variance in the similarity values.

Discussion

Using carefully controlled stimuli, we found evidence supporting the hypothesis that the organization of higher level visual cortex limits the amount of information that can be accessed by awareness. Although overlap among low-level features and early visual areas constrains visual processing in certain instances (Breitmeyer & Ogmen, 2006; Phillips & Wilson, 1984; Legge & Foley, 1980), the stimuli used in this experiment were designed to be difficult to categorize based on lower level properties, and thus, this lower level bottleneck was not evident here: The correlations between early visual areas and behavior were not significant and were reliably lower than the correlations found in occipitotemporal cortex. Instead, these results are consistent with the idea that processing these stimuli in this task was limited by competition within higher level channels.

EXPERIMENT 2: CONTINUOUS FLASH SUPPRESSION

Does the relationship between higher level neural structures and visual awareness hold with paradigms besides visual masking? To answer this question, we used a novel continuous flash suppression paradigm to measure how long it takes target categories to break through suppression from distracting categories (e.g., a chair breaking through suppression from buildings). Although both masking and continuous flash suppression render stimuli invisible, there are many differences between the two paradigms in this particular study. For example, here, stimuli are presented under binocular viewing conditions in the masking experiment but are presented under dichoptic viewing conditions in the continuous flash suppression experiment (but see previous masking studies in which targets and masks were viewed dichoptically; van Boxel, van Ee, & Erkelens, 2007; Schiller, 1965). In addition, it has been repeatedly claimed that different types of unconscious processing are possible under these two techniques. These differences have been observed both behaviorally (Faivre, Berthet, & Kouider, 2014; Almeida, Mahon, Nakayama, & Caramazza, 2008, 2013; Van den Bussche, Van den Noortgate, & Reynvoet, 2009; Kouider & Dehaene, 2007) and with neuroimaging data (Kang, Blake, & Woodman, 2011; Fang & He, 2005; Dehaene et al., 2001). Thus, it is possible that the brain–behavior correlation found with masking might not be observed when using continuous flash suppression. In fact, previous results suggest a clear dissociation between dorsal and ventral processing under continuous flash suppression, with suppressed items being processed

more by the dorsal stream. Evidence for this idea has come from both behavioral and neuroimaging experiments (Almeida et al., 2010; Fang & He, 2005). On the basis of these particular results, one might expect there to be a significant correlation between behavioral performance within occipitoparietal cortex but not within occipitotemporal cortex. However, it has also been suggested that these results only hold for tools and other elongated, graspable objects and suppressed items are also processed by the ventral stream (Ludwig & Hesselmann, 2015; Hesselmann & Malach, 2011). Given these dissimilar results, it is not clear if a converging pattern of results will be obtained for the two experiments.

Methods

Participants

Twenty new participants who did not participate in Experiment 1 performed the behavioral experiment. This number was chosen to match the number of participants used in Experiment 1. The six previous neuroimaging participants and their data were used again for Experiment 2.

Continuous Flash Suppression Task

In this task, participants viewed images through a stereoscope, which presented different images to each eye (Figure 4A). The participant’s task was to monitor for the appearance of a small target item (e.g., a chair) that appeared either above or below the central fixation point. The target item was only presented to one eye. At the same time, large distracting masks (e.g., buildings) were presented to the other eye, with a new image being shown every ~117 msec. This continuously flashing stream of images in one eye tends to suppress perception of images in the other eye (hence, “continuous flash suppression”). As an observer, you cannot tell which eye the images are seen through, and thus, the experience is that you can only see the continuously flashing images. However, after some period, the suppressed image will break through suppression and be perceived. Here, we asked whether the time it takes the target to break through depends on the categories of target and distractor presented (e.g., will it take longer for chairs to break through with house masks than with face masks?).

For each participant, there were five experimental blocks, with each block corresponding to a particular masking category (e.g., “In this block, buildings will be the masks.”). Targets from the other four categories were randomly intermixed within each block. The order in which the five blocks were presented was counterbalanced across observers using a Latin square design. Each block had 10 practice trials and 80 experimental trials, except for the first block, which had 20 practice trials to better familiarize participants with the task. Within each

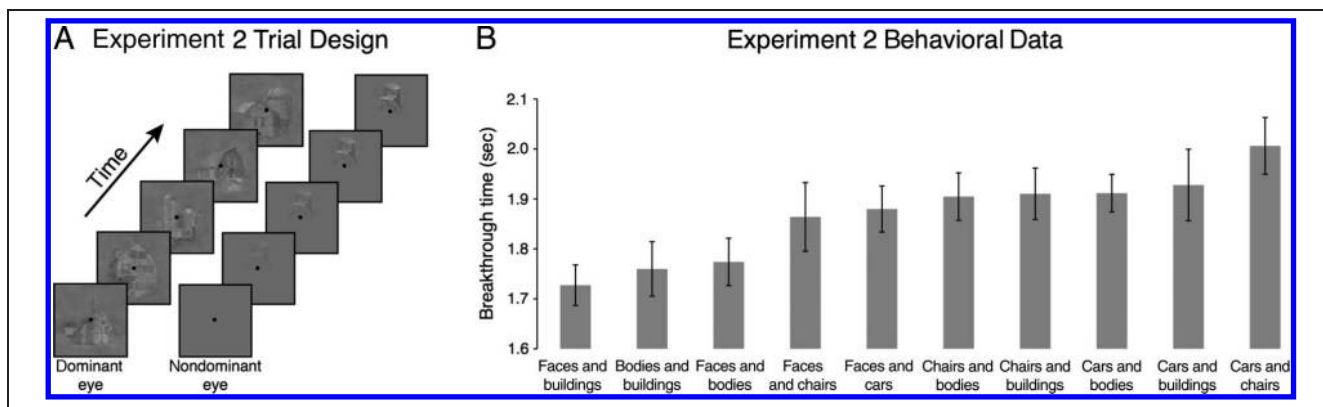


Figure 4. (A) Representation of part of a single trial from Experiment 2. In this case, a target chair is presented above fixation to the nondominant eye, whereas buildings are the masks presented to the dominant eye. Note that the size of the target (i.e., the chair) is not drawn to scale in this figure. Here, the target is bigger and higher contrast than it was during the experiment and is presented this way in the figure for display purposes (see the Methods sections for information on the exact size of the target). (B) Behavioral results from Experiment 2. Average breakthrough time is shown on the *y* axis, reflecting the average amount time needed for participants to stop the trial and then accurately report whether the target was above or below the fixation mark. Each bar corresponds to a particular category pairing of targets and masks. Error bars reflect within-subject *SEM*.

block, there were 20 trials with each of the four remaining categories as the target (i.e., if the building was the mask, there were 20 trials with faces, chairs, cars, and bodies as the targets). Of those 20 trials, 10 had the target appear above the fixation mark, and 10 were below the fixation mark.

On each trial, a black fixation dot ($\sim 0.25^\circ$) appeared at the center of the screen for 250 msec and then turned red for 200 msec to alert the participant that the trial was about to begin. Once the trial began, the dot turned black and remained there for the duration of the trial. Masks were presented to the dominant eye of each participant, subtended $\sim 16.5^\circ$, and changed at a frequency of ~ 8.5 Hz (~ 117 msec per image). The target was presented to the nondominant eye within a circular Gaussian aperture so that it naturally faded into the background and subtended $\sim 3^\circ$. Both target and mask images were presented simultaneously on the monitor and fused with a mirror stereoscope (Figure 4A). A thin black border of $\sim 0.3^\circ$ in width that was centered on the fixation dot was also presented to both eyes (note the black borders in Figure 4A).

Over the course of the trial, targets gradually became more visible, and the masks gradually became less visible. To accomplish this, the opacity of the target item was ramped up from 0% to 100% over the course of approximately 2100 msec. After the target became totally opaque, the masks gradually decreased in opacity from 100% to 40% over approximately the next 6300 msec. After this (~ 8400 msec after the trial began), the trial ended, even if the participant gave no response.

Participants' task was to determine whether the target item was presented slightly above or below a central fixation dot. Participants were told to press the space bar as soon as they could correctly localize the target even if they were unaware of the target's identity. We measured how long it took participants to stop the trial as a function of the mask and target categories. After stopping the

trial, participants used the arrow keys to indicate the location of the target. Visual feedback was given immediately. The dependent variable in this experiment was the amount of time it took participants to press the space bar to stop the trial. Trials in which participants responded incorrectly or did not stop the trial within 8064 msec were excluded from analysis. In addition, all trials in which participants responded in less than 300 msec or slower than 3 *SDs* from their own overall mean response time were excluded from further analysis. These trimming procedures led to the removal of 1.2% of the trials.

Stimuli were the same as in Experiment 1. All items were presented on a 24-in. Apple iMac computer, with a 60-Hz refresh rate and a screen resolution of 1920×1200 pixels, and were created and controlled with MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Participants sat approximately 57 cm away from the display such that 1 cm on the screen subtended 1° of visual angle. Eye dominance was determined before the experiment using the Miles test (Miles, 1930).

Neuroimaging and Brain–Behavior Correlation Analyses

The neuroimaging data from Experiment 1 were used in Experiment 2, and the same analysis procedures were followed, where the behavioral measure here was taken as the average amount of time it takes for target items to break through suppression for all possible category pairings.

Results

In the flash suppression experiment, there were significant differences in the estimated presentation durations needed to equate behavioral performance across the different

category pairings ($F(1, 9) = 2.31, p < .05, \eta_p^2 = 0.11$; Figure 4B). Comparing the behavioral data from the two experiments revealed a significant correlation, with significance determined with a permutation test ($r = .70, p < .001$; Kriegeskorte et al., 2008). Next, we asked if a similar relationship between neural structure and behavioral performance would be observed in the flash suppression experiment as it was in the masking experiment.

Using a group level permutation analysis, we once again found significant correlations in ventral ($r = .64, p < .05$) and lateral ($r = .73, p < .01$) occipitotemporal cortex (Figure 5). However, in this case, no significant correlations were found in occipitoparietal cortex ($r = .44, p = .11$) or early visual cortex ($r = -.39, p = .87$). A largely similar pattern of results was obtained with the LME analysis. Once again, this analysis found significant correlations in ventral (parameter estimate = 0.29, $t = 3.62, p < .001$) and lateral (parameter estimate = 0.26, $t = 2.59, p < .01$) occipitotemporal cortex. However, unlike what was found with the permutation analysis, a significant correlation was found in occipitoparietal cortex (parameter estimate = 0.14, $t = 2.35, p < .05$). This significant correlation is likely because of the LME analysis having more statistical power than a permutation test because it takes into consideration the correlations between all behavioral and neural participants. Again, no significant correlation was found in early visual cortex (parameter estimate = 0.12, $t = 1.16, p = .24$). When using the LME analysis to compare the strength of the correlations in different sectors, we found no difference between ventral and lateral occipitotemporal cortex

(slope estimate = $-0.24, t = -0.02, p = .78$). However, correlations from ventral occipitotemporal cortex were significantly greater than those in both occipitoparietal and early visual cortices (slope estimates $< -0.14, t < -2.14, p < .05$ in both cases). Meanwhile, correlations in lateral occipitotemporal cortex were not significantly different from those in occipitoparietal cortex (slope estimate = $-0.12, t = -1.18, p = .24$) but were significantly greater than those in early visual cortex (slope estimate = $-0.12, t = -1.18, p = .24$). Finally, correlations in occipitoparietal cortex were significantly greater than those in early visual cortex (slope estimate = $-0.27, t = -2.13, p < .05$).

Discussion

Here, we developed a novel continuous flash suppression paradigm to measure the extent to which stimuli from different categories prevent each other from being accessed by consciousness. Rather than varying the target items and holding the masks constant, we systematically varied both the targets and the masks, forcing different categories to directly compete with one another. In this case, we found that the speed with which certain categories break suppression from other categories was predicted by the similarity of the neural patterns elicited by those categories in higher level visual cortex. Thus, we once again find evidence suggesting that visual awareness is limited by the representational architecture of these neural regions. The fact that a similar brain-behavior relationship was found in this paradigm suggests that this

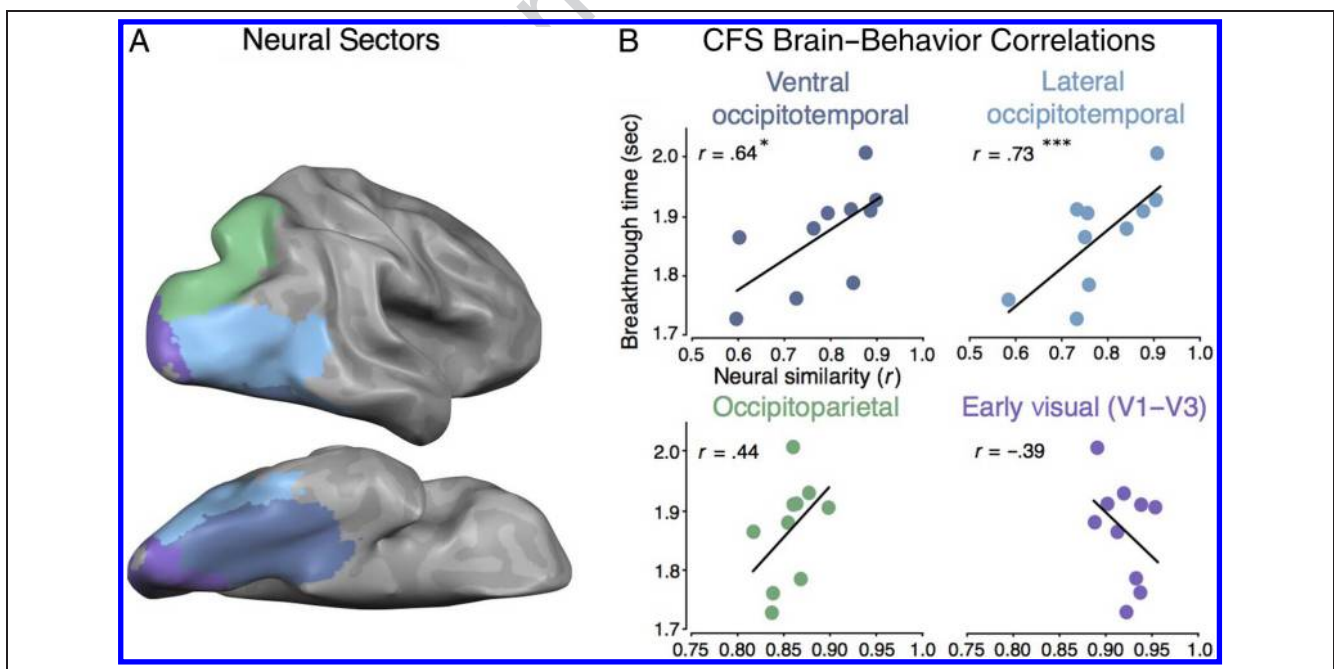


Figure 5. Group level brain-behavior correlations in all sectors in Experiment 2. Neural similarity between category pairings is plotted on the x axis. The y axis represents the amount of time it takes for targets to break through the masks for each category pairing. The values on both axes were calculated by averaging across all behavioral and neuroimaging participants respectively and were plotted accordingly. Note the change in the scales of the x axis between the two occipitotemporal and the occipitoparietal/early visual plots. * $p < .05$, *** $p < .001$.

neural bottleneck is a more general constraint on conscious processing and is not because of the idiosyncrasies of one particular paradigm.

One advantage of the current experiment is that it requires competition between stimuli that vary dramatically in visual scale (the masks are approximately $5.5\times$ the size of the targets). This aspect of the experimental design likely minimizes the contribution of competition in early visual cortex, which is organized into retinotopic maps where stimuli could only interact and compete with one another if they are presented at the same location and spatial scale (Wandell & Winawer, 2015). Whereas the receptive fields in early visual cortex are not large enough to generalize over these differences in size, receptive fields in higher level visual cortex are large enough to be invariant to such transformations (Op de Beeck & Vogels, 2000). Thus, these results reinforce the conclusion that higher level regions impose a bottleneck on visual awareness that is unlikely to simply reflect competition within early visual cortex.

GENERAL DISCUSSION

Here, we found differences in how effectively different visual categories prevent one another from being consciously perceived using both masking and continuous flash suppression. To understand these differences, we asked if performance on these behavioral tasks could be predicted by the similarity of the underlying neural patterns associated with these categories. Using stimuli that could not be easily distinguished by low-level features, we found significant brain-behavior correlations between behavioral performance and neural similarity in both ventral and lateral occipitotemporal cortex, weaker correlations in occipitoparietal cortex, and no reliable correlations in early visual cortex. These results suggest that the representational structure of higher level processing channels, particularly in the ventral stream, plays an important role in determining what information can be accessed by awareness.

These results add to previous work regarding the limits of conscious processing that have focused on perceptual (Kanai et al., 2010) or attentional (Cohen et al., 2011) limitations. These limitations have been largely associated with competitive interactions between stimuli in early visual cortex (Yuval-Greenberg & Heeger, 2013; Anderson et al., 2012; Tong et al., 2006) and the finite processing capacity of the frontoparietal network (Dehaene & Changeux, 2011; Lamme, 2010; Tononi & Koch, 2008). Although we do not dispute the role that these particular regions play, we suggest that the representational architecture within higher level perceptual processing regions imposes a unique bottleneck on conscious processing and needs to be incorporated into existing large-scale models of conscious access.

How might this bottleneck relate to the limitations stemming from early visual cortex and the frontoparietal

network? In terms of early visual cortex, we suggest that the same competitive processes that occur earlier in the visual hierarchy may also occur within higher level visual cortex along both the ventral and dorsal pathways. Under this view, there are multiple bottlenecks limiting visual awareness, with the relevant bottleneck being determined by whichever representational space best separates the items. For example, here, the categories were more distinguishable in a higher level space than a lower level space. Thus, the best predictor of the behavioral results was similarity among higher level neural channels. If the categories differed in terms of their lower level properties, we predict that the behavioral results would be better explained by similarity in early visual cortex. We speculate that access to awareness is limited by a series of bottlenecks, each operating on distinct representational levels that can have dissociable effects on awareness depending on the stimuli/task.

How might these representational constraints relate to mechanisms of attention and frontoparietal processing? We suggest that representational similarity among neural channels is a processing bottleneck that precedes attention. This notion is somewhat similar to Kinsbourne and colleagues' "functional distance" theory in which it was speculated that multiple pieces of information are more efficiently processed when they rely on brain areas that are "functionally distant" (Kinsbourne, 1981; Kinsbourne & Hicks, 1978). When stimuli interact and compete for representational resources within these neural channels in a mutually suppressive fashion, attention is the mechanism that resolves this competition and allows certain stimuli to be processed further (Desimone & Duncan, 1995). Under this view, the amount of overlap between neural channels determines the difficulty of resolving the competition between stimuli. As the amount of overlap between neural channels decreases, it will be easier for attention to resolve the competition between stimuli, which will behaviorally result in improved behavioral performance (e.g., faster RTs). An appealing aspect of this framework is that it could help directly link cognitive theories that posit the existence of separate pools of attentional resources (Alvarez, Horowitz, Arsenio, DiMase, & Wolfe, 2005; Awh et al., 2004) with the organization of the visual system. In this case, independence in cognitive processing may reflect the amount of attentional resources needed to resolve the competition in these neural processing channels.

If this conception of the relationship between neural structure and attentional resources is correct, it could alter the interpretation of previous studies of visual awareness that have focused exclusively on a more general attentional limit. For example, it has been claimed that perceiving animals/vehicles requires little or no attention (Li, VanRullen, Koch, & Perona, 2002; but see Cohen et al., 2011). This claim comes from the fact that detecting animals and vehicles in the periphery is unaffected by simultaneously searching for a "T" among "Ls." However,

it is possible that there is no interference between the tasks simply because these stimulus categories (i.e., animals/vehicles and “Ts/Ls”) are likely supported by different neural channels (Dehaene & Cohen, 2011; Peelen & Kastner, 2011). If, instead, the T-among-Ls task was replaced by a task that also involved animals and vehicles, there may be interference between the two tasks because both tasks would activate the same neural channels. Indeed, various results that are thought to suggest that certain stimuli can be perceived without attention may instead be the product of the representational architecture of the visual system (Reddy, Wilken, & Koch, 2004; Mack & Rock, 1998).

Unlike previous studies, participants in this experiment did not perform the masking or flash suppression task while in the scanner. Why does the similarity in neural response profiles for fully visible stimuli (the neural data) predict the time needed to perceive a masked stimulus (the behavioral data)? It is often claimed that visual information can only be consciously accessed if there is recurrent processing between both the neural channels that encode task-relevant information and the frontoparietal network (Dehaene & Changeux, 2011; Lamme, 2010). Under this view, the masks in both behavioral experiments prevent recurrent processing of the target. When there is relatively high similarity among the channels activated by two categories, it is harder for recurrent processing associated with the target to be maintained because the masks will disrupt interactions between the relevant cortical areas. Thus, recurrent processing will not occur unless the presentation time between items is sufficiently long in the masking experiment (Experiment 1) or until the competition between the two monocularly presented stimuli has been resolved in the target’s favor (Experiment 2).

Overall, these results suggest that a limitation of visual awareness is the organization of the higher level visual system. It is possible that the relationship between representational architecture and visual awareness extends beyond the visual modality and that sensory awareness in all modalities (e.g., audition) may be limited by the organization of the relevant neural pathways. Anatomical constraints on conscious awareness may be a ubiquitous phenomenon, placing limits on conscious processing in all domains of cognition.

UNCITED REFERENCES

Cohen, Cavanagh, Chun, & Nakayama, 2012
Nunnally & Bernstein, 1994

Acknowledgments

Thanks to Daniel Dennett and Nancy Kanwisher for helpful discussions. This work was supported by an NSF-GRFP and NIH-NRSA (F32EY024483, M. A. C.), NIH NEI RO1 (EY01362, K. N.),

NIH-NRSA (F32EY022863, T. K.), and NSF CAREER (BCS-0953730, G. A. A.).

Reprint requests should be sent to Michael A. Cohen, Department of Brain and Cognitive Sciences, MIT, Room 46-4141, 77 Massachusetts Avenue, Cambridge, MA 02139, or via e-mail: michaelthecohen@gmail.com.

REFERENCES

- Almeida, J., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences, U.S.A.*, *105*, 15214–15218.
- Alvarez, G. A., Horowitz, T. S., Arsenio, H. C., DiMase, J. S., & Wolfe, J. M. (2005). Do multielement visual tracking and visual search draw continuously on the same visual attention resources? *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 643–647.
- Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Current Biology*, *22*, 1199–1206.
- Awh, E., Serences, J., Laurey, P., Dhaliwal, H., van der Jagt, T., & Dassonville, P. (2004). Unimpaired face discrimination during the attentional blink: Evidence for multiple processing channels. *Cognitive Psychology*, *48*, 95–126.
- Baayen, R. H. (2009). languageR: Data sets and function with “Analyzing linguistic data: A practical introduction to statistics.” R package version 0.955.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.
- Bates, D. M., & Maechler, M. (2009). lme4: Linear mixed-effects models using Eigen and R syntax. R package version 0.999375-32.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Breitmeyer, B., & Ogmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*. Oxford, UK: Oxford University Press.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, *3*, 296–322.
- Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science*, *22*, 1165–1172.
- Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, *16*, 411–417.
- Cohen, M. A., Konkle, T., Nakayama, K., & Alvarez, G. A. (2014). Exploring the representational geometry of object representation in the ventral stream using brain–behavior correlations. *Journal of Vision*, *14*, 185.
- Cohen, M. A., Konkle, T., Rhee, J., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences, U.S.A.*, *111*, 895–896.
- Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*, 200–227.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*, 254–262.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752–758.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

- Faivre, N., Berthet, V., & Kouider, S. (2014). Sustained invisibility through crowding and continuous flash suppression: A comparative review. *Frontiers in Psychology*, *5*, 475.
- Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *10*, 1380–1385.
- Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (2013). A functional and perceptual signature of the second visual area in primates. *Nature Neuroscience*, *16*, 974–981.
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in visual categorization. *Nature Reviews Neuroscience*, *15*, 536–548.
- Hesselmann, G., & Malach, R. (2011). The link between fMRI-BOLD activation and perceptual awareness is “stream-invariant” in the human visual system. *Cerebral Cortex*, *21*, 2829–2837.
- Kanai, R., Walsh, V., & Tseng, C. H. (2010). Subjective discriminability of invisibility: A framework for distinguishing perceptual and attentional failures of awareness. *Consciousness and Cognition*, *19*, 1045–1057.
- Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *Journal of Neuroscience*, *31*, 13535–13545.
- Kinsbourne, M. (1981). Single channel theory. In D. Holding (Ed.), *Human skills* (pp. 65–89). Chichester, UK: Wiley.
- Kinsbourne, M., & Hicks, R. E. (1978). Functional cerebral space: A model for overflow, transfer, and interference effects in human performance. In J. Requin (Ed.), *Attention and performance VIII* (pp. 345–362). Hillsdale, NJ: Erlbaum.
- Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience*, *33*, 10235–10242.
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipito-temporal cortex. *Neuron*, *74*, 1114–1124.
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 1–28.
- Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*, 204–220.
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, *70*, 1458–1471.
- Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 9596–9601.
- Lin, Z., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87*, 195–211.
- Ludwig, K., & Hesselmann, G. (2015). Weighing the evidence for a dorsal processing bias under continuous flash suppression. *Consciousness and Cognition*, *35*, 251–259.
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge, MA: MIT Press.
- Miles, W. R. (1930). Ocular dominance in human adults. *Journal of General Psychology*, *3*, 412–430.
- Nunnally, J. C., & Bernstein, J. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Op de Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *Journal of Comparative Neurology*, *426*, 505–518.
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *108*, 12125–12130.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Phillips, C. G., & Wilson, H. R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America A*, *1*, 226–232.
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org>.
- Reddy, L., Wilken, P., & Koch, C. (2004). Face gender discrimination in the near-absence of attention. *Journal of Vision*, *4*, 106–117.
- Schiller, P. H. (1965). Monoptic and dichoptic visual masking by patterns and flashes. *Journal of Experimental Psychology*, *69*, 193–199.
- Schiller, P. H. (1966). Forward and backward masking as a function of relatively overlap and intensity of test and masking stimuli. *Perception & Psychophysics*, *1*, 161–164.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Dwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.
- Spearman, C. C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, *3*, 271–295.
- Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Sciences*, *1124*, 239–261.
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*, 1096–1101.
- Ungerleider, L. G., & Bell, A. H. (2011). Uncovering the visual “alphabet”: Advances in our understanding of object perception. *Vision Research*, *51*, 782–799.
- van Boxel, J. J. A., van Ee, R., & Erkelens, C. J. (2007). Dichoptic masking and binocular rivalry share common perceptual dynamics. *Journal of Vision*, *7*, 1–11.
- Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin*, *135*, 452–477.
- VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research*, *18*, 3017–3027.
- Wandell, B. A. (1999). Computational neuroimaging of human visual cortex. *Annual Review of Neuroscience*, *22*, 145–173.
- Wandell, B. A., & Winawer, J. (2015). Computational neuroimaging and population receptive fields. *Trends in Cognitive Sciences*, *19*, 349–357.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*, 671–684.
- Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv: 1305.5499.
- Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *23*, 8327–8328.
- Yuval-Greenberg, S., & Heeger, D. J. (2013). Continuous flash suppression modulates cortical activity in early visual cortex. *Journal of Neuroscience*, *33*, 9635–9643.