

Learning to Sample: Eye Tracking and fMRI Indices of Changes in Object Perception

Lauren L. Emberson¹ and Dima Amso²

Abstract

■ We used an fMRI/eye-tracking approach to examine the mechanisms involved in learning to segment a novel, occluded object in a scene. Previous research has suggested a role for effective visual sampling and prior experience in the development of mature object perception. However, it remains unclear how the naive system integrates across variable sampled experiences to induce perceptual change. We generated a Target Scene in which a novel occluded Target Object could be perceived as either “disconnected” or “complete.” We presented one group of participants with this scene in alternating sequence with variable visual experience: three Paired Scenes consisting of the same Target Object in variable rotations and states of occlusion. A second control group was presented with similar Paired Scenes that did

not incorporate the Target Object. We found that, relative to the Control condition, participants in the Training condition were significantly more likely to change their percept from “disconnected” to “connected,” as indexed by pretraining and post-training test performance. In addition, gaze patterns during Target Scene inspection differed as a function of variable object exposure. We found increased looking to the Target Object in the Training compared with the Control condition. This pattern was not restricted to participants who changed their initial “disconnected” object percept. Neuroimaging data suggest an involvement of the hippocampus and BG, as well as visual cortical and fronto-parietal regions, in using ongoing regular experience to enable changes in amodal completion. ■

INTRODUCTION

The mechanisms of object perception and recognition have received considerable scientific attention. Discovery in this domain holds promise for informing one of the most important questions in cognitive and brain sciences: How do we construct and act on an enduring representation of the external environment? There are a variety of research avenues and levels of analysis, appropriate for addressing this question, that span everything from the size of receptive fields along visual pathways to philosophical discussions about the origins of object concepts. Here we use a novel amodal completion paradigm to examine how visual experience supports changes in object perception. Amodal or perceptual completion is the perception of an occluded object as whole, despite incomplete visual input. Research suggests a role for both effective attentional sampling and regular experience with objects in the development of amodal completion. However, the mechanistic nature of the interaction of these variables is unclear. The current work uses an fMRI/eye-tracking approach in adults to examine this issue.

Data from both the perceptual development and cognitive neuroscience literatures converge to shed light on the mechanisms that support amodal completion. Neonates respond to partly occluded object displays only in terms of what is directly visible. They do not perceptually

complete a center-occluded object (Slater, Johnson, Brown, & Badenock, 1996; Slater et al., 1990; but see Valenza, Leo, Gava, & Simion, 2006). By 2 months, infants are able to perform amodal completion under limited conditions (e.g., where the occluder is narrow). By 4 months, amodal completion is more robust (Johnson, 2004). Neuroimaging investigations that have examined the perception of occluded objects have found involvement of object processing regions, such as the lateral occipital complex (LOC), and portions of the inferior temporal and posterior parietal cortices (Hedg , Feng, Murray, & Kersten, 2008; Shuwairi, Curtis, & Johnson, 2007; Olson, Gatenby, Leung, Skudlarski, & Gore, 2004; Lerner, Hendler, & Malach, 2002; Grill-Spector, Kourtzi, & Kanwisher, 2001). Shuwairi et al. (2007) note that maintaining an active representation of an occluded object may require a series of mechanisms including selective attention (Scholl, 2001; Awh, Jonides, & Reuter-Lorenz, 1998) and visual working memory (Pasternak & Greenlee, 2005).

A series of developmental behavioral studies have determined a role for developing attention-driven sampling mechanisms in the emergence of amodal completion (Bhatt & Quinn, 2011; Amso & Johnson, 2006). We use sampling here to mean visually orienting to object-relevant locations in a visual scene. Theoretically, sampling may serve to support the extraction of object feature correlations in the service of efficient perception and recognition (Bhatt & Quinn, 2011). For example, Amso and Johnson (2006) found that 3-month-old infants who indicate unity

¹University of Rochester, ²Brown University

perception of an occluded object also targeted scans and fixations to the object surfaces during initial visual inspection, indicating a relationship between where infants look and what they perceive. A question key to the development of object perception and recognition remains open. How does the organism learn what to sample when encountering novel objects in variable and cluttered scenes?

Clues are offered by research showing that visual experience plays an essential role in the development of object perception (Kersten, Mamassian, & Yuille, 2004; Gibson, 1969). Infants as young as 4.5 months can use variable visual experience with an object to segment a novel scene (Needham, Dueker, & Lockhead, 2005). Similarly, the adult visual system can develop sophisticated visual processing of a novel class of objects with a short training period (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999).

It is important to consider how similar two statistically regular object experiences must be to support perceptual change. Invariance is an important property of perceptual systems, whereby objects are recognized as the same independent of changes in perceptual information such as scale, location, and orientation. This phenomenon involves the entire ventral visual pathway (V1, V2, V4, and inferotemporal cortex or IT). Along the pathway, IT neuronal firing has been found to most highly correlate with conscious perception and to show various degrees of invariance to image transformations (Leopold & Logothetis, 1999; Booth & Rolls, 1998). Inferotemporal cortex is highly connected to medial-temporal lobe (MTL) regions, including the hippocampus and parahippocampal cortex. For example, research has suggested that the MTL has a role in the encoding of object categories (Gross, 2000) and identity across variable input (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). These representations may be a by-product of relations among variable experiences. Using the principle of trace learning, a version of Hebbian learning, modeling work has shown that different views of an object that occur close together in time are associated or bound (Stringer, Rolls, & Tromans, 2007). Trace learning generated invariant neuronal responses to different transforms or variations of an object.

We used a training paradigm combined with functional neuroimaging (fMRI)/eye-tracking methods to provide insight into the role of visual experience and sampling in the perception of a novel occluded object. In an effort to mimic the naive visual experience, we tested adults on whether variable exposure to a novel occluded object would support change from perceiving the object as two “disconnected” surfaces to perceiving it as one occluded whole or as “connected” (Figure 1). Only participants with an initial “disconnected” percept completed the experiments. In the Training condition, we generated alternative views of the same Target Object in variable rotations and occluded in different locations, embedded in other simple visual scenes. In the Control condition, the same Target Scene is paired with equally complex images without the Target Object.

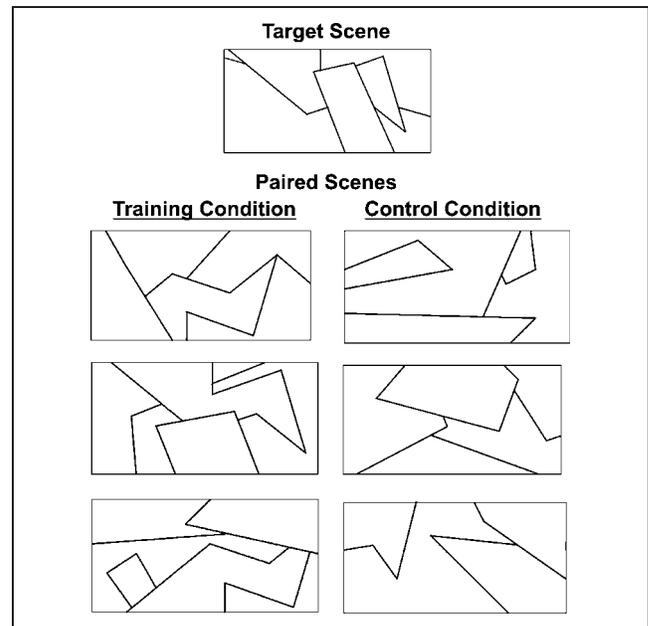


Figure 1. Illustrates stimuli employed in this study. The Target Scene (top) was presented in alternation with Paired Scenes (bottom). The only difference between conditions is the identity of the Paired Scenes. In the Training condition (left) the scenes contained the Target Object in varying orientations and states of occlusion. In the Control condition (right), the Paired Scenes did not contain the Target Object.

We predict that efficient attentional sampling in the Training condition will support change in amodal completion of the Target Object. Our specific prediction is derived from the reviewed developmental literature (Amso & Johnson, 2006; Johnson, Slemmer, & Amso, 2004). Those data demonstrated that infants who perceived a novel occluded object as connected had a larger proportion of scans and fixations to the surfaces of the occluded object, relative to infants who indicated a “disconnected” percept. Likewise, we predict here that that visually attending to Object Surfaces (sections of the visual scene pertaining to the Target Object) in the novel scenes will be related to change in amodal completion. Critically, we predict that this object-targeted visual sampling will be driven by effective use of variable exposure to the Target Object in various orientations and states of occlusion in the Training condition. On the basis of the reviewed neuroimaging and computational literatures, we further predict involvement of ventral visual and MTL pathways in changes in amodal completion in the Training condition.

GENERAL MATERIALS AND METHODS

The following sections pertain to participants tested both behaviorally and with fMRI.

Participants

Sixty-one volunteers (37 women, M age = 21.39 years, SD = 2.8 years) completed this study. Twenty-one of these

volunteers were neuroimaging participants (14 women, M age = 21.5 years, SD = 2.6 years, all right-handed). All participants had no history of neurological or psychiatric disorders, were born full-term with no major birth complications, and had normal or corrected-to-normal vision. Participants were recruited using flyers and the departmental undergraduate subject pool and were compensated for their time and travel with course credit or money. Written informed consent, in accord with the policies of the institutional review board, was obtained from all volunteers before participation in the study.

Stimuli and Design

Stimuli were two-dimensional, black-and-white line drawings depicting visual scenes. Throughout the article, we will refer to “Target Object”, “Surfaces”, and “Object Surfaces.” Target Object refers to the manipulated object that is presented in both the Target Scene and the Paired Scenes of the Training condition. The term surface refers to closed, clearly defined regions of the visual scenes where each scene has five to seven surfaces. Each surface could be perceived as a single object or multiple surfaces can perceptually be grouped into a single object. For example, in the Target Scene, the Target Object is occluded such that it occupies two spatially distinct object surfaces. Object Surfaces refers specifically to the two spatially distinct surfaces of the Target Scene that comprise the Target Object. Since the Target Object is comprised of two separate surfaces it is unclear whether the Target Object is two separate objects (a “disconnected” percept) or a single object behind another (a “connected” percept; Figure 2A, B).

The experiment examined two exposure Conditions (Training and Control conditions). Each condition provided equal experience with the same Target Scene but were differentiated by the Paired Scenes. In the Training condition, Paired Scenes included additional varying views of

the Target Object. In the Control condition (tested only behaviorally), the Paired Scenes had the same number of surfaces and approximately the same number of perceived objects are determined through extensive pilot testings (Figure 1). We will refer to the difference between Target and Paired Scenes as an effect of “Scene Type.”

Pretraining and Posttraining Tests

For the pretraining test, all participants colored all seven scenes in Figure 1. Only participants who colored the Target Object as “disconnected” in the pretraining test were enrolled into the experiment (Figure 2A). For the posttraining test, all participants colored the Target Scene and only the Paired Scenes for their exposure condition (4).

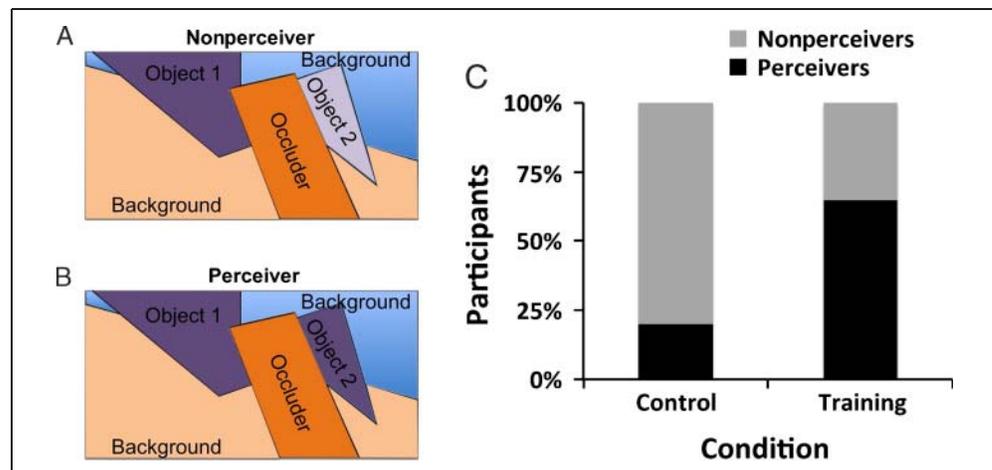
Before participating in the pretraining test coloring task, the experimenter colored an example scene, which was not subsequently used in the study, and explained that the purpose of the study was to understand the participant’s visual perception. Participants were told that the scenes were intentionally designed to be unfamiliar and abstract and to color surfaces that comprised the same object the same color. An additional 25 participants were found to have an initial “completed” percept of the Target Object and were subsequently excluded.

We gathered two additional measures of perceptual change over the course of the task, in addition to the eye movements and pre- and posttraining coloring task. Before and after each exposure block, participants were asked to report the number of objects in each scene, without time limit, as an indirect index of whether they perceive the Target Object as “complete” or “disconnected.” Participants indicated the number of objects using the keyboard and to pressed the space bar to proceed to the next scene.

After the final training block, participants viewed all the visual scenes for their condition with numbers on each surface and were asked to use the numbers to describe

Figure 2. Sample pretraining and posttraining tests and behavioral results. Left: AOI labels depict surfaces defined for analysis of eye movements (Object 1, Object 2, Occluder, and Background) and did not appear in the images colored by the subjects. (A) Depicts an example of a “disconnected” percept of the Target Object, where the two Object surfaces (Object 1 and Object 2) are reported as separate. All participants invited into the experiment began the experiment with a “disconnected” percept. Nonperceivers persisted with

this disconnected percept as assessed in post-training test. (B) Illustrates a “connected” percept where the two Object surfaces are reported as being part of the same object. Perceivers changed their percept from “disconnected” to “connected” during exposure. (C) Illustrates percentage of Perceivers and Nonperceivers after exposure in the Training and Control conditions ($n = 40$, behavioral participants only).



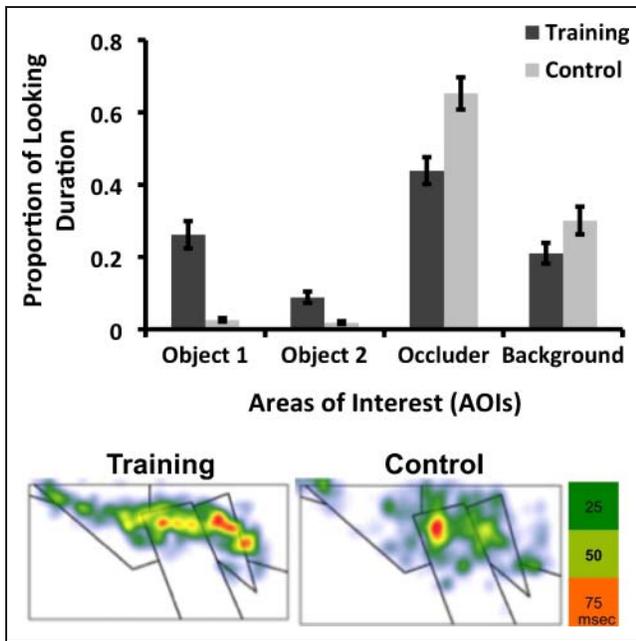


Figure 3. Gaze distribution of non-perceivers across exposure conditions. Top: Illustrates the distribution of looking to the Target Scene for Nonperceivers only in the Training and Control conditions. Proportion of looking (msec) is presented for each of the four AOIs used in the eye tracking analyses. Error bars represent the *SEM*. Bottom: Heat maps depict fixation durations for a representative Nonperceiver in the Training and Control Conditions (left and right, respectively). Colors in the heat map represent the average fixation duration in milliseconds for a given region of the Target Scene for the duration of exposure. These regions are subsets of the AOIs and sampled at high resolution; thus, the average duration for each region illustrated is much lower than the proportion of looking for a given AOI. The legend depicts the average fixation in milliseconds for three sample colors. Warmer colors depict longer average fixations.

how they viewed the scene (e.g., 1 and 2 are part of the same object, which is behind 3 and 4 is the background). The purpose of this task was to test correspondence between verbal report of perception during the experiment with participant’s posttraining test.

In the posttraining test, we labeled those who persisted in their initial “disconnected” percept as Nonperceivers, whereas Perceivers changed their percept and identified the Target Object as “complete” (Figure 2A and B, respectively). We will refer to the difference between Perceivers and Nonperceivers as an effect of the Posttest Group.

Task Procedure

Behavioral participants were randomly assigned to one of two conditions (Training condition $n = 20$, Control condition $n = 20$). Participants received three blocks of passive exposure to the visual scenes. During each block, participants were presented with sequential alternating Target and Paired Scenes. Each Scene Type was presented 14 times per block; Paired Scene order was randomly determined with equal frequency. Scenes were

presented for 3 sec with a 3-sec ISI during which a blank, white screen was presented.

Scenes were presented using SMI Experiment Center (2.5, 3.0) on a white screen measuring $24.5^\circ \times 12.5^\circ$ of visual angle. Stimuli were presented on a 22" monitor. Tracking distance for the SMI system was approximately 70 cm.

Eye-tracking Apparatus and Preprocessing

Eye position was tracked using an SMI system (Sensorimotor Instruments, Needham, MA) sampling at 60 Hz and native iView software. For behavioral participants, an SMI RED system was used with a 5-point calibration, 4-point validation routine. Average error was 0.47° and 0.50° in the x and y coordinates, respectively.

We used software specific to the SMI (BeGaze 3.0) to identify and extract fixations the four areas of interest (AOIs) for the Target Scene. Fixations were defined with a maximum dispersion of 100 pixels (2° of visual angle) and a minimum duration of 80 msec (native settings). AOIs included each surface of the Target Object (Object 1, Object 2), the occluder, and the background (Figure 2A, B). We extracted duration of looking to these four AOIs for each participant and trial. We calculated the proportion of looking durations in milliseconds for each AOI per trial as the total duration of looking per AOI/the total duration of looking for the trial for all AOIs. Figures depicting eye movement analyses (Figures 3 and 4) present the averages and standard errors of the raw proportions of looking duration. These raw proportions were submitted to an arcsine transform before the data were submitted to parametric tests. The arcsine transformed data are in radians.

To consider fixation durations against chance performance, we considered the case where fixations are randomly distributed and assumed that the proportion of fixations to each AOI would be equivalent to the proportionate surface area of each AOI. The proportion of pixels relative to the total scene size was determined for each AOI. Then, for each trial for each AOI, the proportion of

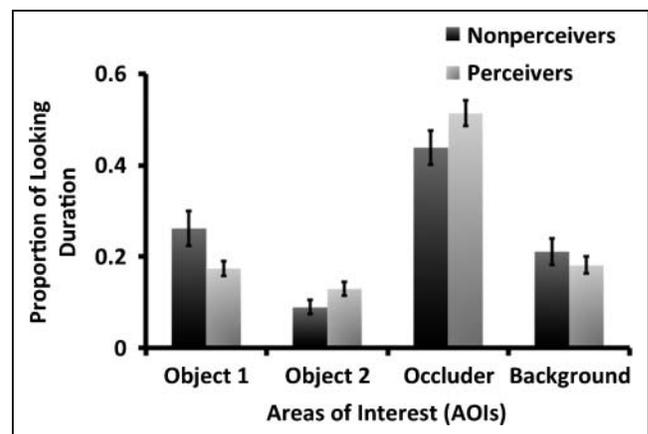


Figure 4. Illustrates distribution of looking across AOIs for Perceivers and Nonperceivers in the Training condition only. Error bars represent *SEM*.

fixations expected by this measure was subtracted from the proportion of fixations observed. Both the proportions were submitted to an arcsine transform before subtraction. To consider eye movements as a function of exposure, data were binned across every seven presentations of the Target Scene, equaling six Training Intervals.

MATERIALS AND METHODS SPECIFIC TO NEUROIMAGING PARTICIPANTS

Pretraining and Posttraining Tests

To ensure their appropriateness for the study before scheduling a scan, the pretraining test for neuroimaging participants was conducted in a separate session. The scanning session was scheduled as soon as possible (M difference = 9.62, SD = 7.53 days). Fourteen participants were found to have a “completed” percept during pretraining test and were not enrolled in this study’s imaging session but participated in another study in the laboratory. Two enrolled participants were found to have a “connected” percept at the beginning of the scan, as indicated by their on-line behavioral responses before the first block of exposure and verbal debriefing. Their data were excluded from subsequent analyses.

Task Procedure

Neuroimaging participants were assigned to the Training condition only. Stimulus presentation was identical with the exception of a longer ISI. The ISI for neuroimaging participants lasted 12 sec to accommodate the time course of the BOLD response. Stimuli were presented by rear projection onto a screen viewed through a mirror box mounted above the head coil. Visual angle was calculated from the screen to the position of mirror within the magnet and was the same as the behavioral experiment. An SMI iView X MRI-LR infrared eye tracker was used with a 9-point calibration, 4-point validation routine. Average error for participants in the scanner was 1.04° and 1.72° in x and y coordinates, respectively.

fMRI Data Acquisition

Images were acquired using a 3-T Siemens Trio MRI scanner (Siemens, Erlangen, Germany). A 3-D localizer was run (AAScout) to position the slices for the remainder of the sequences. This localizer was rotated to ensure whole-brain coverage; all other images were collected in the same oblique angle. A high resolution anatomical image (MultiEcho MP-RAGE: 1.20 mm isotropic voxel size, repetition time (TR) = 2200 msec, inversion time = 1100 msec, flip angle = 7°, 4× acceleration, 144 slices, bandwidth = 651 Hz/Px) was collected for 3-D localization and morphometric analyses.

A slow event-related design was used for functional imaging runs allowing for a direct mapping of the hemo-

dynamic response onto each visual presentation. EPI was used to measure the BOLD signal as an indication of cerebral brain activation during three blocks of exposure. EPI images were aligned to the whole brain MP-RAGE anatomical image (TR = 3000 msec, echo time [TE] = 28 msec, flip angle = 90°). Forty-two oblique slices were collected of 3-mm thickness and 0-mm gap (64 × 64 mm in-plane resolution) were collected for 168 repetitions (including two discarded acquisitions at the onset of each of the three runs). Two participants had fewer than the 508 total repetitions collected. Because of an accidental squeezing of the emergency squeeze ball, one participant had the last 53 repetitions from the second run not collected. Because of a scanner error, another participant had only 102 repetitions from the final run collected. However, both participants received the same stimulus exposure, and eye-tracking measures continued to be collected.

After EPI sequences, we acquired some additional MRI data. First, anatomical T2-SPACE images (high-resolution turbo-spin-echo with high sampling efficiency, 1.20 mm isotropic voxel size, TR = 2800 msec, TE = 327 msec, 144 slices, bandwidth = 651 Hz/Px) were collected at locations identical to the functional image for localization purposes. Finally, functional images (T2* BOLD) were collected while participants were in a resting state (3.0 mm isotropic voxel size, TR = 3.0 sec, TE = 30 msec, flip angle = 85 degrees, 47 transverse slices aligned approximately to the AC–PC plane, no skip, no dummy scans, fat saturation on). All lights were turned off. Participants were instructed to simply let their mind wander and rest but to keep their eyes open and to stay awake during this scan.

Image Processing and Analysis

Functional imaging data were processed and analyzed with the Analysis of Functional NeuroImages (AFNI) software package (Cox, 1996). EPI and anatomical images were deobliqued. Images then underwent (1) registration to the first image volume, (2) alignment to the high-resolution anatomical data set (MP-RAGE), and (3) smoothing with an isotropic 6.0-mm Gaussian kernel. Time series were normalized to percent signal change by dividing signal intensity at each time point by the mean intensity for that voxel for that run and multiplying the result by 100. A model was fit for each subject that included regressors for the Target and Paired Scenes (collapsing across the three Paired Scenes). General linear modeling was performed to fit the percent signal change time courses to each regressor. Linear and quadratic trends were modeled in each voxel to control for correlated drift. Motion regressors, calculated during preprocessing in three dimensions (roll, pitch, and yaw), were also included.

A linear mixed-effects model included factors for subject (random effect), Scene Type (Target vs. Paired) and Posttest Group (Perceiver vs. Nonperceiver) was run within AFNI using functions from the R software package (www.R-project.com, Vienna, Austria, 2005). Correction

for multiple comparisons was applied at the cluster level following Monte Carlo simulations conducted in the AlphaSim function within AFNI. This calculation determines the probability of obtaining a false positive for the 3-D image using individual voxel probability threshold in combination with a cluster size threshold. Spatial correlation between voxels was assumed. For the prescription used (64×64 voxel matrix, 42 slices, 3.0 mm^3 voxels) and preprocessing techniques (Gaussian filter applied of 6.0 mm), a 1000-iteration Monte Carlo simulation was run. The simulation revealed that when probability threshold is set at $p = .05$, 63 contiguous voxels are required to correct for false positives to $p < .05$. Follow-up tests determined the direction of results for linear mixed-effects interaction regions using extracted beta weights.

For the interaction of Scene Type and Posttest Group, the region encompassing the right MTL also spanned the right hippocampus. To consider the individual contributions of hippocampus in isolation, a separate mask was created to cover the right hippocampus only. This mask was a sphere with its origin at coordinates $-32, 22, -11$ with a radius of 3.5 mm (Figure 5A).

RESULTS

Each participant's percept of the Target Scene was tested before and after Training or Control exposure. In addition to verbal report, participants colored the black and white scenes, allowing them to report their perception without time constraints or the ambiguity of verbal responses. All subjects included in the analyses reported a pretraining test "disconnected" percept where the two

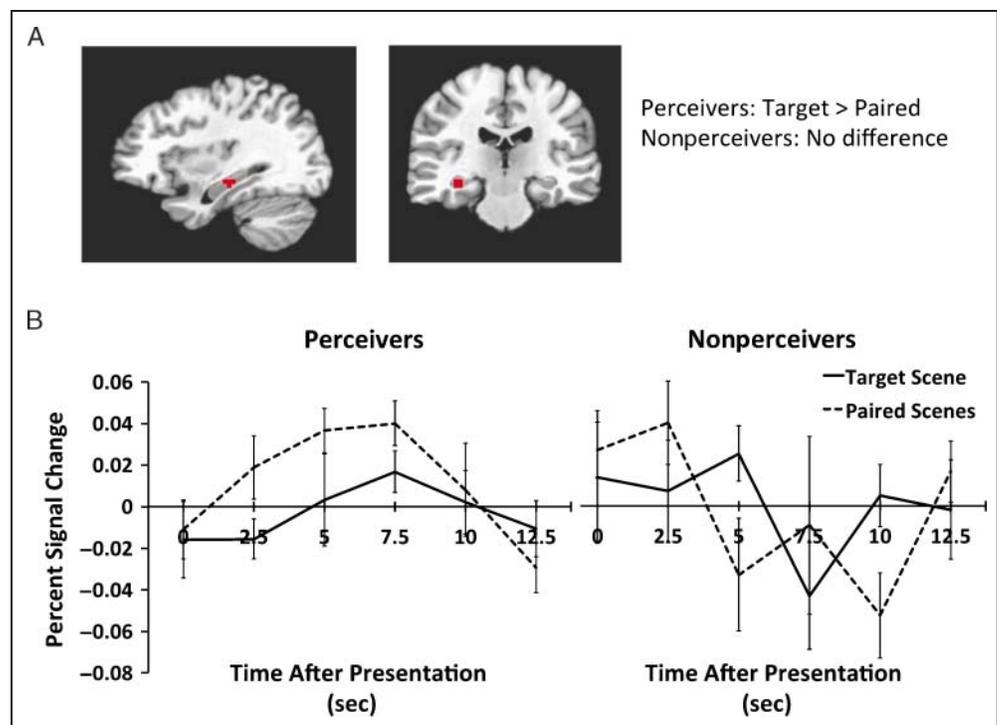
surfaces of the Target Object were colored as two separate objects and not as an occluded whole. In the posttraining test, we determined the proportion of Nonperceivers (those who persisted in their initial "disconnected" percept) and Perceivers (those changed their percept and identified the Target Object as complete, Figure 2A and B, respectively).

Behavioral Outcomes

Training with Variable Exposure to the Target Object Is an Effective Means of Driving Perceptual Change

Participants in the Training condition were more likely to become Perceivers, as indicated by posttraining test performance. In the Training condition, 65% were Perceivers whereas 35% were Nonperceivers. The Control condition yielded only 20% Perceivers whereas 80% remained Nonperceivers (Figure 2C). Behavioral outcomes for behavioral participants were submitted to a Pearson chi-squared test of independence with Condition (Training vs. Control) and Posttest Group (Perceiver vs. Nonperceiver). This test established a nonuniform distribution across Conditions: $\chi^2(1, n = 40) = 8.29, p < .005$. Thus, receiving variable views of the novel object supported in changes in perception of the occluded Target Object from "disconnected" to "complete." Moreover, repeating the same Target Scene in isolation, as in the Control condition, was not sufficient to power such a perceptual shift. Repeating this analysis with neuroimaging participants included does not change the rejection of the null hypothesis, $\chi^2(1, n = 59) = 7.11, p < .01$. An additional chi-squared

Figure 5. The right hippocampus responds to Paired Scenes in Perceivers only. (A) Visual depiction of this ROI, which was calculated as a subset of a larger ROI encompassing the right MTL. (B) The time course of percent signal change for Perceivers (left) and Nonperceivers (right) to both Scene Types (Target Scene and Paired Scenes). Error bars depict the SEM.



test, including factors of Posttest Group and Location (behavioral vs. neuroimaging participants) revealed no differences in the distribution of Perceivers versus Nonperceivers across Testing Locations, $\chi^2(1, n = 39) = 1.232, p = .267$, suggesting that the modest differences in methods across Testing Locations does not significantly affect behavioral outcomes in this task.

Eye Movement Patterns

We parsed the Target Scene into four AOIs (Object 1, Object 2, Occluder, and Background; Figure 2). The arcsine transformed proportion of looking duration (radians) for each AOI were submitted to parametric tests (e.g., ANOVAs) to consider patterns of looking to those AOIs as a function of Condition and Posttest Group.

Regularity in the Training Condition Biased Looking to the Target Object

We submitted the proportion of looking per Target Scene to a mixed-design ANOVA using the within-subject factors of AOI: 4 (Object 1, Object 2, Occluder, and Background) and Exposure Interval: 6 and the between-subject factor of Condition: 2 (Training \times Control) and Post-Test Group: 2 (Perceiver \times Non-Perceiver) for the behavioral participants ($n = 40$). We found a main effect of Condition, $F(1, 36) = 9.78, p < .005$. However, this subtle effect (difference in proportion of looking duration = 0.004) did not interact with Posttest Group. The analysis also yielded a main effect of AOI, $F(1.3, 108) = 98.7, p < .001$, and an AOI \times Condition interaction, $F(1.3, 108) = 16.9, p < .001$. Because of a violation of the assumption of sphericity, all tests involving within-subject factors were Greenhouse–Geisser corrected. Follow-up tests revealed greater proportion of looking allocated to the Object Surfaces (Object 1 and Object 2) in the Training relative to the Control condition and a greater proportion allocated to the Occluder and Background in the Control condition (all $ts(39) > |3.6|, ps \leq .001$, Bonferroni-corrected alpha set to .05/4). We include Figure 3 to specifically illustrate the difference in sampling between Nonperceivers in the Training and the Control conditions. Variable exposure in the Training condition powers efficient object-centered sampling even when the participant does not ultimately connect the Object Surfaces.

Sampling Differences between Perceivers and Nonperceivers in the Training Condition Are Specific to the Target Object

A mixed-design ANOVA of the Training condition ($n = 39$; Figure 4) with within-subject factors of AOI: 4 and Training Interval: 6 and between-subject factors of Posttest Group: 2 ($n = 22$ Perceivers, 17 Nonperceivers) and Location: 2 (behavioral vs. neuroimaging participants) revealed a main effect of AOI, $F(1.95, 105) = 54.5, p < .001$,

Greenhouse–Geisser corrected. Participants in the Training condition looked most at the Occluder and they looked least at the Object 2 surface (all tests with Occluder and Object 2, $ts(38) > |3, 6|, ps \leq .001$, Bonferroni-corrected alpha $p = .008$). This was qualified by a marginally significant AOI by Posttest Group interaction, $F(1.92, 105) = 2.7, p = .08$. We used planned comparisons to examine differences in looking distributions for Perceivers and Nonperceivers. Nonperceivers had a significant difference in looking between the two Object Surfaces, $t(16) = 3.9, p = .001$. However, this effect was not reliable in Perceivers, $t(21) = 2.0, p = .057$, suggesting a more even distribution of looking between the Object Surfaces. Specifically, Nonperceivers looked significantly more at Object 1 than at Object 2, whereas Perceivers only looked marginally more at Object 2. A more even distribution, exhibited by Perceivers, may be conducive to extracting feature correlations across Object Surfaces and Paired Scenes.

We found no main effects associated with Training Interval or Testing Location in the Training condition. As above, all tests with within-subject factors employed a Greenhouse–Geisser correction where sphericity was violated. There was a significant Training Interval by Posttest Group by Location interaction, $F(2.6, 175) = 3.5, p = .005$. Follow-up tests were conducted on behavioral and neuroimaging participants separately. A mixed-design ANOVA including within-subject factors of AOI: 4 and Training Interval: 6 and between-subject factors of Posttest Group: 2 reveal a significant Training Interval by Posttest Group interaction for neuroimaging participants, $F(2.0, 85) = 4.34, p = .019$, but not for behavioral participants. This subtle effect is likely the result of the differences in timing across Testing Locations. However, these differences do not seem to meaningfully or systematically be reflected in the AOI distributions.

Correcting for Size of AOIs Confirms These Patterns

AOIs in the current task vary in their surface area and shape. We re-examined looking patterns while controlling for differences in surface area. If fixations were randomly distributed, the proportional looking to each AOI would be equivalent to their proportionate surface area. We calculated the proportion of looking expected per AOI as the proportion pixels for each AOI relative to total scene size and then compared the proportion of looking observed with this baseline measure. This difference score was statistically compared with zero (difference score if observed and expected values are equal). For the Control condition, we found that participants looked significantly more at the Occluder than would be expected by the surface area of this AOI and less to all other surfaces, $ts(19) > |6.6|, ps < .01$. This is perhaps because the Occluder is centrally located in the Target Scene. We found a markedly different pattern for the Training condition. Participants looked at both Object Surfaces, in addition to the Occluder, more than would be expected by surface area, and looked

less at the Background, $t(38) > |2.8|$, $ps < .01$. Perceivers in the Training condition showed significantly more looking to the Object 2 surface, $t(21) = 3.7$, $p = .001$, than would be expected as a function of surface area, while Nonperceivers showed greater than expected looking to the Object 1 surface, $t(16) = 2.7$, $p = .017$. Given the relative differences in surface area between these regions (Figure 2A, B), the bias in looking toward Object 2 in Perceivers is again indicative of a more even net distribution of looking (Figure 4).

Finally, we submitted these corrected data to the ANOVAs reported for the raw proportion of looking described in detail above. This resulted in virtually identical pattern of effects confirming that these analyses are not confounded by surface area differences across AOIs. The exception is an additional main effect of Training Interval, $F(5, 180) = 2.41$, $p = .038$, in the mixed-design ANOVA with within-subject factors of AOI and Training Interval and between-subject factors of Condition and Posttest Group for behavioral participants only. We find no interaction of Training Interval and Posttest Group. A marginal interaction between Training Interval, AOI, and Condition failed to reach significance, $F(15, 540) = 1.65$, $p = .058$.

There Is No Indication that Sampling Differences Are Related to General Scanning or Mechanics of Eye Movements Independent of the Task Manipulation

We explored the possibility that differences in gaze distribution across AOIs could simply be accounted for in differences mechanics of scanning behavior and not directed eye movements resulting from the presence of visual regularities. To this end, we examined the number of fixations, average and total duration of fixations (in microseconds) and scan path length measured in pixels across Condition and Posttest Group. These data were submitted to separate mixed-design ANOVAs including the within-subjects factor of Training Interval (6) and the between-subjects factors of Condition (Training vs. Control) and Posttest Group (Perceiver vs. Nonperceiver). Looking at the number of fixations per trial, we find a main effect of Training Interval, $F(5, 270) = 2.53$, $p = .037$, with fixations decreasing over exposure in both Perceivers and Nonperceivers and across Conditions. There are neither main effects nor interactions associated with either Posttest Group or Condition. Examining the total duration of fixation, we find a marginally significant main effect of Training Interval, $F(4.001, 270) = 2.107$, $p = .081$, with a Greenhouse–Geisser correction applied. As with fixation count, there is a general decline in total fixation duration per trial over exposure. We also find a marginally significant main effect of Condition, $F(1, 54) = 3.417$, $p = .07$, with participants in the Training condition fixating on average for less time per trial ($M = 2081.85$, $SD = 393.56$) than those in the Control condition ($M = 2276.69$, $SD = 448.88$). There are no effects across Conditions or Posttest Group in average fixation duration per trial. Turning to scan path length, we again find a main

effect of Training Interval, $F(3.68, 270) = 2.52$, $p = .047$, with a general decrease with exposure in both Perceivers and Nonperceivers and across Condition and a Greenhouse–Geisser correction applied. There were no main effects of Posttest Group or Condition nor interactions with these factors. In summary, across the measures examined—number of fixations, total and average fixation duration, and scan path length—there were no consistent and significant differences across either Posttest Group or Conditions, suggesting that differences in proportion of looking across AOIs are not driven by general differences in looking behavior.

Eye Movement Patterns: Conclusions

Taken together, we find that variable exposure is relevant for sampling. However, differences between Perceivers and Nonperceivers are restricted to a bias for a more even distribution of looking to the Object Surfaces in Perceivers relative to Nonperceivers. That is, there is no convincing evidence that sampling alone is sufficient to catalyze the shift in object perception between Perceivers and Nonperceivers. Indeed, sampling differences in the Training relative to the Control condition in Nonperceivers are obvious in Figure 3, even as the subjects in the Training condition do not make the perceptual shift. We turned to neuroimaging data to expose other relevant processing differences between Perceivers and Nonperceivers.

Neuroimaging Results

We collected fMRI data in the Training condition and conducted a whole-brain analysis for Posttest Group (Perceivers vs. Nonperceivers) by Scene Type (Target vs. Paired; Figure 1). All main effects and interaction statistics are corrected for multiple comparisons at $p < .05$. t Values are reported for follow-up analyses. We focus first on the interaction of Scene Type and Posttest Group.

fMRI Data Indicate Important Differences in Neural Activations between Perceivers and Nonperceivers

To preview the data, Perceivers are showing a larger activation for the Paired relative to the Target Scenes, and Nonperceivers are showing no differences across Scene Type (see Table 1). This pattern is evident in the right hippocampus (see Methods Image Processing and Analysis section; Perceivers Paired > Target, $t(8) = 2.29$, $p = .05$, no effect for Nonperceivers; Figure 5) as well as bilateral portions of both the thalamus (Perceivers Paired > Target, $ts(8) = 2.56$, $p < .05$, no effect for Nonperceivers) and caudate nucleus (Perceivers Paired > Target, $ts(8) > 3.2$, $ps \leq .012$, no effect for Nonperceivers).

This pattern was also evident in regions spanning bilateral portions of the middle and superior frontal cortices, as well as the right anterior cingulate. Specifically, these activations included bilateral premotor cortex (~BA 6,

Table 1. Interaction of Scene Type \times Posttest Group

Side	Areas	Coordinates		
		X	Y	Z
R	Anterior cingulate	12	25	15
R	Caudate nucleus	15	14	13
R	Cerebellar tonsil	24	-57	-49
R	Cingulate gyrus	11	9	28
R	Fusiform gyrus	46	-3	-23
R	Hippocampus	36	-19	-11
R	Inferior temporal gyrus	46	-3	-30
R	Middle frontal gyrus	47	-18	-7
R	Middle temporal gyrus	47	-11	-9
R	Superior frontal gyrus	19	28	49
R	Superior temporal gyrus	45	24	32
R	Thalamus	5	-16	8
L	Anterior cingulate	13	-39	0
L	Caudate nucleus	-14	14	14
L	Inferior frontal gyrus	-32	39	2
L	Inferior temporal gyrus	-50	-8	-22
L	Middle temporal gyrus	-52	-8	-15
L	Superior frontal gyrus	-22	14	47
L	Superior temporal gyrus	-45	-11	-5

Corrected to $p < .05$.

Perceivers Paired $>$ Target, $t(8) > 2.7$, $ps < .05$, no effect for Nonperceivers) and the right dorsolateral pFC (\sim BA 9/46, Perceivers Paired $>$ Target, $t(8) = 2.5$, $p = .037$, no effect for Nonperceivers; Figure 6).

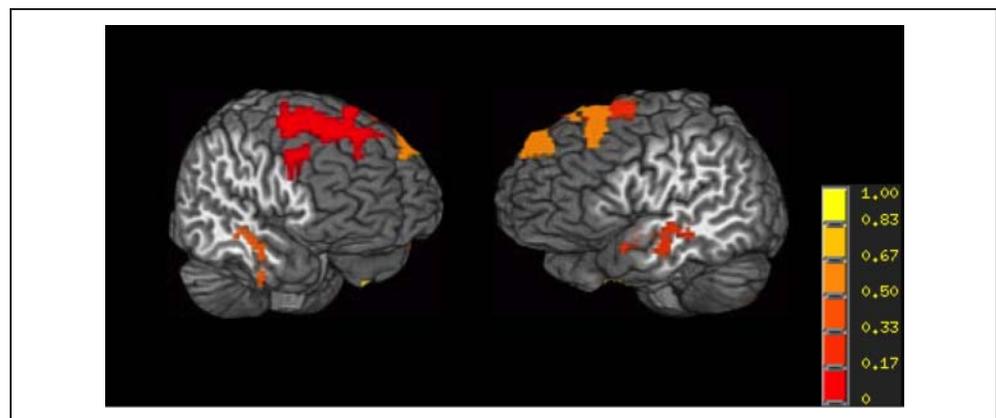
Temporal lobe regions spanning the left superior, middle temporal, and inferior temporal gyrus (Figure 6) exhibited the same pattern as reported above (inferior temporal gyrus, \sim BA 20, Perceivers, Paired $>$ Target, $t(8) =$

3.32, $p = .01$, no effect for Nonperceivers). The set of regions that did not follow this pattern included the right inferior, middle, and superior temporal gyri. This set of regions only showed a marginally reliable effect for Paired $>$ Target scenes in Perceivers, $t(8) = 1.85$, $p = .1$. However, they also showed the reverse Target $>$ Paired scene effect in Nonperceivers, $t(9) = 2.36$, $p = .04$.

We asked whether activation of regions in this interaction was related to distribution of looking on task. As discussed, Perceivers and Nonperceivers were largely differentiated by differences in proportion of looking directed at the Object Surfaces. We conducted simple AOI (Object 1 \times Object 2 \times Occluder) \times Posttest Group (Perceiver \times Nonperceiver) ANOVAs separately modeling Target and Paired Scene beta weights per ROI as covariates. The only region that explained any of the variance in the distribution of proportional looking was the thalamus. Activations in both the Target, $F(2, 30) = 12.6$, $p < .001$, and Paired Scenes, $F(2, 30) = 4.89$, $p = .015$, interacted with the AOI variables. A positive correlation between thalamic activity in the Target scene and proportion of looking directed at the Occluder was found in Perceivers, $r(9) = 0.81$, $p = .003$, and Nonperceivers, $r(10) = 0.78$, $p = .003$. A positive correlation between thalamic Paired Scene activity and proportional looking at Object 2 obtained marginally in Nonperceivers, $r(10) = -0.62$, $p = .03$, and reliably in Perceivers, $r(9) = 0.73$, $p = .01$. We verified that simple fixation count and average duration of fixation metrics did not interact with this activation in the same analysis, indicating that this finding is not being driven by general eye movement differences.

See Table 2 for a list of regions active in the Scene Type main effect. We highlight here regions relevant to object recognition. Specifically, we found greater activations for the Paired relative to the Target scenes in bilateral portions of the LOC (\sim BA 19), right inferior temporal and fusiform gyri (\sim BA 37), and bilateral parahippocampal gyrus. Because of the differences in scene frequency within each Scene Type, we examined all regions for patterns consistent with repetition suppression (Grill-Spector, Henson, & Martin, 2006) but found no clear indication of this effect.

Figure 6. Depicts activations generated by the Posttest Group \times Scene Type interaction (listed in Table 1). The ROIs presented here include the right dorsolateral pFC and bilateral temporal lobe regions. The colored bar reflects the value intensity for the interaction per region.



Many of the regions active in the Posttest Group main effect were subsumed by the Posttest Group by Scene Type interaction. However, activations in portions of the parietal lobe were unique to the main effect of Posttest Group and showed greater activity for the Perceivers relative to the Nonperceivers (Table 3). These specifically spanned the bilateral postcentral gyri and the inferior parietal lobule. Perceivers also had greater activation in the LOC (~BA 19), relevant to recognition of object shape (Grill-Spector et al., 2001), than Nonperceivers.

DISCUSSION

Using an eye-tracking/fMRI methods approach, we examined how variable experience support changes in object perception. Exposure to variable views of the Target Object in the Training condition catalyzed perceptual change; more participants were able to perceptually complete the Target Object after exposure in the Training (65%) than the Control condition (20%). This result establishes that, despite never seeing the Target Object in its entirety, participants were able to integrate across variable and locally

Table 2. Main Effect of Scene Type

Side	Areas (Paired > Target)	Coordinates		
		X	Y	Z
R	Cuneus	18	-75	24
R	Fusiform gyrus	40	-66	-13
R	Inferior parietal lobule	39	-42	26
R	Inferior temporal gyrus	51	-61	-10
R	Middle occipital gyrus	40	-77	-10
R	Middle temporal gyrus	59	-49	-4
R	Parahippocampal gyrus	32	-21	-24
R	Posterior cingulate	16	-6	11
R	Preceuneus	24	-49	31
L	Caudate nucleus	-23	-34	2
L	Declive	0	-57	-11
L	Fusiform gyrus	-40	-66	-11
L	Inferior parietal lobule	-41	-54	46
L	Middle occipital gyrus	-31	-61	2
L	Parahippocampal gyrus	-24	-42	2
L	Posterior cingulate	-24	-61	18
L	Precuneus	24	-76	35
L	Superior parietal lobule	-31	-61	57
L	Thalamus	-21	-27	5

Corrected to $p < .05$.

Table 3. Main Effect of Posttest Group

Side	Areas (Perceivers > Nonperceivers)	Coordinates		
		X	Y	Z
R	Cingulate gyrus	6	1	42
R	Inferior parietal lobule	47	-27	28
R	Inferior temporal gyrus	52	-9	-30
R	Insula	32	-9	17
R	Medial frontal gyrus	12	40	25
R	Middle frontal gyrus	35	26	42
R	Middle occipital gyrus (LOC)	48	-71	5
R	Middle temporal gyrus	51	1	-15
R	Postcentral gyrus	56	-15	28
R	Precentral gyrus	50	-15	38
R	Precuneus	24	-51	48
R	Superior frontal gyrus	12	46	44
R	Superior temporal gyrus	41	16	-34
L	Cingulate gyrus	-16	24	29
L	Declive	-45	-65	16
L	Inferior parietal lobule	-54	-27	37
L	Insula	-40	-18	15
L	Medial frontal gyrus	-15	40	25
L	Middle frontal gyrus	-27	40	37
L	Precentral gyrus	-34	16	37
L	Postcentral gyrus	-54	-15	38
L	Superior temporal gyrus	-44	16	-24

Corrected to $p < .05$.

ambiguous experiences to arrive at a globally “connected” percept.

We found differences in eye movements to the identical Target Scene across exposure conditions. This confirms the prediction that variable experience with an object is relevant for efficient sampling from scenes containing that object. Additionally, we found differences in eye movements between Perceivers and Nonperceivers in the Training condition. As predicted based on the developmental data (Amso & Johnson, 2006; Johnson et al., 2004), results demonstrate that Perceivers looked more evenly at both Object Surfaces, a pattern that may be reflecting object-based attention. Theeuwes, Mathôt, and Kingstone (2010) found that participants prefer to make eye movements within an object rather than between objects. Matsukura and Vecera (2006) found that object-based attention was evident when perceptual grouping cues were robust. While attention was not directly manipulated in this work, we did find that participants sample more from the two Object Surfaces in the

Training condition and, importantly, Perceivers in this condition show the more equal distribution of looks across the two surfaces of the Target Object. However, it is unlikely that a connected percept catalyzed eye movements to the Object Surfaces in this task. All participants confirmed a “disconnected” percept in the pretraining test. Indeed, even Nonperceivers in the Training condition showed more robust eye movements to the surfaces of the Target Object than Nonperceivers in the Control condition. An even distribution of looking to the Object Surfaces in Perceivers, in addition to this more basic object-based attentional sampling difference seen in all subjects in the Training condition, indicates that object-based attentional sampling effects may emerge from variable environmental exposure to a novel object.

Thalamic activity has been found to be involved in visual perception to the extent that it supports integration across eye movements (Hafed & Krauzlis, 2006; Sommer & Wurtz, 2006). Thalamic activity during Paired Scene inspection related to proportion of looking directed at the Object 2 surface, a small surface that would attract little looking based on relative surface area alone. Only Perceivers looked at this AOI more than would be expected by chance. Integrating information within and across the Paired Scenes during free viewing may have been a necessary component of picking up on the regularity in those scenes and relating that to an even distribution during Target Object sampling.

We found evidence for the involvement of classic object perception regions in Perceivers, indicated by preferential activity across bilateral temporal regions, the LOC, the fusiform gyrus and the IT cortex. The latter three cortical regions have been implicated in object perception. The LOC has been shown to be involved in recognition of object shape (Grill-Spector et al., 2001) and also in amodal completion (Hedg  et al., 2008; Shuwairi et al., 2007; Olson et al., 2004; Lerner et al., 2002; Grill-Spector et al., 2001). IT has been shown to be sensitive to image transformations, including rotation or orientation (Booth & Rolls, 1998). Greater activity in Perceivers during the Paired Scene is an indication that this region may be involved in detecting the Target Object across varying viewpoints. The right IT and fusiform gyrus specifically were the only regions that showed greater activity for Target relative to Paired Scenes in the Nonperceivers. This may reflect object processing at the level of the Target Scene, but not as related to the varying viewpoints of the same object presented in the Paired Scenes.

This work was designed with the goal of examining the role of statistically regular exposure, through variable views of the same object, in changes in object perception in general and amodal completion more specifically. The hippocampus has classically been implicated in the generation of cognitive maps and episodic memory formation (Mishkin, 1978; O’Keefe & Nadel, 1978; O’Keefe & Dostrovsky, 1971). However, previous research has found that both the BG and the hippocampus, in addition to visual (LOC and ventral

occipito-temporal cortex) and frontal cortical regions, contribute to statistical learning in the visual modality (Turk-Browne, Scholl, Chun, & Johnson, 2009; Amso, Davidson, Johnson, Glover, & Casey, 2005). Consistent with these findings, we find parallel visual cortex, caudate, and hippocampal activation in Perceivers in this task. Activity in the BG is often considered in relation to the frequency or predictability of events (Amso et al., 2005; Redgrave, Prescott, & Gurney, 1999). The Target Scene is presented with greater frequency than the individual Paired Scenes. Thus, Perceivers may be tracking the frequency or predictability of individual scenes.

The hippocampus also responded preferentially to Paired Scenes in Perceivers but did not respond to either Scene Type in Nonperceivers. The exact nature of the input that the hippocampus is picking up on is something we can only speculate about. However, in contrast with the BG, the hippocampus may be involved here in learning to associate across sequentially presented visual episodes (Tubridy & Davachi, 2011). Relatedly, Howard, Kumaran,  lafsd ttir, and Spiers (2011) recently showed that the hippocampus responds to changes in the spatial relationship between an object and the contextual scene. They conclude that the hippocampus is involved in rapid item context binding. Perceivers may be using the hippocampus to detect the regular presentation of the same object in different viewpoints across the varying contextual Paired Scenes. As reviewed in Introduction, modeling work has suggested that binding across variable views of an object can in turn give rise to invariant object representations (Stringer et al., 2007). This result dovetails with the findings from Shohamy and Wagner (2008), suggesting that the hippocampus is involved in integrating across overlapping past events to build an internal representation of the structure of the environment.

How might involvement of the hippocampus contribute to change in amodal completion of the Target Object? Stokes, Atherton, Patai, and Nobre (2012) recently identified a role for the hippocampus in using prior experience to guide attentional control in the service of perception. Specifically, they showed that memory shapes perception by optimizing sensory analysis of task-relevant information. As noted, Perceivers distributed looking patterns more equitably across the Object Surfaces in the Target Scene during the Training Condition, indicating effective use of variable experience in ongoing object perception. These behavioral data, paired with fronto-parietal activations (discussed below) mirroring the pattern exhibited by the hippocampus (greater activity for Paired Scenes in Perceivers only), indicate that this may be a fruitful avenue for future research.

Specifically, the myriad frontal cortex activations that differentiated Perceivers from Nonperceivers for Paired relative to Target Scenes are consistent with the view that both bottom-up and top-down processes occur in parallel during complex object processing (Oliva, Torralba, Casthelano, & Henderson, 2003). pFC has specifically been shown to be involved in object recognition (Fenske,

Aminoff, Gronau, & Bar, 2006). As well, inferior parietal cortex has been shown to be involved in the monitoring of target objects. Perceivers preferentially engaged regions of the parietal cortex throughout the task. These regions have been associated with orientation coding and visuospatial transformation of objects (Harris, Benito, Ruzzoli, & Miniussi, 2008; Shikata et al., 2001). These findings suggest that Perceivers more strongly engaged attentional processes that maintain object information across variable views. Indeed, the inferior parietal cortex has been shown to be involved in amodal completion, perhaps in the service of attentional guidance (Shuwairi et al., 2007).

That perception is not static is both a strength and weakness of this study. As a strength, it allows us to present subjects with a training condition and mark the systems involved in perceptual change in one group relative to another. As a weakness, it means that Perceivers may at some point shift back to a “disconnected” percept of the Target Object or Nonperceivers might shift to a “completed” percept. We take these data only as “in the moment” indices of a broader process and not as reflecting anything immutable about processing in either of our Perceiver and Nonperceiver groups. This raises a related point about the possibility of an a priori processing difference driving neural activations between Perceivers and Nonperceivers. The behavioral data described, 65% Perceivers subsequent to Training condition exposure versus 25% subsequent to Control condition exposure, indicate that Posttest Group differences were resultant from the task manipulation. Nonetheless, it will be important to test the Control condition using fMRI in the future for certain verification. The challenge, of course, is that very few participants change their initial percept in the Control condition.

These experiments were motivated by the question of how the visual system develops to support object perception in the first postnatal year. As a model case, we examined the on-line neural and behavioral mechanisms that adults employ in learning to complete an occluded, novel object. The systems we highlight are available in some form early in postnatal development. Behavioral research has established that the hippocampus is available for memory formation early in infancy (Little, Lipsitt, & Rovee-Collier, 1984) and, as reviewed in the Introduction, sampling develops during the first months of postnatal development. Finally, recent research suggests that prior experience with statistical information can support effective categorization and segmentation of perceptual input (Romberg & Saffran, 2010; Fiser & Aslin, 2002). Our data demonstrate that variable experience with an object supports targeted sampling from a visual scene that encompasses the same occluded object. While these results are not restricted to participants who change their perception, Perceivers show a more even looking distribution across the Object Surfaces than Nonperceivers. Neuroimaging data suggest a role for the hippocampus and BG, in concert with visual cortical and fronto-parietal

regions, in using statistically regular experience with an object to enable changes in perception, as evidenced by amodal completion. Future work will focus on testing the predictions generated by these results in infant learning and object perception.

Reprint requests should be sent to Lauren L. Emberson, Brain and Cognitive Science Department, University of Rochester, Meliora Hall, Rochester, NY 14627-0268, or via e-mail: lemberson@bcs.rochester.edu.

REFERENCES

- Amso, D., Davidson, M. C., Johnson, S. P., Glover, G., & Casey, B. J. (2005). Contributions of the hippocampus and the striatum to simple association and frequency-based learning. *Neuroimage*, *27*, 291–298.
- Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, *42*, 1236–1245.
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 780–790.
- Bhatt, R. S., & Quinn, P. C. (2011). How does learning impact development in infancy? The case of perceptual organization. *Infancy*, *16*, 2–38.
- Booth, M. C., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex*, *8*, 510–523.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: Object-based and context-based contributions. *Progress in Brain Research*, *155B*, 3–21.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of new feature combinations by infants. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 15822–15826.
- Gauthier, I., Tarr, M. J., Anderson, A., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.
- Gibson, E. J. (1969). *Principals of perceptual learning and development*. East Norwalk, CT: Appleton-Century-Crofts.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital cortex and its role in object perception. *Vision Research*, *41*, 1409–1422.
- Gross, C. G. (2000). Coding for visual categories in the human brain. *Nature Neuroscience*, *3*, 855–856.
- Hafed, Z. M., & Krauzlis, R. J. (2006). Ongoing eye movements constrain visual perception. *Nature Neuroscience*, *9*, 1449–1457.
- Harris, I. M., Benito, C. T., Ruzzoli, M., & Miniussi, C. (2008). Effects of right parietal transcranial magnetic stimulation on object identification and orientation judgments. *Journal of Cognitive Neuroscience*, *20*, 916–926.
- Hedg e, J., Feng, F., Murray, S. O., & Kersten, D. (2008). Preferential responses to occluded objects in the human visual cortex. *Journal of Vision*, *8*, 1–16.
- Howard, L. R., Kumaran, D.,  lafsd ottir, H. F., Spiers, H. J. (2011). Double dissociation between hippocampal and

- parahippocampal responses to object-background context and scene novelty. *Journal of Neuroscience*, *31*, 5253–5261.
- Johnson, S. P. (2004). Development of perceptual completion in infancy. *Psychological Science*, *15*, 769–775.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, *6*, 185–201.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304.
- Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*, 254–264.
- Lerner, Y., Hendler, T., & Malach, R. (2002). Object-completion effects in the human lateral occipital complex. *Cerebral Cortex*, *12*, 163–177.
- Little, A. H., Lipsitt, L. P., & Rovee-Collier, C. (1984). Classical conditioning and retention of the infant's eyelid response: Effects of age and interstimulus interval. *Journal of Experimental Child Psychology*, *37*, 512–524.
- Matsukura, M., & Vecera, S. P. (2006). The return of object-based attention: Selection of multiple-region objects. *Perception & Psychophysics*, *68*, 1163–1175.
- Mishkin, M. (1978). Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. *Nature*, *273*, 297–298.
- Needham, A., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. *Cognition*, *94*, 215–240.
- O'Keefe, J. N., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*, 171–175.
- O'Keefe, J. N., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon: Oxford University Press.
- Oliva, A., Torralba, A., Castelhano, M., & Henderson, J. (2003). Top-down control of visual attention in object detection. *Proceeding of the IEEE International Conference Image Processing*, *1*, 253–256.
- Olson, I. R., Gatenby, J. C., Leung, H.-C., Skudlarski, P., & Gore, J. C. (2004). Neuronal representation of occluded objects in the human brain. *Neuropsychologia*, *42*, 95–104.
- Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*, 97–107.
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*, 1102–1107.
- Redgrave, P., Prescott, T. J., & Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward error? *Trends in Neurosciences*, *22*, 146–151.
- Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews of Cognitive Science*, *1*, 906–914.
- Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*, 1–46.
- Shikata, E., Hamzei, F., Glauche, V., Knab, R., Dettmers, C., Weiller, C., et al. (2001). Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: An event-related fMRI study. *Journal of Neurophysiology*, *85*, 1309–1314.
- Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*, 378–389.
- Shuwairi, S. M., Curtis, C. E., & Johnson, S. P. (2007). Neural substrates of dynamic object occlusion. *Journal of Cognitive Neuroscience*, *19*, 1275–1285.
- Slater, A., Johnson, S. P., Brown, E., & Badenock, M. (1996). Newborn infant's perception of partly occluded objects. *Infant Behavior & Development*, *19*, 145–148.
- Slater, A., Morison, V., Somers, M., Mattock, A., Brown, E., & Taylor, D. (1990). Newborn and older infants' perception of partly occluded objects. *Infant Behavior & Development*, *13*, 33–49.
- Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*, 374–377.
- Stokes, M. G., Atherton, K., Patai, E. Z., & Nobre, A. C. (2012). Long-term memory prepares neural activity for perception. *Proceedings of the National Academy of Sciences*, *109*, E360–E367.
- Stringer, S. M., Rolls, E. T., & Tromans, J. (2007). Invariant object recognition with trace learning and multiple stimuli present during training. *Network: Computation in Neural Systems*, *18*, 161–187.
- Theeuwes, J., Mathôt, S., & Kingstone, A. (2010). Object-based eye movements: The eyes prefer to stay within the same object. *Attention, Perception, & Psychophysics*, *72*, 597–601.
- Tubridy, S., & Davachi, L. (2011). Medial temporal lobe contributions to episodic sequence encoding. *Cerebral Cortex*, *21*, 272–280.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*, 1934–1945.
- Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual completion in newborn human infants. *Child Development*, *77*, 1810–1821.