Spring 5-23-2019

Changes in hemodynamic response to faces, scenes, and objects in a visual statistical learning task: An fMRI analysis

Aaron T. Halvorsen
University of Nebraska - Lincoln

Changes in hemodynamic response to faces, scenes, and objects in a visual statistical learning task: An fMRI analysis

An Undergraduate Honors Thesis
Submitted in Partial Fulfillment of
University Honors Program Requirements
University of Nebraska-Lincoln

by
Aaron Halvorsen, BS
Biological Sciences and Psychology
College of Arts and Sciences

May 23, 2019

Faculty Mentors:
Matthew Johnson, PhD, Psychology
Mike Dodd, PhD, Psychology

ABSTRACT

Learning causes changes in brain activity and neural connections. Statistical learning is an implicit learning process that involves extracting regularities from the environment and finding patterns in stimuli based on their transitional probabilities. The following study describes an attempt to elucidate temporal changes in hemodynamic activity for three category-specific brain areas using functional magnetic resonance imaging (fMRI). Blood oxygen-level dependent signal (BOLD) responses were collected while subjects viewed faces, scenes, and objects with high and low transitional probabilities in an fMRI scanner. We expected brain activity to show a temporal shift in timing of activation when comparing BOLD signal responses before and after visual statistical learning. Instead, a general, yet insignificant, trend in the magnitude of activation was found. Although these findings suggest category-specific brain areas may undergo magnitude changes in activation for item-specific stimuli in response to visual statistical learning, further confirmatory analyses and comparisons to behavioral data are needed.

Key Words: fMRI, statistical learning, BOLD signal, brain activity

DEDICATION/APPRECIATION

I would like Dr. Matthew Johnson and Dr. Cheng Lim for the opportunity to take the lead on an fMRI project as an undergraduate student. Your patience, guidance, and trust as I completed this project was invaluable to me. I would also like to thank the NSF/EPSCoR Grant #1632849 for funding this research project, as well as the Undergraduate Creative Activities and Research Experience (UCARE) program for their financial support. Thank you also to Dr. Mike Dodd for his support and encouragement through my undergraduate years and this project. Finally, I would like to thank my parents, John and Cheryl, for their unwavering support.

INTRODUCTION

We are constantly bombarded with an enormous number of stimuli from our environment. Despite the ongoing arrival of new sights and sounds, one way we are able to avoid becoming overwhelmed is by finding regularities in those stimuli and forming patterns that help us to make sense of the world. Some of this pattern association is done through intentional and explicit methods requiring mental effort, such as learning where each key is located on the keyboard. Other methods operate below our awareness and happen automatically, such as statistical learning. Statistical learning is a primarily implicit process that allows us to find patterns of regularities in the environment based on their chances of occurring together, or their transitional probabilities (Kirkham, Slemmer, & Johnson, 2002). This type of learning was first demonstrated in a study in which infants were able to distinguish "words" from "non-words" after only two minutes of listening to pseudospeech (Saffran, Aslin, & Newport, 1997). Infants showed differences in listening times depending on the transitional probability of the syllables that comprised the "words" from the "non-words". These infants were demonstrating recognition of a pattern and providing an example of learning taking place without conscious effort.

A defining feature of statistical learning is its implicit nature, which makes it an intriguing, yet elusive, process to study. In the earlier years of statistical learning research, most experiments were behaviorally based and did not include measures of brain activity (Fiser & Aslin, 2002). Since statistical learning is an implicit process, it can be difficult to measure its presence through explicit memory tests and reaction times (Turk-Browne, Scholl, Chun, & Johnson, 2008), although implicit association of the stimuli can sometimes form explicit knowledge of the patterns (Batterink, Reber, Neville, & Paller, 2015). More recent research has focused on elucidating the neural basis of this learning process through neuroimaging. Several

studies have shown differences in brain activation before and after statistical learning in areas such as the superior temporal gyrus, the inferior frontal gyrus, and several locations in the medial temporal lobe (Karusa et al., 2013; McNealy, Mazziotta, & Dapretto, 2006; Schapiro, Kustner, & Turk-Browne, 2012). Additionally, Alba, Katahira, and Okanoya (2008) examined auditory event-related potentials (ERP) and found differences in the N100 and N400 peaks based on subjects' level of auditory statistical learning (Abla, Katahira, & Okanoya, 2008). However, to our knowledge, there has been no research conducted on the time shifts in blood oxygen-level dependent (BOLD) signals during the course of a statistical learning task. Finding differences in the timing of these signals could suggest a mechanism for how the brain represents this type of learning in different areas of the brain and provide a method for determining whether an individual is making pattern associations in the stimuli.

One challenge in investigating these time shifts, as with many fMRI studies, is the time delay between stimulus onset and the resulting BOLD activation peak (Buxton, Wong, & Frank, 1998). BOLD signals are a measure of the amount of deoxygenated hemoglobin present in the blood and can thus be used as a proxy for the amount of oxygen currently located in different areas of the brain. When an area of the brain is activated, this leads to the uptake of oxygen by neurons in that area and necessitates the oxygen that was used to be replenished. The amount of oxygen replenished by blood flow is more than the original baseline amount, and this overshoot is the underlying mechanism behind the observed peak in the BOLD response (Huettel, Song, & McCarthy, 2004, pg. 224-225). However, unlike the fast changes of electrical activity recorded by electroencephalography (EEG), changes in oxygen levels occur more slowly with peaks taking 4-5 seconds to fully form and return to baseline taking up to 12 seconds or more (Buxton, Wong, & Frank, 1998).

This time delay can make studying temporal differences in activation difficult. If a region of interest (ROI) is activated by stimuli that are only seconds apart, many of the BOLD signal peaks would overlap with each other and make monitoring timing of activation for a single item presentation very difficult. In order to separate the temporal patterns of BOLD activity during analysis, category-specific brain areas were used that have been shown in previous research to demonstrate a preference for certain categories of visual stimuli. The three areas we examined were the fusiform face area (FFA), the parahippocampal place area (PPA), and the lateral occipital complex (LOC), which show preferential activity for faces, scenes, and objects, respectively (Kanwisher, McDermott, & Chun, 1997; Epstein & Kanwisher, 1998; Grill-Spector, Kourtzi, & Kanwisher, 2001, respectively). By observing category-specific activation from separate areas in the brain in response to their preferred categories, we were able to parse apart individual activations and isolate BOLD signal peaks despite the stimuli being presented closely together in time. By creating timelines of these peaks, we were able to observe if there were any changes in the timing of activation due to statistical learning.

The purpose of this study is to investigate these temporal shifts in category-specific brain activation when comparing BOLD signal timing before and after visual statistical learning (VSL). We hypothesized that there would be temporal differences in BOLD responses for category-specific brain areas in response to implicitly learned patterns (Figure 1). The experimental design was similar to a VSL study of implicit anticipation when viewing pairs of faces and scenes conducted by Turk-Browne and colleagues (2010). We expanded this paradigm to include faces, scenes, and objects and view items as triplets instead of pairs. We predicted that after VSL has taken place there would be time shifts in the BOLD signal as the implicit awareness of the stimuli patterns would allow for anticipatory activation. Thus, we hypothesized

that the third item in the triplets would start to be detectable during the presentation of the second item, and the second item would be detectable during the presentation of the first item but to a lesser extent since only one item preceded it instead of two.

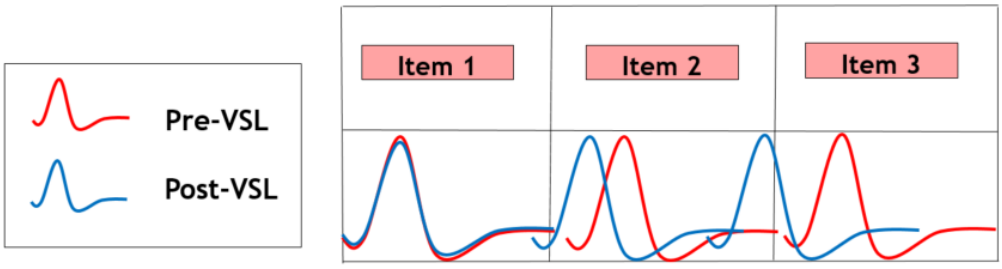

Figure 1. Hypothesis for temporal shifts in activation. As the pattern of items in the high transitional probability triplets are learned, we predicted that BOLD signals would shift earlier in time.

METHODS

Participants. Twenty subjects with normal or corrected-to-normal vision participated in the study for monetary compensation. One subject was removed due to excessive movement and two were removed for problems with falling asleep in the scanner. Thus, seventeen subjects (12 females; mean age = 23.23 sd = 3.19; 14 right-handed, 2 left handed, 1 ambidextrous) were used in the analysis. Although the standard for most fMRI studies is to limit participation to only right-handed subjects to reduce lateralization variability, we did not believe this was necessary since our task did not involve language or words and all brain areas used are present on both the left and right hemispheres. Subjects viewed a series of images while in an fMRI scanner, responded to each image via a cover task, and completed subsequent surprise memory tests.

Stimuli and Task Design. Stimuli comprised faces, scenes, and objects that were presented one at a time with an inter-stimulus interval of one second. Each trial began with a scrambled image that descrambled over the course of one second to reveal the original item

(Figure 2). Descrambling of items was incorporated in hopes of increasing anticipation and focus for each item presentation. The images were recognizable at approximately 600 ms of descrambling, so the expectation was that subjects would be able to recognize what the image would be earlier in the descrambling process as they implicitly learned patterns in the stimuli. The item remained on screen for two seconds after it was fully descrambled. During the entire task, a white dot was shown in the center of the screen to discourage eye movements and aid fixation. Between items, only the white dot appeared on screen. Subjects were encouraged to focus on the white dot and to not move their eyes anywhere else as much as possible.

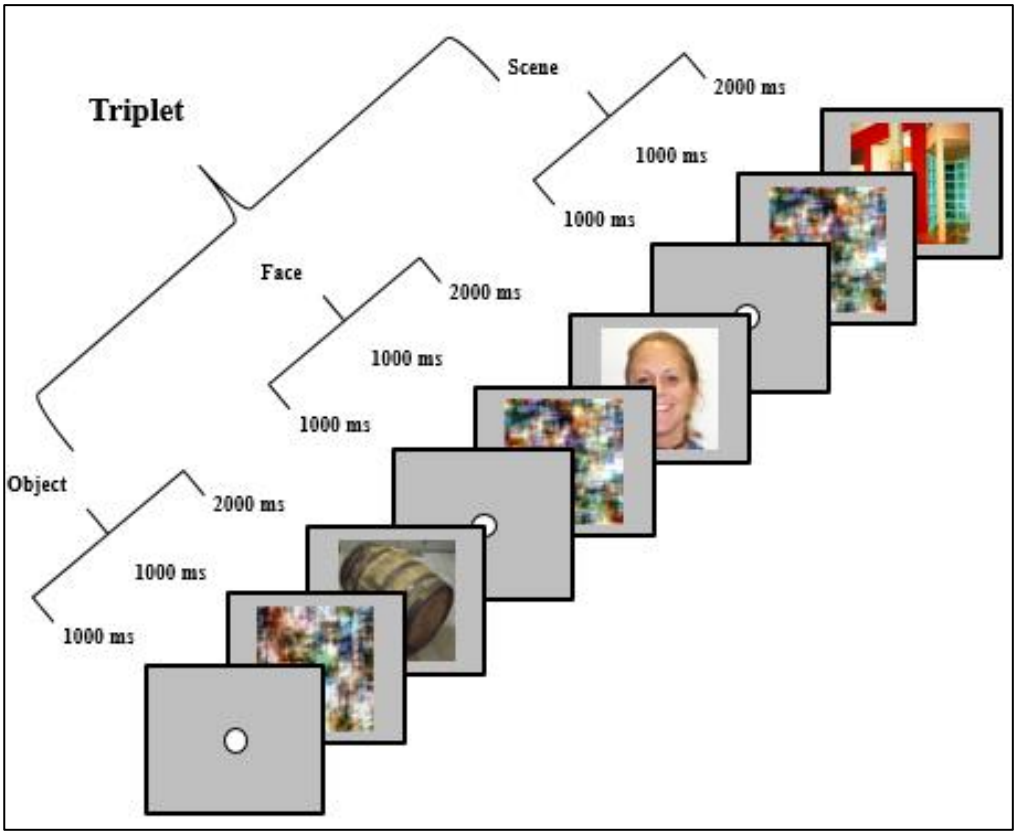

Figure 2. Task presentation of an object-scene-face triplet. Each triplet presentation lasted 12 seconds. Participants completed a cover task of category determination. Using a three-button signaling device, participants were instructed to press a button corresponding to the category of the item being presented as quickly as possible without making errors. Participants responded to faces with their index finger, scenes with their middle finger, and objects with their ring finger.

Unbeknownst to subjects, stimuli were presented as part of triplets with different probabilities of occurring together. Triplets were divided into "strong" and "weak" triplets. In strong triplets, items had 100% transitional probability (e.g. a barrel, a blonde girl's face, and a living room appeared in the same triplet every time). Weak triplets had unique patterns for every presentation (e.g. a clock was presented with a certain male face and a store-front building only one time). Each triplet contained one of each image category. The triplets were balanced so that all six possible permutations of category orders were used equally for both strong and weak triplets (Figure 3).

The main task comprised six runs that were divided into two sets of three runs. Each set contained its own set of stimulus items. The first set consisted of Runs 1, 2 and 3, and the second set consisted of Runs 4, 5, and 6. There were 36 unique stimuli in total for each set (12 faces, 12 scenes, and 12 objects). Every item was viewed an equal number of times to prevent familiarity bias from seeing one item more often than others. Each run consisted of three presentations of every strong and weak triplet combo (for a total of 108 items per run), and thus subjects were exposed to each of the strong triplets nine times per set. The neural signatures of statistical learning can be observed within just a few exposures to the temporal probabilities of the stimuli (Turk-Browne, Scholl, Chun, & Johnson, 2008), so we expected there to be clear, observable differences in brain activation after nine exposures.

Counterbalancing. The experiment design was coded and counterbalanced using the programming language MATLAB and the Psychtoolbox extension. The order of the stimulus sets was counterbalanced such that odd numbered subjects received Set A stimuli first and even numbered subjects received Set B first. Every participant observed the same sets of stimuli, but individual items within the sets were randomly assigned to triplets. For example, every

participant saw a picture of a barrel nine times, but that barrel could have been in a strong triplet for one subject and in a weak triplet for another.

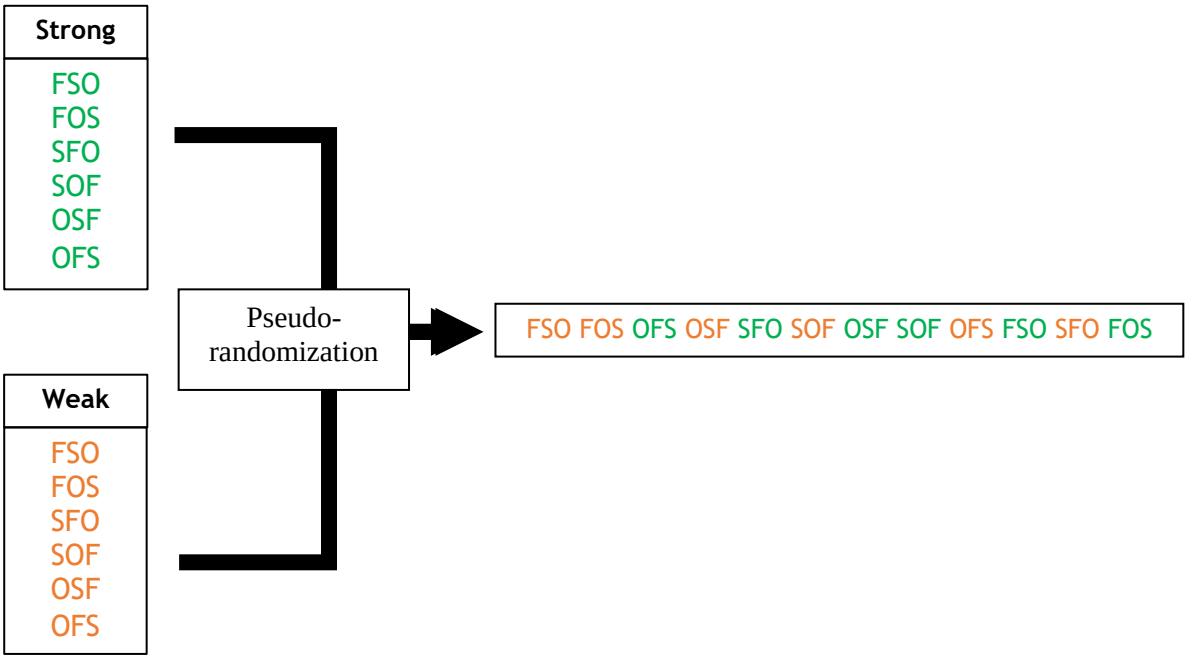

Figure 3. Each run contained the same number of strong and weak triplets and the same number of each order (i.e. FSO). Triplets were randomized as much as was allowed through our counterbalancing conditions. This process was used to create one sequence of the 12 possible triplets (six strong and six weak). Three of these sequences were combined to form one run. The runs were then checked against the counterbalancing conditions again.

Item presentation was counterbalanced in six separate ways to reduce as many category and presentation order confounds as possible. Triplets were pseudo-randomly ordered to create each run (Figure 3). Each triplet contained one face, one scene, and one object. All possible combinations for presentations of faces (F), scenes (S), and objects (O) were removed if any of the following conditions were not met:

1. Repetitions of category between triplets (i.e. FSO OSF)
2. Three or more strong/weak triplets appeared in a row
3. Repetitions of triplet back-to-back (FSO FSO)
4. Image was shown more than once within nine item presentations

5. Two items in weak triplets appeared back-to-back more than once*
6. Two items in weak triplets appeared as the first and third items more than once*

*Since weak items appear in different unique weak triplets every time, these conditions were included to ensure that no implicit association occurred from weak items appearing in the same triplet more than once.

Although statistical learning requires attending to the task, it does not require a conscious effort in order to implicitly recognize patterns in the stimuli (Turk-Browne, Jungé, & Scholl, 2005). To encourage implicit learning of the patterns, an unrelated cover task was included to keep the subjects' focus on the stimuli while trying to prevent conscious awareness of the underlying transitional probabilities. Subjects were given an orthogonal cover task of category determination and instructed to press the button that corresponded to the category that was on screen as quickly as possible while avoiding errors. Instructing subjects to respond quickly was done with the intention to increase anticipation to the stimuli and respond earlier and during the descrambling process. They were given a three-button box in their right hand and told to respond to faces with their index finger, scenes with their middle finger, and objects with their ring finger. A practice task was given before going into the scanner to ensure full understanding of the instructions and to become proficient with pressing the correct button in response to each category of item.

Localizer. After six runs of the main task, a functional localizer task was administered to help identify regions of interest (ROI) for the FFA, PPA, and LOC. The localizer consisted of blocks of items that contained one of four categories: faces, scenes, objects, or scrambled objects. A series of the same category of images was presented one at a time for one second per image. Images would appear one after the other to elicit a BOLD response specific to that category of item. This process was repeated for each category of item. This task provided raw

BOLD signal responses by each brain area that would later be used to show which areas express preferential activity to one type of item (face, scene, or object).

Memory Tests. Immediately following the scanner tasks, subjects were asked if they noticed anything about the stimuli they had just seen or if they noticed any patterns in the way the items were presented. Responses were recorded, and the subjects were informed that some items followed each other every time, while some did not follow each other at all. A surprise memory test was then administered for both sets of stimuli. The two-part memory test was similar to other self-paced familiarity tests used in previous studies of statistical learning. The first part asked participants to rate each triplet on a sliding familiarity score from "Very unfamiliar" to "Very familiar" (Figure 4a). Subjects were instructed to base their rating on how familiar the *sequence* of items was and not the familiarity of the individual images. The second part of the test displayed two triplets on screen, and participants were instructed to choose the triplet that seemed more familiar to them using the same criteria as before (Figure 4b). Unbeknownst to the participants, each pair of triplets contained one strong and one weak triplet.

Figure 4a. First memory test. Participants rated each triplet based on how familiar the sequence of items was to them using a sliding scale. Responses were recorded as 0 (very unfamiliar) to 100 (very familiar) based on dot placement. The middle area was grayed out in order to force participants to respond with at least a slight preference towards one of the two descriptions.

Figure 4b. Second memory test. Participants were shown two triplets and asked to choose the triplet sequence that looked more familiar. Unbeknownst to them, there was always one strong and one weak triplet shown on screen.

Data acquisition/scanning parameters. Functional and anatomical data were collected using a Siemens Skyra 3T scanner with a 32-channel head coil. T1 and T2 anatomical images were obtained for each subject, and functional data was acquired with moderately high resolution (voxel size: 2.5 x 2.5 x 2.5 mm). Whole-brain functional acquisition was collected with TR=1000ms, TE=30ms, field of view = 210mm, multiband factor of 3 slices, 51 slices per volume, and 60º flip angle. Each run lasted 444 seconds, which equated to 444 brain volumes acquired per run. The first four volumes were collected during initial fixation and settling of baseline BOLD signal and were discarded prior to analysis.

Preprocessing. Preprocessing of data was completed using the Statistical Parametric Mapping 8 (SPM8) program (Penny, Friston, Ashburner, Stefan, & Nichols, 2006). Each functional run was corrected for any motion in the scanner. Each anatomical T1 image was aligned to the Montreal Neurological Institute (MNI) template to allow for accurate location comparisons within and between subjects. The subjects' functional data were then aligned to

their own respective anatomical data. Finally, functional data were smoothed using a Gaussian smoothing kernel with a full width at half maximum value of 6mm to increase signal-to-noise ratio (Tabelow, Piëch, Polzehl, & Voss, 2009) and still allow for comparison across subjects (Chen & Calhoun, 2018).

ROI determination. To isolate each ROI, the localizer task was analyzed with a general linear model analysis to determine which brain areas showed the strongest preferential activity to one category or another (i.e. which voxels activated most strongly to faces compared to scenes). A contrast of activity was found for each category to isolate preferential activity. The face contrast was formed by subtracting all of the scene activation from all of the face activation. The scene contrast was formed by subtracting all of the face activation from the scene activation. The object contrast was formed by subtracting all of the activation due to scrambled objects from the object activation. Scrambled objects have been used in previous research as a contrast to distinguish LOC from other areas that are not as object specific (Grill-Spector, Kourtzi, & Kanwisher, 2001).

ROIs were manually located for each subject by using these contrasts and the Neurosynth coactivation maps for reference of general locations of the FFA, PPA, and LOC used in past research (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). The cluster of activity that most closely matched the ROIs on Neurosynth was selected, and then a central voxel of activation was chosen using SPM8 and finding the local maximum activation to find the voxel that had the highest activation in the cluster. The left and right regions for all three areas were approximated for each subject, totaling six areas for each subject (Figure 5).

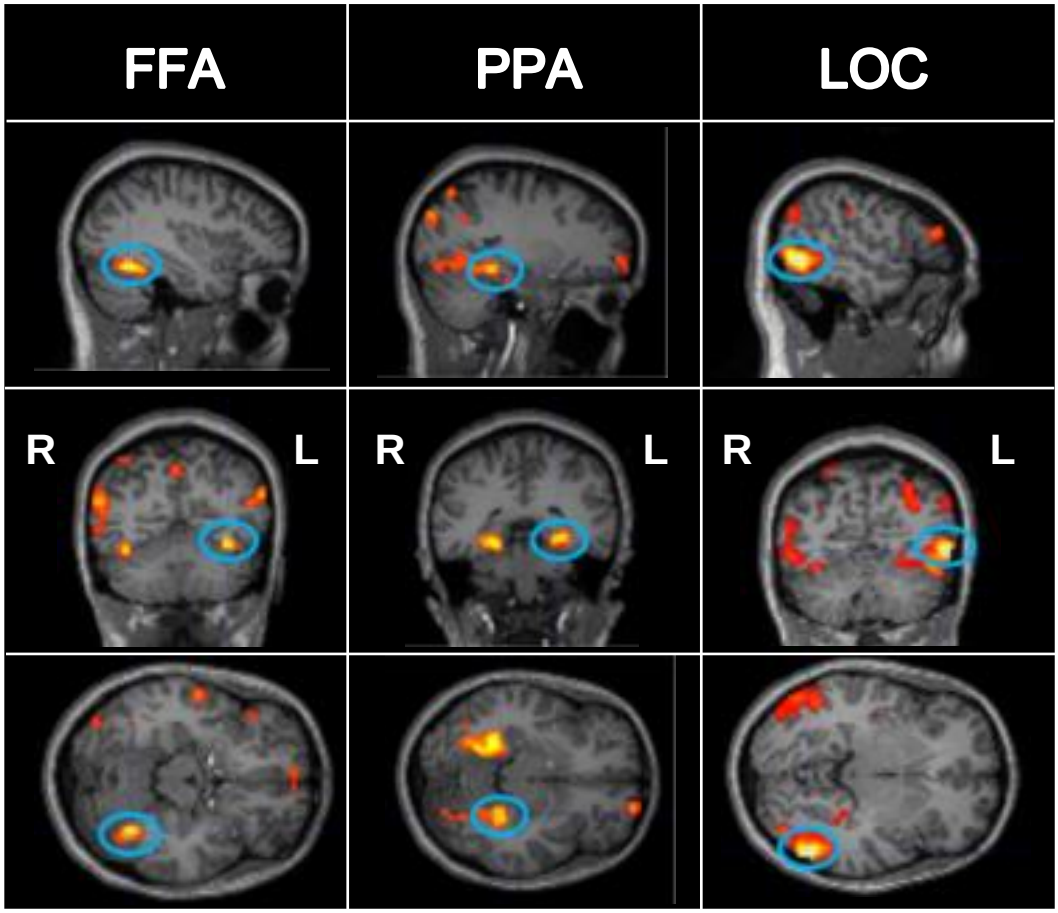

Figure 5. Approximate locations of the left-side ROIs for one subject from three different views (top – sagittal, middle – coronal, bottom – axial).

Sphere extraction. To approximate the general area for a specific ROI, a sphere of activation for each ROI was obtained. Using the ROI coordinates, voxels for analysis were acquired by taking a sphere of activation (radius = 6 mm) around the central voxel of activation for each ROI location. BOLD signals for each voxel for the left and right side of each area were averaged together for each item category. Since the sphere of activation for an ROI does not perfectly encompass the actual voxels activated, there were often some voxels that were included in the sphere average that did not show any activation. This was especially the case when the ROI was located close to the skull or the ear canal. To see if this inclusion of non-activation

voxels affected the analysis, a secondary analysis was run using a large sphere of activation (radius = 10 mm). The 50 voxels from each ROI that showed the highest preferential activation in the larger spheres were found and then averaged and analyzed separately. By taking the voxels with the highest activation, this ensured that no voxels were part of the skull or ear canal.

RESULTS

Implicit/Explicit Awareness. When asked about recognizing a pattern in the stimulus presentation, two participants had a general sense that there was some pattern present but could not fully explain it. Only one participant had full explicit knowledge of the pattern, so BOLD signal ANOVAs and familiarity score analyses were run with and without their data to observe if they had an effect on the implicit VSL data.

Memory Test 1. Familiarity scores for strong and weak triplets were averaged for each subject across the two sets. Average familiarity score was 56.6% for strong triplets and 48.2% for weak triplets. Overall, when combining sets, there was no difference between familiarity scores for strong and weak triplets ($t = 1.778$, $df = 16$, $p = 0.094$).

Memory Test 2. The number of times that each subject correctly picked the strong triplet was recorded and averaged. Subjects correctly chose the strong triplet 55.3% of the time, which was not significant when compared to chance ($t = 1.594$, $df = 16$, $p = 0.130$).

Timelines. The aim of this study was to find shifts in the time course of activation in response to implicitly learned patterns. To compare how BOLD signal timing changed over the course of learning, timelines (Figure 6) were created by averaging the activation of each brain area during the presentation of triplets when their preferred item was third in the triplet (i.e. averaging the FFA activation during the OSF and SOF triplets). The average BOLD response peaks after approximately 4-5 seconds post-stimulus and mostly returns to baseline around 12

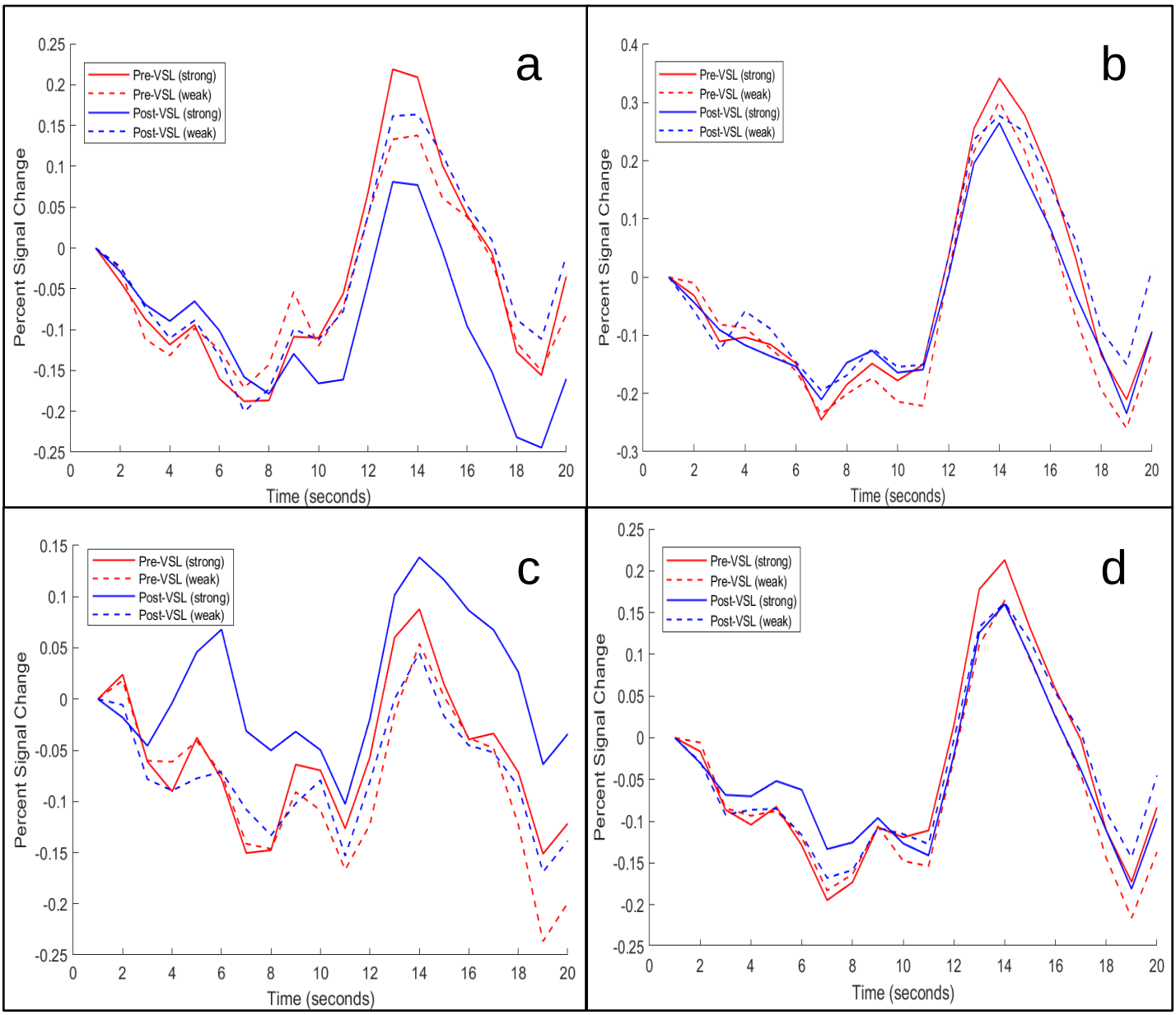

Figure 6. Timelines of activation before and after VSL. Stimuli were presented at 0, 4 and 8 seconds. Peak activation of BOLD responses can be seen at about 4 seconds after each item presentation. Pre-VSL is Run 1 and Post-VSL is Run 3. The graphs show only the data for the triplets that have the preferred item of activation in the third position for each brain area (i.e. only the SOF and OSF triplets used for FFA). Graphs are shown for each brain area: (a) FFA, (b) PPA, (c) LOC, and (d) average of all areas.

seconds post-stimulus (Buxton, Wong, & Frank, 1998). Thus, timelines were 20 seconds long with item presentation at 0, 4, and 8 seconds to allow for full BOLD response to each item. Pre-VSL was defined as the activation averaged in Run 1, which included the first, second, and third exposure to each triplet. Post-VSL was defined as the activation averaged in Run 3, which included the seventh, eighth, and ninth exposures. Timelines for pre-VSL and post-VSL were shown for both strong and weak triplets. Activations were averaged over subject, stimulus set, and exposures during each run. Timelines for the highest 50 voxel analysis and the two analyses without the outlier showed slightly different patterns but were not significantly different.

BOLD Signals. BOLD signals were averaged for each of the three brain areas for each subject and compared in a 2x2x2 repeated measures ANOVA using runs (Run 1 vs. Run 3), strength (strong vs. weak), and item position (first item vs. last item in triplet) as the factors. Averages for each BOLD signal peak were found by averaging across three functional brain acquisitions: directly before the peak, the peak itself, and directly after the peak. Thus, time points 5, 6, and 7 were used for Time 1 and time points 13, 14, and 15 were used for Time 3. Time 1 refers to the first item in the triplet, and Time 3 refers to the last item in the triplet. Separate tests were run for BOLD signal activations for the smaller sphere (Table 1a and 1b) and for the 50 most highly activated voxels in the larger sphere (Table 2a and 2b). The same two tests were repeated after excluding the outlier subject (Table 3a, 3b, 4a, 4b). The main effect of time was present in every analysis, which shows that each brain area responded more strongly to the preferred item than to a non-preferred item. The Run x Time interaction was only significant in the analyses run without the outlier (Table 3b and 4b). Separate follow-up 2x2x2 repeated measures ANOVA analyses for the FFA, PPA, and LOC were run individually to observe the Run x Time interaction for each brain area (Table 5).

Table 1a. Descriptive statistics of voxel activations in the small spheres.

Descriptives

Run	Strength	Time	Mean	SD	N
Run 1	Strong	Time 1	−0.135	0.131	17
		Time 3	0.174	0.083	17
	Weak	Time 1	−0.130	0.113	17
		Time 3	0.123	0.134	17
Run 3	Strong	Time 1	−0.082	0.095	17
		Time 3	0.127	0.119	17
	Weak	Time 1	−0.123	0.184	17
		Time 3	0.137	0.119	17

Table 1b. ANOVA results of voxel activations. Only the main effect of time was significant, which shows that each brain area activated more strongly to its preferred category.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Run	0.002	1	0.002	0.079	0.782
Residual	0.305	16	0.019		
Strength	0.012	1	0.012	0.721	0.408
Residual	0.276	16	0.017		
Time	2.263	1	2.263	164.044	< .001
Residual	0.221	16	0.014		
Run ✳ Strength	4.501e −4	1	4.501e −4	0.020	0.890
Residual	0.362	16	0.023		
Run ✳ Time	0.018	1	0.018	2.650	0.123
Residual	0.110	16	0.007		
Strength ✳ Time	8.438e −5	1	8.438e −5	0.018	0.894
Residual	0.073	16	0.005		
Run ✳ Strength ✳ Time	0.024	1	0.024	2.887	0.109
Residual	0.134	16	0.008		

Note. Type III Sum of Squares

Table 2a. Descriptive statistics of the 50 most highly activated voxels in the larger spheres.

Descriptives

Run	Strength	Time	Mean	SD	N
Run 1	Strong	Time 1	−0.164	0.136	17
		Time 3	0.212	0.096	17
	Weak	Time 1	−0.150	0.119	17
		Time 3	0.165	0.135	17
Run 3	Strong	Time 1	−0.100	0.098	17
		Time 3	0.160	0.132	17
	Weak	Time 1	−0.137	0.195	17
		Time 3	0.181	0.115	17

Table 2b. ANOVA results of the 50 most highly activated voxels. Only the main effect of time was significant, which shows that each brain area activated more strongly to its preferred category.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Run	0.003	1	0.003	0.149	0.704
Residual	0.366	16	0.023		
Strength	0.005	1	0.005	0.273	0.608
Residual	0.309	16	0.019		
Time	3.417	1	3.417	188.138	< .001
Residual	0.291	16	0.018		
Run ✻ Strength	6.834e −4	1	6.834e −4	0.030	0.865
Residual	0.365	16	0.023		
Run ✻ Time	0.027	1	0.027	3.274	0.089
Residual	0.131	16	0.008		
Strength ✻ Time	2.686e −5	1	2.686e −5	0.005	0.944
Residual	0.083	16	0.005		
Run ✻ Strength ✻ Time	0.030	1	0.030	2.627	0.125
Residual	0.181	16	0.011		

Note. Type III Sum of Squares

Table 3a. Descriptive statistics of small sphere activations with outlier removed.

Descriptives

Run	Strength	Time	Mean	SD	N
Run 1	Strong	Time 1	−0.140	0.133	16
		Time 3	0.174	0.086	16
	Weak	Time 1	−0.141	0.105	16
		Time 3	0.116	0.134	16
Run 3	Strong	Time 1	−0.069	0.080	16
		Time 3	0.127	0.123	16
	Weak	Time 1	−0.120	0.189	16
		Time 3	0.137	0.123	16

Table 3b. ANOVA results for sphere activations with outlier removed. The main effect of time is significant again. The Run x Time interaction was significant such that the first item had lower activation in Run 1 and higher activation in Run 3, and the third item had higher activation in Run 1 and lower activation in Run 3.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Run	0.009	1	0.009	0.529	0.478
Residual	0.251	15	0.017		
Strength	0.021	1	0.021	1.214	0.288
Residual	0.255	15	0.017		
Time	2.101	1	2.101	143.872	< .001
Residual	0.219	15	0.015		
Run ✻ Strength	7.266e −4	1	7.266e −4	0.030	0.864
Residual	0.361	15	0.024		
Run ✻ Time	0.028	1	0.028	4.577	0.049
Residual	0.090	15	0.006		
Strength ✻ Time	3.647e −5	1	3.647e −5	0.008	0.931
Residual	0.070	15	0.005		
Run ✻ Strength ✻ Time	0.028	1	0.028	3.266	0.091
Residual	0.129	15	0.009		

Note. Type III Sum of Squares

Table 4a. Descriptive statistics of the 50 most highly activated voxels in the larger spheres while excluding outlier.

Descriptives

Run	Strength	Time	Mean	SD	N
Run 1	Strong	Time 1	−0.168	0.139	16
		Time 3	0.213	0.099	16
	Weak	Time 1	−0.160	0.116	16
		Time 3	0.159	0.137	16
Run 3	Strong	Time 1	−0.089	0.088	16
		Time 3	0.161	0.136	16
	Weak	Time 1	−0.136	0.202	16
		Time 3	0.180	0.119	16

Table 4b. ANOVA results of the 50 most highly activated voxels in the larger sphere while excluding outlier. Similar results were seen as compared to the analysis in Table 3b.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Run	0.011	1	0.011	0.488	0.496
Residual	0.329	15	0.022		
Strength	0.011	1	0.011	0.563	0.465
Residual	0.289	15	0.019		
Time	3.205	1	3.205	165.551	< .001
Residual	0.290	15	0.019		
Run ✻ Strength	6.387e −4	1	6.387e −4	0.026	0.873
Residual	0.365	15	0.024		
Run ✻ Time	0.036	1	0.036	4.702	0.047
Residual	0.115	15	0.008		
Strength ✻ Time	3.330e −5	1	3.330e −5	0.006	0.939
Residual	0.081	15	0.005		
Run ✻ Strength ✻ Time	0.034	1	0.034	2.862	0.111
Residual	0.176	15	0.012		

Note. Type III Sum of Squares

Table 5. Results for the Run x Time interactions from the 2x2x2 repeated measures ANOVAs for each individual brain area, with and without outlier. None of the interactions were significant.

Run x Time Interactions

ANOVA	SS	df	MS	F	p
FFA – All subj.	0.030	16	0.030	4.414	0.052
FFA – No outlier	0.029	15	0.029	3.980	0.065
PPA – All subj.	0.023	16	0.023	1.662	0.216
PPA – No outlier	0.030	15	0.030	2.222	0.157
LOC – All sub.	0.007	16	0.007	0.191	0.668
LOC – No outlier	0.023	15	0.023	0.759	0.397

DISCUSSION

This study overall suggests that VSL could possibly extend to category-specific areas in the brain such as the FFA, PPA, and LOC, but the differences in brain activity from VSL are more likely due to magnitude changes than temporal shifts. The current study provides no evidence of temporal shifts in activation due to VSL. Although there appears to be a trend in both the functional and behavioral data to suggest VSL has taken place, there was not a significant difference in subject behavioral responses on the memory tests. Other fMRI studies have also found neural differences during learning paradigms that do not always culminate in observable behavioral differences on subsequent explicit memory or performance tests (Turk-Browne, Scholl, Chun, & Johnson, 2008; McNealy, Mazziotta, & Dapretto, 2006; Landau, Schumacher, Garavan, Druzgal, & D' Esposito, 2004). Consequently, we could be observing visual statistical learning differences in brain activity that are not strong enough to cause changes in behavioral responses.

Interestingly, by removing the outlier subject that had conscious awareness of the transitional probability patterns, the Run x Time interaction became more significant in both the small sphere and highly activated voxel analyses ($p = 0.049$ and $p = 0.047$, respectively, compared to $p = 0.089$ with outlier retained), such that BOLD signals were higher for the first

item in the triplet during Run 3 and higher for the third item in the triplet in Run 1. Since removing the behavioral outlier had such a large impact on the overall interaction, this result could possibly suggest that the outlier had different brain activations due to the outlier's conscious awareness of the patterns.

One possible explanation for these magnitude difference results is that instead of observing differences in the timing of activation, we are seeing a pre-activation, per se, of the expected item. For example, perhaps the FFA shows increased activation to the non-preferred first item (object) in an OSF triplet as it begins to anticipate the presentation of an upcoming face in the third position of the triplet. This pre-activation could also possibly explain the lowered activation seen in response to the face in the triplet during Run 3 after VSL has taken place.

Additionally, although these magnitude changes were only observed with the explicit memory outlier removed, there are factors that could be masking this effect. One assumption we made during task design and analysis was that all three ROIs would experience the same statistical learning effects. By looking at the patterns of activity in Figure 6, it is clear that each area experienced VSL in different ways. For example, the FFA and PPA have some similarity in their timeline trends, but the LOC post-VSL activity for strong triplets is drastically different. Part of this discrepancy could be due to the ambiguity of the LOC ROI location. When selecting the center voxels of activation for each ROI, the LOC for each subject had a large amount of variability, while the FFA and PPA locations were fairly consistent.

Consequently, these timelines suggest that the three areas should be analyzed separately, as the pattern does not seem generalizable to all three. Individual ANOVA analyses of the three brain areas provided some evidence for these differences in VSL for each area. The Run x Time interaction for the FFA was very close to significant, but the Run x Time interactions of the PPA

and LOC were not (Table 5). Future research could focus more specifically on the FFA, as these results suggest that the FFA could be experiencing the largest effects of VSL.

Another assumption we made was that the third item in the triplets would experience the largest difference due to VSL since it was preceded by two items instead of one. Our timeline analyses were limited to activation of each brain area when its preferred category of item was in the third position of triplets. Future analyses could determine if there are timing or magnitude changes present for each category-specific area when its preferred item is in the second position of the triplet.

Another possible effect that could have contributed to the results of the study is individual differences in brain activation during statistical learning. For example, an event-related potential ERP study by Abla et al. (2008) suggests that there are individual differences in the ability to use statistical learning, positing that participants can possibly be grouped into "high", "middle", and "low" degrees of statistical learning based on their behavioral data, which can then be used to find differences in the neural signatures. However, we did not have a large enough sample size or significant behavioral data in order to separate subjects into various levels of statistical learning ability. These degrees of statistical learning could possibly explain the one subject that had explicit knowledge of the patterns; perhaps the patterns were solidified enough through VSL that the associations became strong enough for their explicit awareness.

Further statistical analyses could be completed on this data set to see if changes in the time course of activation are present in other areas of the brain, although we believe that other brain areas would show magnitude changes as well, instead of differences in timing. Possible follow up analyses could include areas that have been shown in previous research to exhibit changes in activation during statistical learning, such as the hippocampus and striatum.

Exploratory whole brain analysis could also be conducted to more broadly search for areas that are sensitive to VSL. Overall, the trends seen in the magnitude differences for the averaged areas, and specifically the FFA, suggest that further research is warranted.

ACKNOWLEDGEMENTS

We would like to thank Lauren Bandel, Joanne Murray, and Kerry Hartz for assistance with data collection and scanning. Additionally, we thank Cheng Lim for her help with counterbalancing and general consulting.

REFERENCES

Abla, D., Katahira K., Okanoya K. (2008). Online assessment of statistical learning by event related potentials. *Journal of Cognitive Neuroscience, 20(6),* 952-964.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language, 83,* 62-78.

Bays, B. C., Turk-Browne, N. B., Seitz, A. R. (2015). Dissociable behavioural outcomes of visual statistical learning. *Visual Cognition, 23,* 1072-1097.

Buxton, R. B., Wong E. C., & Frank L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine, 39:* 855–864.

Chen, Z., & Calhoun, V. (2018). Effect of spatial smoothing on task fMRI ICA and functional connectivity. *Frontiers in Neuroscience*, *12,* 15.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392,* 598-601.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology, 28,* 458-467.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41,* 1409-1422.

Huettel, S.A., Song, A. W., & McCarthy, G. (2004). *Function Magnetic Resonance Imaging*.

Sunderland, MA: Sinauer Associates, Inc.

Kanwisher, N. G., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311

Kanwisher, N., Tong, F., Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68,* B1-B11.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain Language, 127(1),* 46-54.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83,* B35-B42.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *Journal of Neuroscience, 26,* 7629–7639.

Penny, W., Friston, K., Ashburner, J., Stefan, K., & Nichols, T. (2006). *Statistical parametric mapping: The analysis of functional brain images*. Cambridge: Academic Press.

Saffran, J. R., Aslin R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *NeuroImage, 30,* 1088-1096.

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology, 22,* 1622–1627.

Schapiro, A., & Turk-Browne, N. B. (2015). Statistical learning. *Brain Mapping: An*

Encyclopedic Reference, 3, 501-506.

Tabelow, K., Piëch, V., Polzehl, J., & Voss, H. U. (2008) High-resolution fMRI: overcoming the signal-to-noise problem. *Journal of Neuroscience Methods, 178(2),* 357-365.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134,* 552–564.

Turk-Browne, N. B., Scholl B. J., Chun M. M., & Johnson M. K. (2008). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*. *21(10),* 1934-1945. doi:10.1162/jocn.2009.21131.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K, Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience, 30(33),* 11177-11187.

Yarkoni T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T.D. (2011). NeuroSynth: a new platform for large-scale automated synthesis of human functional neuroimaging data. *Front. Neuroinform. Conference Abstract: 4th INCF Congress of Neuroinformatics.*