Browse all Theses and Dissertations

Theses and Dissertations

2016

How configural is the Configural Superiority Effect? A neuroimaging investigation of emergent features in visual cortex

Olivia Michelle Fox
Wright State University

HOW CONFIGURAL IS THE CONFIGURAL SUPERIORITY EFFECT? A

NEUROIMAGING INVESTIGATION OF EMERGENT FEATURES IN VISUAL

CORTEX

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

By

OLIVIA MICHELLE FOX
B.A., Clemson University, 2011

2016
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

<u>January 3, 2017</u>

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY <u>Olivia Michelle Fox</u> ENTITLED <u>How Configural is the Configural Superiority Effect? A Neuroimaging Investigation of Emergent Features in Visual Cortex</u> BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF <u>Master of Science</u>.

_____
Kevin Bennett, Ph.D.
Thesis Director

_____
Scott Watamaniuk, Ph.D.
Graduate Program Director

_____
Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Committee on
Final Examination

_____
Kevin Bennett, Ph.D.

_____
Assaf Harel, Ph.D.

_____
Joseph Houpt, Ph.D.

_____
Jason Parker, Ph.D.

_____
Robert E. W. Fyffe, Ph.D.
Vice President for Research and
Dean of the Graduate School

ABSTRACT

Fox, Olivia Michelle. M.S., Department of Psychology, Wright State University, 2016. How Configural is the Configural Superiority Effect? A Neuroimaging Investigation of Emergent Features in Visual Cortex.

The perception of a visual stimulus is dependent not only upon local features, but also on the arrangement of those features. When stimulus features are perceptually well organized, a global configuration with a high degree of salience emerges from the interactions between these features, often referred to as emergent features. Emergent features can be demonstrated in the Configural Superiority Effect (CSE): presenting a stimulus within an organized context relative to its presentation in a disarranged one results in better performance. Prior neuroimaging work on the perception of emergent features regards the CSE as an "all or none" phenomenon, focusing on the contrast between configural and non-configural stimuli. However, it is still not clear how emergent features are processed between these two endpoints. The current study examined the extent to which behavioral and neuroimaging markers of emergent features are responsive to the degree of configurality in visual displays. Subjects were tasked with reporting the anomalous quadrant in a visual search task while being scanned. Degree of configurality was manipulated by incrementally varying the rotational angle of low-level features within the stimulus arrays. Behaviorally, we observed faster response times with increasing levels of configurality. These behavioral changes were accompanied by increases in response magnitude across multiple visual areas in occipito-temporal cortex, primarily early visual cortex and object-selective cortex. Our findings suggest that the neural correlates of emergent

features can be observed even in response to stimuli that are not fully configural, and demonstrate that configural information is already present at early stages of the visual hierarchy.

TABLE OF CONTENTS

LIST OF FIGURES

I. INTRODUCTION

One of the more well-known ideas to emerge from Gestalt psychology is that the whole is different than the sum of its parts (e.g., Koffka, 1935; Wagemans et al., 2012). A closely related phenomenon has been referred to as emergent features: the subjective perception of a visual stimulus is dependent not only upon the local features but also on the joint co-occurrence and arrangement of those features (e.g., Pomerantz et al., 1977). When the stimulus features are perceptually well-organized (i.e., form a cohesive structure) a global configuration with a high degree of perceptual salience emerges from the interactions between local features. When these emergent features correspond to task demands behavioral performance improves dramatically (e.g.,Pomerantz, 1986). Emergent features have been found to play a key role in visual display design (Bennett and Flach, 1992), as they provide powerful tools for decision-making by effectively leveraging the natural perceptual skills of human observers (Woods, 1991). If a display has been designed successfully the salient emergent features will represent meaningful properties and relationships within complex work domains, a notion critical in the ecological approach to display design (e.g., Bennett and Flach, 2011).

The perception of emergent features is often regarded as an "all or none" phenomenon, present only when the local features are strongly grouped to produce salient higher order global properties that are not contained within any of the constituent parts (Pomerantz and Portillo, 2011). The majority of work on emergent features largely

focuses on stimuli that are either configural (i.e., producing salient emergent features) or non-configural (not producing emergent features). Configural stimuli are comprised of local features intentionally arranged by the experimenter to produce the emergence of a global percept (i.e., an emergent feature). For example, many of the studies use experimental stimuli that are arranged to form either perfect bilateral symmetry or parallelism (e.g., Pomerantz and Garner, 1973; Pomerantz et al., 1977; Pomerantz and Schwaitzberg, 1975). And while the identity of the local features are the same for both configural and non-configural stimuli, those of non-configural stimuli have been intentionally arranged to be void of such emerging properties, or arranged to form an extreme case of poor configuration (Kubilius et al., 2011; Pomerantz et al., 1977). In other words, configural stimuli are optimally arranged to produce high-level, salient properties emerging from the interactions of the elements in the stimulus, whereas the non-configural stimuli are designed to be void of such emergent features, representing an extreme case of poor configuration (Pomerantz et al., 1977).

Critically, however, this dichotomous approach to emergent features overlooks the possibility that stimuli between these ends of the continuum may vary in their degree of configurality. Thus, a possibility that has not explicitly been considered is that emergent features vary in a quantitative rather than qualitative manner. This would imply that configurality might be achieved even in conditions that do not entirely satisfy the geometric properties characteristic of emergent features; for example, emergent features may be perceptually accessible in configurations that approximate but are not in themselves symmetric and parallel orientations (for a similar argument in the face perception literature, see Schwaninger and Mast, 2005). Stated alternatively, will

2

observers be capable of tolerating some degree of imperfection in the arrangement of the local features, and if so, to what extent? Accordingly, will they still retain perceptual access to the emergent features that aid task performance?

The current neuroimaging study aims to address these questions by examining how far perceptual stimuli can stray from what is usually considered a configuration before their emergent features break down. Such a break down may be reflected as a decrease in behavioral performance, as well as a decrease in the magnitude of neural activity associated with configural processing. To investigate this possibility we adapted the stimuli set used to demonstrate *the configural superiority effect* (CSE, Pomerantz et al., 1977). A set of stimuli were developed that continuously and parametrically varied the deviation of their constituent features from perfect symmetric and parallel orientations. We chose to focus on symmetry and parallelism, as they are two prime examples of emergent features often discussed in the context of the CSE (e.g., Pomerantz and Cragin, 2015; Mersch, 2014; Pomerantz and Garner, 1973; Pomerantz et al., 1977). This allowed us to explore the extent to which emergent features are continuous, and correspondingly, how behavioral outcomes (response times, accuracy) are impacted by degrees of configurality (e.g., continuous degradation or a dichotomous break?).

We further used neuroimaging to help us determine whether emergent features are continuous. We sought to both answer an open question in the literature regarding which areas in ventral visual pathway are correlates of configural processing, and to assess the extent to which the neural activity in those candidate regions is impacted by gradual changes in configurality. We measured the neural responses to the above set of stimuli, focusing on regions along the ventral visual pathway that have been previously shown to

3

support configural processing (e.g., Chechlacz et al., 2015; Ward and Chun, 2015). This enabled us to determine the extent to which the CSE is manifest along the visual hierarchy; specifically, we investigated whether CSE reflects activation across multiple levels of the visual hierarchy (e.g., Altmann et al., 2003; Ban et al., 2006; Chechlacz et al., 2015), or whether it is supported by specific high-level areas in occipito-temporal cortex (OTC) (e.g.,Kubilius et al., 2011; Lerner et al., 2001).

While early visual areas (e.g., V1, V2, V4) are known to be primarily responsible for processing simple local features, these regions have also been implicated in global shape processing, suggesting a more varied and complex role for early visual cortex (EVC) than traditionally thought (e.g.,Kourtzi et al., 2003). Neuroimaging work focused specifically on the perception of global shapes emerging from local elements arranged at various orientations shows both early visual areas and higher level visual areas are involved in the processing of emergent features (e.g. Altmann et al., 2003). In contrast, other works have shown configural effects primarily in higher areas in the visual hierarchy such as object-selective lateral occipital complex (LOC: Malach et al., 1995), and were not able to demonstrate configural effects in EVC. For example, Kubilius et al. (2011) a CSE in LOC in the form of higher decoding accuracy to whole stimuli relative to parts, whereas in EVC decoding accuracy was higher to parts than to wholes.

In the present study, we utilize knowledge of the functional properties of visual areas in order to infer how emergent features are processed by examining both behavioral performance and activity in regions of interest (ROIs). We ask whether configural effects will be seen in only high-level, or in both high-level and lower-level areas of the visual

4

hierarchy, or whether one might observe gradual changes along the visual hierarchy as a function of the degree of deviation from perfect configurality.

II. METHOD

Participants

Eight Wright State University students participated (6 female, age range: 18-29, *M*=24.5). All participants were right-handed, had normal or normal-corrected vision, normal color perception, the ability to read and write in English, and the ability to complete a MRI safely. All participants signed an informed written consent according to the institutional review board of Wright State University and received due compensation for their participation.

Stimuli

The stimuli were comprised of 144 unique stimulus arrays. 128 of these stimulus arrays consisted of eight parentheses arranged to create four pairs aligned in a 2x2 matrix (*Figure 1*). Each stimulus array contained a target; 3 of the parenthesis pairs were identical while the target pair was distinct with its leftmost parenthesis facing the opposite direction of the leftmost parentheses of the other 3 pairs. Each pair subtended a visual angle of approximately 1.56° vertically and 1.79° horizontally. The overall array subtended a visual angle of approximately 7.15° vertically by 7.63° horizontally. The orientation angle of the rightmost parenthesis of each pair was manipulated to result in stimuli at 16 different rotational angles beginning with zero degrees rotation and continuing in six-degree increments (0°, 6°, 12°, 18°, 24°, 30°, 36°, 42°, 48°, 54°, 60°, 66°, 72°, 78°, 84°, & 90°). The target location and target direction were also manipulated
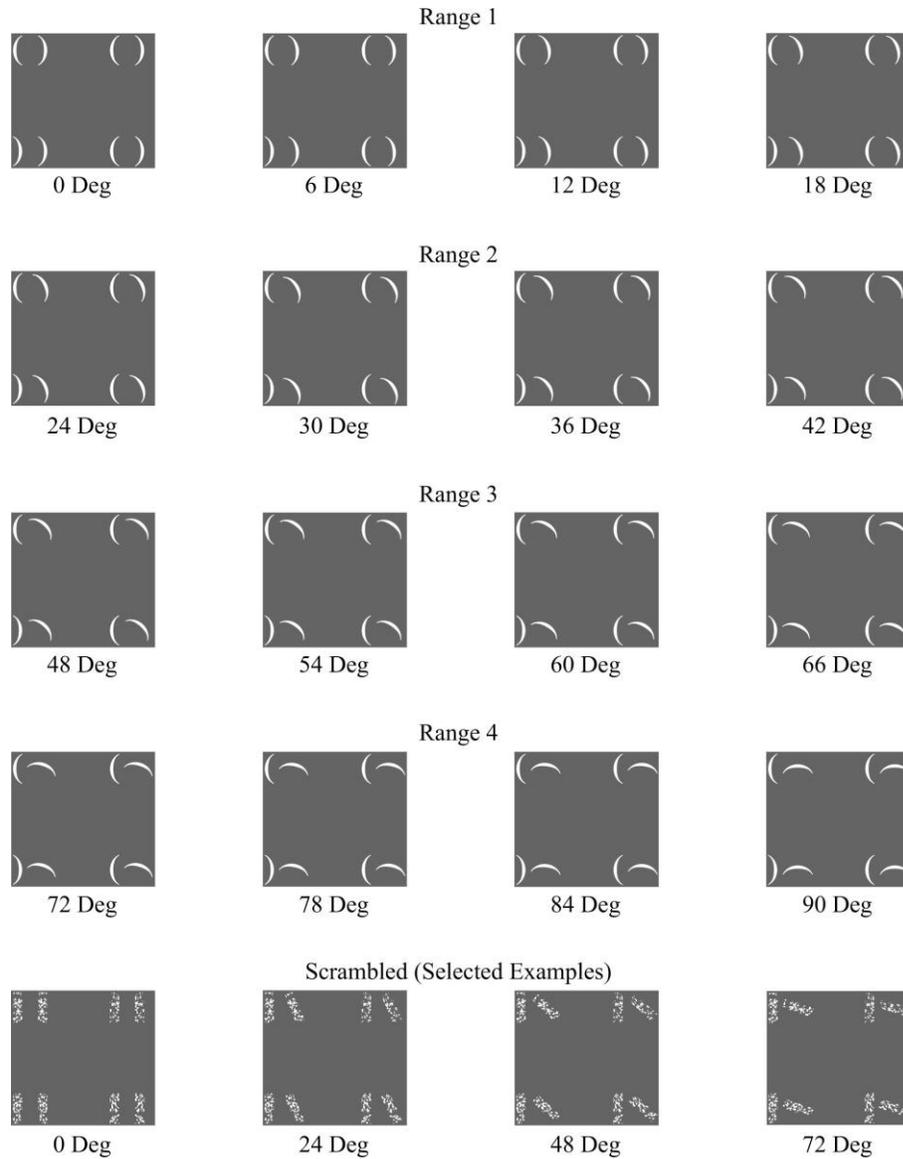
Figure 1. A representative set of the full range of Configural stimuli separated by Configurality Range, with a selected set of Scrambled stimuli.

such that on any given trial, the target could appear in any of the four quadrants, and be facing either left or right. This culminated in 128 unique stimulus arrays (16 angles X 4 quadrant locations x 2 target directions). In addition, 16 control stimulus arrays were used, which retained the 2x2 matrix structure, but using textural rectangular elements

rather than the parentheses (*Figure 1*). Specifically, these "scrambled" arrays were made up of one parenthesis pair per quadrant with the rectangular area surrounding each parenthesis (2.1 mm x 7.0 mm) having been divided into 120 squares (0.35 mm^2) and then randomly rearranged while retaining each square's original orientation. There were 16 versions of the Scrambled condition, one with the rightmost parenthesis of each pair oriented to each of the 16 rotational angles.

fMRI Procedure and Experimental Design

An event-related fMRI design was used for the study, with participants performing an anomalous quadrant discrimination task on the stimulus arrays. Stimuli were presented in a pseudo-randomized fashion, so that each sequence of 18 trials contained all 16 rotational angles stimuli, with an additional Scrambled stimulus array, and a null event (blank screen). The target location and target facing direction randomly varied across the sequences of stimuli, spanning the full range of the stimulus set. The stimuli were presented on the screen until a response was made (participants were instructed to maintain fixation throughout the length of the experiment). Immediately following the participant's response a fixation cross screen of variable duration was presented (duration of the interstimulus interval was determined by multiplying the latency of the response time by two and adding a randomly chosen value ranging from +0.5 s to -0.5 s). The scan sessions comprised of two experimental runs (10.5 min each), separated by a 6 min high-resolution anatomical scan, and followed by a functional localizer run (7.7 min). Prior to entering the scanner, the participants received training, familiarizing them with the experimental task. The participants indicated the location of the target (i.e. the anomalous quadrant in each array) using response pads, with each

quadrant corresponding to a unique button. Participants had a response pad in each hand with two buttons; the top buttons corresponded to the upper quadrants, and the bottom buttons corresponded to the lower quadrants. Participants indicated the left half of the display using the response pad in their left hand while the response pad in the right hand indicated the right half of the display. Since the arrays of the Scrambled condition did not contain a target, participants could respond with any button on the response pad in order to advance to the next trial.

Category Localizer Experiment

In addition to the main experiment, an independent functional localizer experiment was conducted in order to identify category-selective regions in visual cortex. This block-designed fMRI experiment included 4 stimulus conditions (faces, houses, objects, and simple textures; see Figure 2 for stimulus examples). Each condition was repeated 7 times in pseudorandom order. Blocks consisted of 9 images of the same category, each displayed for 800 ms followed by a 200-ms blank screen (a total of 9 s, interleaved with 6-s fixation periods). All stimuli were grayscale photographs of 300 by 300 pixels each, subtending a visual angle roughly equal to that of the experimental arrays. The task was a standard one-back memory task, with an image repetition occurring once or twice in each block.
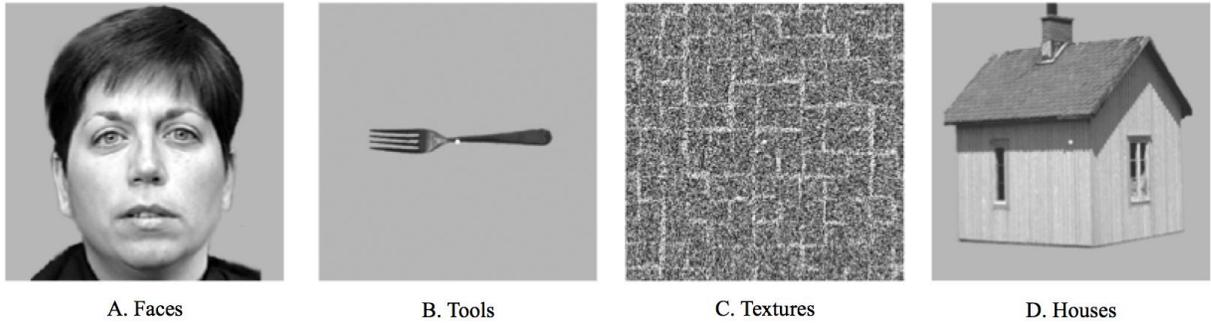
A. Faces   B. Tools   C. Textures   D. Houses

Figure 2. A subset of the stimuli used in the category localizer experiment. **(A)** An example from the faces category. **(B)** An example from the tools category. **(C)** An example from the textures category. **(D)** An example from the houses category.

Magnetic Resonance Imaging Setup

A 1.5 Tesla MRI scanner (General Electric Excite HDX; General Electric, Milwaukee, Wisconsin) with an eight-channel head coil was used for all acquisitions. A high resolution T1-weighted anatomical scan was acquired for each participant using a 3D Magnetization-Prepared Rapid-Acquisition-Gradient-Echo (MPRAGE) sequence (512x512 matrix, 120 slices, 1mm x 1mm x 1mm voxel size, TR/TE = 500/15ms, flip angle = 15°). Three functional MRI acquisitions were also acquired for each participant using a Gradient-Recalled-Echo (GRE) sequence (64x64 matrix, 24 slices, 4.5mm x 4.5mm x 5mm voxel size, 1mm slice gap, TR/TE = 2000/10ms, and flip angle = 90°).

fMRI Data Preprocessing

We used the BrainVoyager software package (Brain Innovation, Maastricht, The Netherlands) to analyze the fMRI data. The first image was discarded from each functional scan prior to analysis. Next, the scans were normalized into Talairach space. Preprocessing steps performed on the functional images included 3D motion correction, slice scan time correction, linear trend removal, and mean intensity adjustment.

Statistical Analysis

     Behavior. For all analyses, the 16 levels of configurality were condensed into four Configurality Ranges in order to increase the number of repetitions: Range 1 (0°, 6°, 12°, 18°), Range 2 (24°, 30°, 36°, 42°), Range 3 (48°, 54°, 60°, 66°) and Range 4 (72°, 78°, 84°, 90°). Additionally, a rate of change measure was derived on an individual basis by calculating the slope of each participant's RT function across the Configurality Ranges.

     Neuroimaging. For each subject, after the time courses of the 2 experimental scans were transformed into Talairach space and preprocessed (see fMRI data preprocessing and analysis), they were z-normalized and concatenated. For the ROI time course analysis, the data were deconvolved using the deconvolution analysis for rapid-event-related paradigms that consists of a general linear model analysis in BrainVoyager software package (Brain Innovation, Maastricht, The Netherlands) in order to extract the estimated hemodynamic response in each voxel for each condition. The analysis was done separately for each subject on a voxel-by-voxel basis. A rate of change measure was also calculated for each participant within each ROI. The rate of change was determined by calculating the slope of each participant's response magnitude function across the Configurality Ranges, within each ROI.

ROI Selection and Analysis

     ROIs were identified in each subject separately based on the category localizer experiment as described above. They were defined on the basis of a minimum cluster size of 6 contiguous functional voxels that exhibited selective activations in response to a specific category ($p < 0.01$). EVC ROIs were defined as those contiguous voxels that

responded preferentially to textures relative to faces and objects. LOC ROIs were defined as those contiguous voxels that responded preferentially to objects as compared to textures. FFA ROIs were defined as those regions showing preferential activation for faces relative to houses. PPA ROIs were defined as those regions showing preferential activation for houses relative to faces and textures. We applied a deconvolution analysis using the BrainVoyager software package (Brain Innovation, Maastricht, the Netherlands) to the time course of each voxel within the ROIs to extract an estimated hemodynamic response. Next, the estimated responses were averaged across five time points starting two seconds after stimulus onset and surrounding the peak of activation. The estimated response values were then averaged across hemisphere, as no significant differences were found between the ROI time courses for the left and right hemispheres. The resulting values, averaged across hemisphere, were then used in an analysis of variance (ANOVA). ANOVAs were also conducted using the peak activations of the estimated hemodynamic responses for each condition. However, while descriptively in the same direction as our time-point averaged response values, these analyses yielded no significant results.

III. RESULTS

Behavior.

To examine the effect of configurality on behavioral performance, we conducted a within subjects repeated measures ANOVA with Configurality Range (5 levels: Range 1-Range 4, scrambled) and Run (2 levels) as independent variables on participants' mean response times. Trials in which participants provided inaccurate responses (2.37% of Range 1 trials, 1.41% of Range 2 trials, 2.39% of Range 3 trials and 2.63% of Range 4 trials), and trials identified as outliers (15 out of 1900 trials, <0.01%) were removed from the dataset using the method described by Lovie (1986, p. 55-56): $T1 = (x_{(n)} - x) / s$, where $x_{(n)}$ is a particular observation (one of n observations), x is the mean of those observations, and s is the standard deviation of those observations. Response times were averaged across both location (quadrant) and target direction (facing left or right) as initial analyses did not reveal significant effects of these factors ($F(3,21) < 1$ and $F(1,7) = 3.04$, MSE = .370, $p > .12$, respectively).

Significant main effects were found for both Configurality Range ($F(4,28) = 32.36$, MSE = .17, $p < .001$) and Run ($F(1,7) = 41.82$, MSE = .012, $p < .001$). Overall, participants responded faster in the second run ($M = 1.61$, SE = 0.11) relative to the first run ($M = 1.77$, SE = 0.11). However, there was no significant interaction of Run and Configurality Range ($F(4,28) = 1.30$, MSE = .008, $p > .29$), meaning that performance in the experimental conditions was not differentially impacted across the runs.

Consequently, response latencies were averaged across runs for subsequent comparisons.

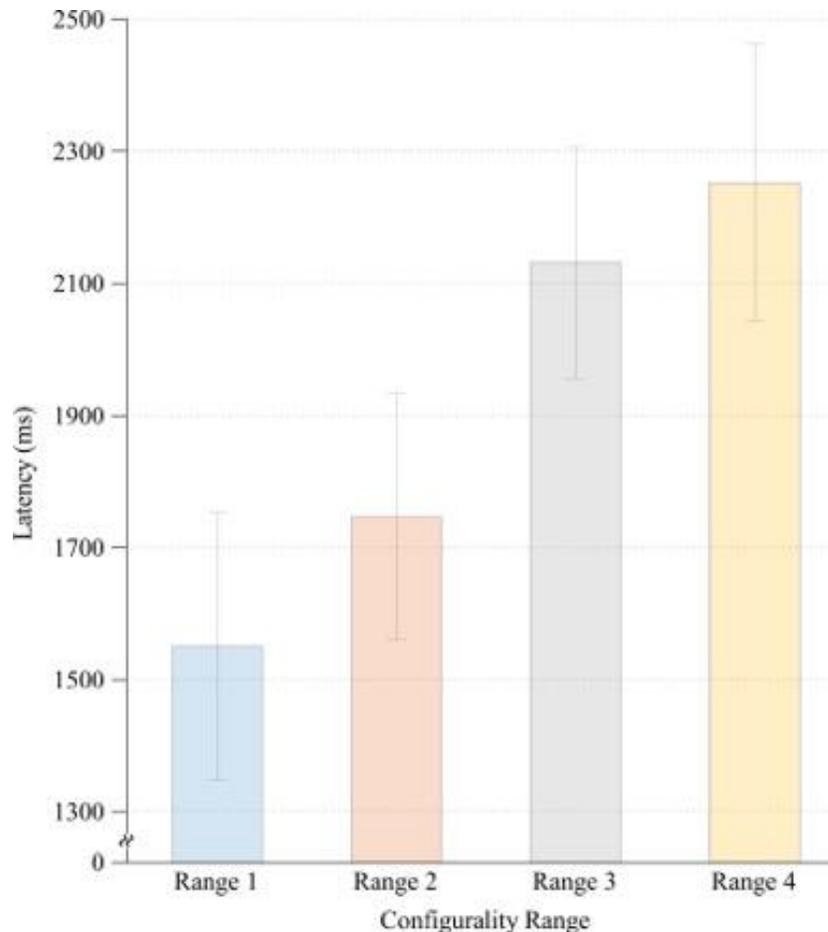Mean response times for each range are presented in Figure 3.

Figure 3. Mean response times for each Configurality Range
with error bars denoting standard error.

To further explore the effect of Configurality Range, we performed a series of

planned comparisons on each successive pairs of our conditions. We found that in

general, the smaller the deviation from configurality the faster the response latencies

were. Thus, response latencies in Range 1 were significantly lower than those in Range 2

($F(1,7) = 14.27$, MSE $= .022$, $p < .007$), and response latencies in Range 2 were

significantly lower than those in Range 3 ($F(1,7) = 12.12$, MSE $= .231$, $p < .01$). Average

response latencies in Range 3 were faster than in Range 4, although this effect was not

14

significant (F(1,7) = 4.72, MSE = .024, p < .07). Finally, the response latencies in the

Scrambled condition were significantly faster than in each of the four Configurality

Ranges (all p's < .003, MSE = .175 - .263).

Neuroimaging

In order to determine how early along the ventral visual pathway configurality

emerges, we first examined how our parametric manipulation of the relations between

visual features impacted the magnitude of activation in early visual areas (EVC) and

higher-level visual areas (LOC), two areas of interest in prior neuroimaging examinations

of configural processing (our other two ROIs were intended to serve as control regions;

see below). We started by asking whether the configural stimuli would produce unique

activation patterns relative to any type of general visual stimulation by comparing the

differences in activation between the configural stimuli (averaged across all levels of

configurality) and arrays of task-irrelevant, scrambled elements. We conducted a two-

way ANOVA with ROI (EVC/LOC) and Stimulus Type (Configural/Scrambled) as our

independent variables (no significant effects were found for Run, Hemisphere or their

interactions, all p's > .25, MSE < .001). We found significant main effects of ROI and

Stimulus Type (F(1,7) = 9.23, MSE= 0.001, p < 0.02; F(1,7) = 9.65, MSE = 0.005, p <

0.02, respectively), and critically, also an interaction between ROI and Stimulus Type,

F(1,21) = 8.92, MSE < .00005, p < 0.02). Post hoc analyses revealed that the difference

between configural and scrambled stimuli was more pronounced in EVC (t(7) = -3.58, p

< .01) than in LOC (t(7) = -2.38, p < .05). However, and somewhat counter-intuitively,

the Scrambled condition elicited a higher response than the configural stimuli across all

ROIs (Figure 4). This may be accounted for by the fact that while the scrambled and

experimental stimuli were originally designed to have identical physical size (in terms of

pixel count; for examples of stimuli, see Figure 1), it might be that the density and

distribution of these pixels differ across the different types of stimuli, leading to this
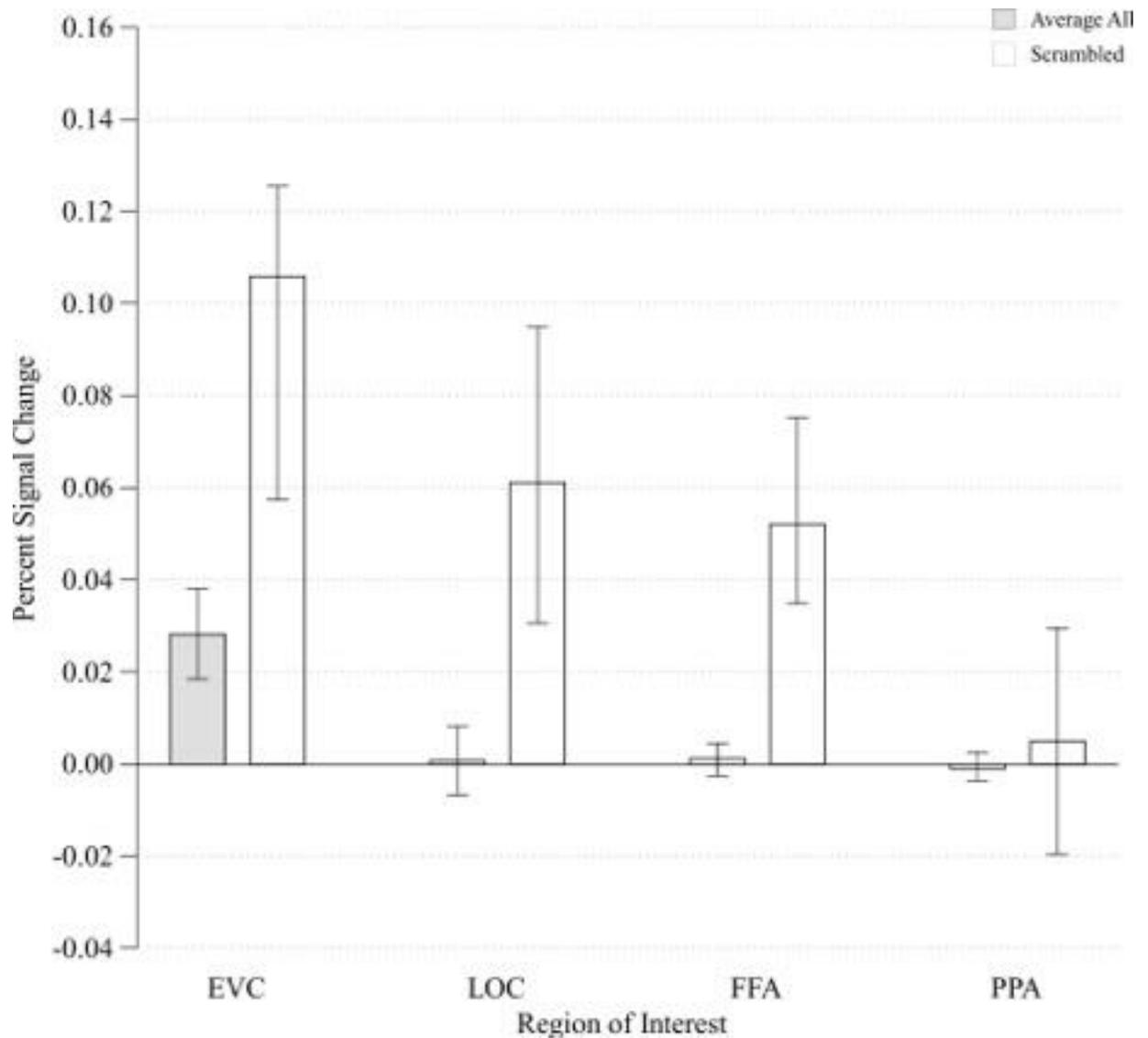
unexpected effect.

Figure 4. Mean percent signal change of all Configurality Ranges compared to the percent signal change of the Scrambled condition within each ROI. Error bars represent standard error.

We next examined the extent to which finer differences within the configural stimuli are manifest in the response magnitude of LOC and EVC. We conducted a three-way ANOVA in each ROI with Run (first run/second run), Hemisphere (right/left) and Configurality Range (Range 1- Range 4) as independent variables. We excluded the Scrambled condition from the analysis given its noticeable difference in response magnitude from the other configural stimuli.

LOC. The three-way ANOVA showed no main effects of either Run or Hemisphere ($F(1,3) = 2.08$, MSE $= 0.00$, $p > .24$; $F(1,3) < 1.00$, respectively). No significant interactions between any of the independent variables were noted (all p's $> .22$). We therefore averaged our data across these two factors, and performed a one-way ANOVA with Configurality Range as the independent variable. This ANOVA revealed a significant main effect of Configurality Range ($F(3,21) = 5.33$, MSE $= .001$, $p < .007$). It is important to note that given the small sample size, these results should be considered with caution. However, a post-hoc power analysis revealed an observed power of 0.88. Post-hoc comparisons of each consecutive pair of configurality ranges, revealed a significant difference in response magnitude between Range 2 and Range 3 ($t(7) = 1.913$, $p < .05$), with no significant differences between the upper two and the lower two levels of configurality (all p's $> .45$) (Figure 5). A linear trend analysis across all levels of configurality revealed a significant linear trend ($F(1,7) = 9.71$, $p < .02$).
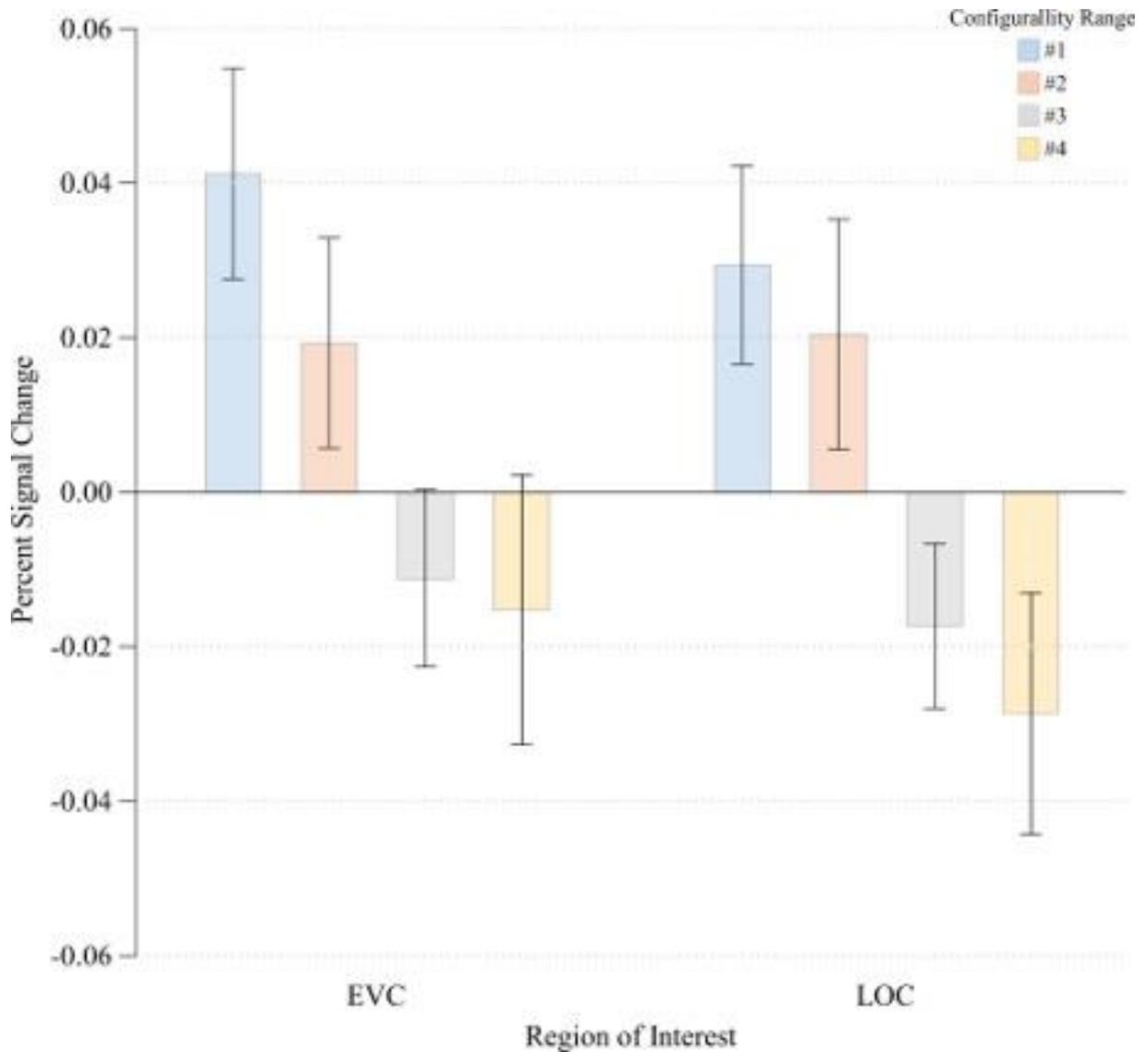
Figure 5. Percent signal change at each Configurality Range within EVC and LOC, with error bars representing standard error.

EVC. The three-way ANOVA showed no main effects of either Run or Hemisphere ($F_{(1,5)} < 1.00$; $F_{(1,5)} < 1.00$, respectively). No significant interactions between any of the independent variables were noted (all p's > .12). We therefore averaged our data across these two factors, and performed a one-way ANOVA with Configurality Range as the independent variable. A significant main effect of

Configurality Range ($F(3,21) = 3.36$, MSE $= .002$, $p < .04$) was obtained. Despite the small sample size, a post-hoc power analysis yielded an observed power of 0.68. Post-hoc pairwise comparisons following the main effect of Configurality Range showed that Range 1 stimuli evoked a significantly greater response than Range 2 ($t(7) = 2.383$, $p <$ .05). There were no other significant differences between each two successive levels of configurality (all $p > .23$). A linear trend analysis across the four levels of configurality did not reveal a significant linear trend ($F(1,7) = 4.37$, $p < .08$).Thus, while EVC was sensitive to changes in configurality, it seems to respond in a different manner than LOC.

Other Regions of Interest. To verify that the main effect of configurality was due to our experimental manipulation rather than reflecting a general effect, such as global attention, we examined the extent to which configurality modulates the magnitude of response of two additional high-level visual regions: the parahippocampal place area (PPA) and the fusiform face area (FFA). We wanted to verify that our findings are not simply due to the fact that participants were more engaged, or more attentive to the most configural stimuli. Given that in addition to their category selectivity, FFA and PPA are susceptible to manipulation of selective attention (O'Craven et al., 1999), we should expect to find a general increase in their response magnitude if indeed our participants were more engaged with the configural stimuli. Furthermore, while FFA has been associated with configural processing, its response has been shown to be limited to face stimuli (Yovel and Kanwisher, 2004), and is therefore an interesting control region. We first found that activation in response to the Scrambled condition was significantly higher relative to the Configural condition in the FFA ($t(7) = -3.08$, $p < .02$). Response magnitude in PPA was not sensitive to the difference between the two conditions ($t(7) = -$

19

0.31, $p > .75$). We then assessed the effects of Configurality Range on both ROIs (excluding Scrambled, and averaging across runs and hemispheres) by conducting a one-way ANOVA with Configurality Range as an independent factor for each ROI (Figure 6). Critically, neither FFA nor PPA showed a significant effect of our conditions ($F(3,21) <$ 1.00 for both ROIs), making the possibility of general attentional effect unlikely.
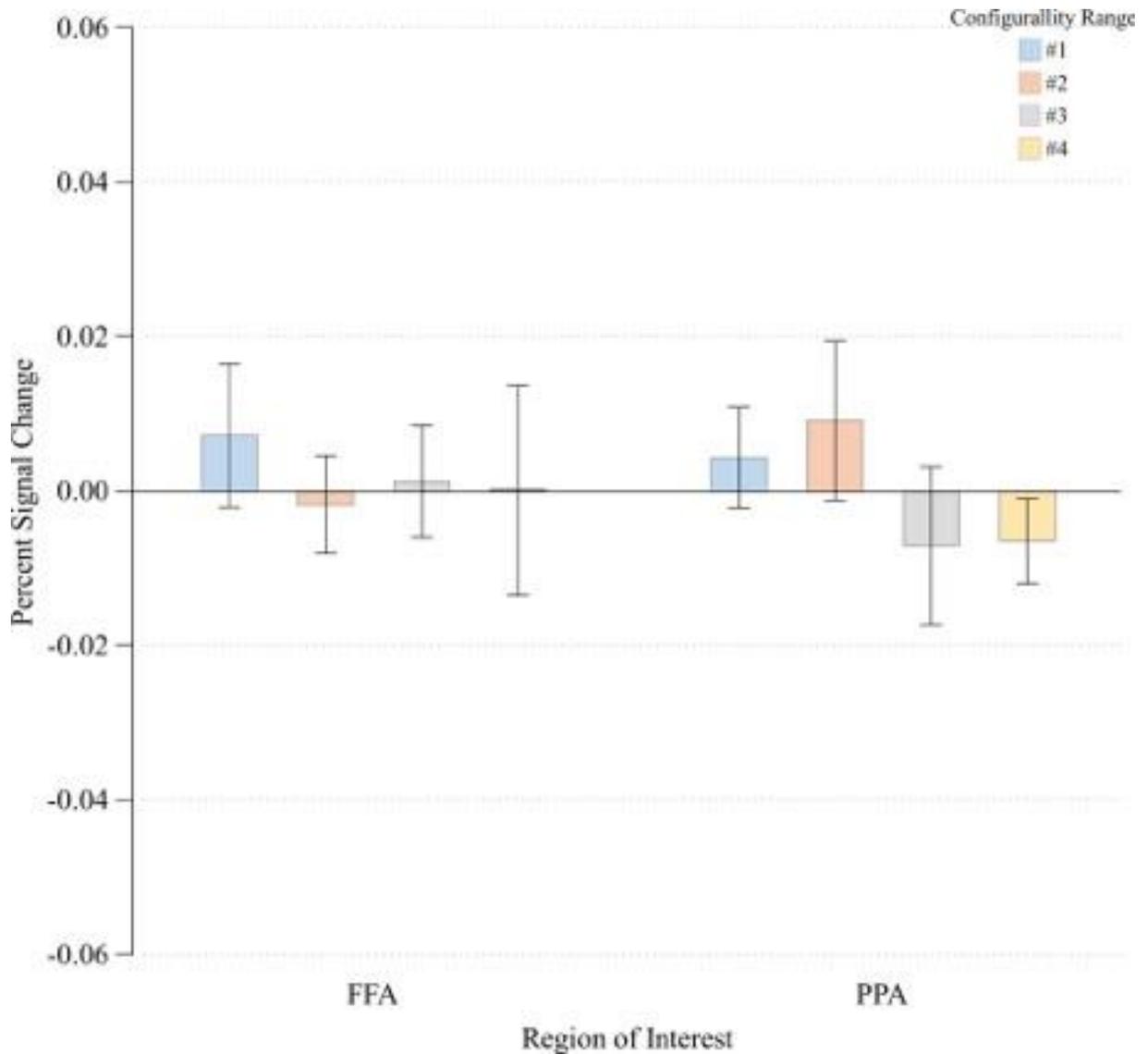

Figure 6. Percent signal change at each Configurality Range in FFA and PPA. Error bars denote standard error.

20

Correlations with Behavior

To directly assess how the neural measures of configurality reported above are related to the behavioral performance measures (which, admittedly were more continuous in nature), we correlated the rate of change in reaction times as a function of configurality range with the rate of change in response magnitude as a function of configurality range (calculating the slope of each variable for each individual; see Methods for details). We found a significant negative correlation between the rate of change in reaction times and the rate of change in response magnitude in both EVC ($r = -.80$; $p < .02$) and LOC ($r = -.81$; $p < .01$), indicating that the change in the speed of response to the varying degrees of configurality is tightly linked to the increase in magnitude as a function of configurality. In contrast, FFA and PPA showed no significant correlation between the behavioral change measure and the neural change measure (FFA: $r = -.34$, $p > .40$; PPA: $r = .36$, $p > .37$).

IV. DISCUSSION

Emergent features are global perceptually salient properties that arise from interactions among local stimulus elements but are not visible within any one element alone (e.g., Pomerantz, 1986). The goal of the present study was to examine how flexible emergent features are, that is, to determine the extent to which the perception of emergent features is tolerant to changes in the arrangement of their constituent local elements. We manipulated the arrangement of local elements within stimulus arrays, and assessed the extent to which the Configural Superiority Effect (CSE), an index of stimulus configurality (Pomerantz and Cragin, 2015), can be obtained across a range of deviations from an ideal configuration. We used behavioral performance to establish whether increasing configurality produces faster and more accurate responses; we used neuroimaging to determine where increasing configurality produces increased activation along the ventral visual pathway.

Behaviorally, we found a clear relationship between response times and degree of stimulus configurality. The closer the stimulus was to the ideal configuration, the faster the response latencies were. Neurally, we also found a clear relationship between stimulus configurality and response magnitude in two key visual areas (early visual cortex, EVC, and Lateral Occipital Complex, LOC). Further, the effect of configural information was already evident in early stages of processing in the visual hierarchy.

Taken together, our findings suggest that the perception of emergent features is robust to variations in stimulus configurality.

We hypothesized that configurality is not an all or none phenomenon, and therefore, that variations in symmetry and parallelism would impact the strength of the CSE. In line with our predictions, we found that the CSE weakened as the stimuli departed from what is usually considered to be the optimal configuration (Range 1 stimuli). This decrease was evident until Range 3, after which further deviations no longer resulted in additional decrements to performance. This monotonic relationship between performance and stimulus organization demonstrates that configurality effects can be achieved for stimuli other than the prototypical configural stimuli (e.g., perfectly symmetrical stimuli, as conceptualized in prior studies) and are perceived in a continuous fashion. In other words, emergent features are still useful for improving task performance even in "suboptimal" forms; that is, in contexts that do not entirely satisfy the geometric properties that usually characterize emergent features. Importantly, we were able to establish when "sub-optimal" becomes "non-optimal", reflected in the plateauing of the RT function at the third and fourth Range levels.

The current study is not the first to suggest that traditional Gestalt properties possess a more continuous nature. For example, Pomerantz and Schwaitzberg (1975) sought to discover the limits of proximity on configural grouping by gradually increasing the separation of individual elements in stimulus arrays. They found that performance declined monotonically as the stimulus elements were more widely spaced, although this effect of proximity was eventually restricted in range (<4° of visual angle). Our findings are similar to those of Pomerantz and Schwaitzberg (1975) who also found a similar

monotonic relationship, as well as an established point at which no further advantages were conferred by the stimulus configural structure. Moreover, the current study extends Pomerantz and Schwaitzberg's findings to additional stimulus properties other than proximity, namely, parallelism and bilateral symmetry. More generally, one implication of the current findings is that other seemingly binary principles of object perception may prove to be more flexible than they are often considered. One potential example might be relatability, a criterion used to solve the grouping problem (Kellman and Shipley, 1991; for a review see van Lier and Gerbino, 2015). This criterion predicts a-modal completions by a smooth curve when linear extensions would meet behind the occluding surface at angles of 90° or larger, but not at smaller angles (van Lier and Gerbino, 2015) While this implies a binary rule, an outstanding question would be to what extent the percept of completion is affected by a continuous manipulation of the angle as it progresses farther away from 90°.

It should be noted that the CSE is traditionally defined based on a comparison between display elements in isolation and the same display elements within a context, and thus, technically speaking, we have not measured the CSE here. However, the chosen stimuli of parenthesis pairs we have used in the current study have been repeatedly shown to elicit strong CSEs when contrasted with isolated parentheses (Mersch, 2014; Pomerantz and Garner, 1973; Pomerantz et al., 1977), thereby allowing us to exclude an "isolated" condition, and focus instead on the range of configural variations.

Having established the non-dichotomous nature of CSE using behavioral performance, we also asked whether this effect could also be observed at the neural level. We found that configurality plays a significant role in driving the response of both object-

selective ('high') and early ('low') visual cortex, as reflected in a main effect of Configurality Range in LOC and EVC, respectively. However, while configurality effects were found in both EVC and LOC, the pattern of information usage was not entirely uniform across the ROIs. LOC responded with equal magnitude to both Range 1 and Range 2 stimuli. In other words, LOC's response is tolerant to at least some deviation from the 'prototypical' configuration (although LOC did not differentiate among the other two ranges). In contrast, the response of EVC was more selective, sensitive only to the most configural stimuli (Range 1), as its response was highest to that condition and did not further distinguish among the other configurations. Further resonating this pattern of differences between the two regions, we found that response magnitude in LOC scaled linearly with an increase in configurality range, while in EVC such a linear trend was not observed as significant. The lack of a linear trend in EVC adds to our conclusion that high-level, configural effects are driving the observed response in this low-level area. If the results were due to a low-level stimulus quality such as the orientation, we would expect linear changes in the response magnitudes.

These differential patterns of information usage in EVC and LOC might suggest that visual processing of configural information varies across different regions in the ventral visual pathway. One potential interpretation for this functional difference is that the degree of tolerance to deviations from configurality might be determined as a function of location along the visual pathway. Thus, farther upstream, EVC is very specific and responds only to the most configural stimuli (Range 1: 0º - 18º of rotation). In contrast, farther downstream LOC responds equally to all stimuli across Range 1 & Range 2 (0º - 42º of rotation). This increase in tolerance to variations in configurality

parallels similar increases in tolerance observed with other object dimensions along the ventral visual pathway (for a review, see DiCarlo et al., 2012).

Interestingly, in spite of this increase in representational complexity, the current study shows that one can observe a difference in activity based on the perceptual organization of the features as early as EVC. This is in contrast to prior works showing that EVC (primarily V1) activity does not capture the difference between local stimulus features and the perception of global shape they produce, whereas LOC supports global processes that go beyond the processing of isolated features (Hasson et al., 2001; Lerner et al., 2001). Under this view, EVC would not be anticipated to show any difference in its response across the range of configural stimulus arrays, whereas LOC would. Consistent with this prediction, a recent fMRI study of the CSE (Kubilius et al., 2011), comparing the processing of contextually configured stimuli ('wholes') relative to similar stimuli that did not contain configural context ('parts') found that LOC showed a higher classification performance for wholes versus parts, whereas EVC (V1, V2, V3) showed an opposite pattern, with lower classification to wholes versus parts. Based on these results, the authors concluded that the CSE manifests exclusively in higher visual areas, not in 'low-level' visual areas.

While seemingly at odds with the current results, several methodological differences make a direct comparison between the two studies difficult. For instance, the experimental design of the current study used a univariate analysis of the response magnitude rather than a multi-voxel pattern analysis (see Ostwald et al., 2008 for a discussion of the latter point). Interestingly, while the authors also report response magnitude data, they did not find any significant difference between the whole and parts

conditions, either in EVC or in LOC (Kubilius et al., 2011, supplemental material). Furthermore, whereas Kubilius and colleagues contrasted display elements in isolation with the same display elements within a context, we did not use an "Isolated" condition, but rather employed a range of context-bound configural variations (see above).

These differences between the prior work and the current study are suggestive of differing goals. The current study focused on establishing how gradually varying the relationship between stimulus elements impacts visual cortex, rather than contrasting parts and wholes in 'low' and 'high' level visual cortex. In other words, instead of assuming a clear-cut separation between 'low-level' parts and 'high-level' wholes, we sought to investigate the nature of the transition between configural "wholes" and non-configural "parts". Asking how early in the visual processing stream we can find sensitivity to subtle changes in configuration leads us to conclude that emergent features (or "wholeness", to maintain the analogy) are not binary, and critically, that the perceptual processing of such features is supported by the activity of multiple visual regions. Arguably, the joint activity of these visual regions – each differentially representing various aspects of the configural range – culminates in the characterization of the full range of the stimulus arrays.

The current idea that emergent features are neurally represented by the conjoint activity of multiple regions along the ventral visual pathway is in line with prior neuroimaging works that pointed to the role of both early and higher visual cortex in utilizing emergent features such as collinearity and orientation for the perception of global shapes (Altmann et al., 2003; Kapadia et al., 1995; Kourtzi et al., 2003; Ostwald et al., 2008; Sigman and Gilbert, 2000), and support a top-down interactive view of the

ventral visual pathway (Croner and Albright, 1999; Kravitz et al., 2013; Lamme and Roelfsema, 2000). Under this view, the ventral visual pathway is a recurrent and highly interactive occipitotemporal network that bridges early visual areas and anterior temporal lobe (ATL) along multiple routes through which visual information is processed. According to one specific conceptualization, this occipitotemporal network is hypothesized to be involved in the processing of 'stimulus quality', that is, the formation of specific representations or associations involving stable aspects of visual information (Kravitz et al., 2013).

The term 'stimulus quality' is particularly relevant to the current discussion, as this term aims to capture available information at its broadest sense, referring to the processing of both perceptual dimensions and their conjunction. Any stimulus according this view can be represented as "a coordinate or configuration along all of the dimensions that the occipitotemporal network represents" (p. 28). Arguably, the reason for having multiple representations of configurality might stem from the need for flexible object categorization, whereby different visual cues become diagnostic as a function of the specific task or recognition goal of the observer (Harel and Bentin, 2009; Schyns, 1998). Notably, these top-down modulations of visual processing are supported by both intrinsic interaction in the ventral visual pathway and extrinsic interactions between the ventral visual pathway and other amodal higher-level areas outside of occipito-temporal cortex (e.g., Chiou and Lambon Ralph, 2016; Harel et al., 2014). This hypothesis is obviously beyond the scope of the current work, as the current study used only a single task (visual discrimination). However, future studies that will parametrically manipulate both stimulus configurality and task relevance constitute a promising direction of research.

V. REFERENCES

Altmann, C. F., Bülthoff, H. H., and Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Curr. Biol.* 13, 342–349. doi:10.1016/S0960-9822(03)00052-6.

Ban, H., Yamamoto, H., Fukunaga, M., Nakagoshi, A., Umeda, M., Tanaka, C., et al. (2006). Toward a common circle: interhemispheric contextual modulation in human early visual areas. *J. Neurosci.* 26, 8804–8809. doi:10.1523/JNEUROSCI.1765-06.2006.

Bennett, K. B., and Flach, J. M. (1992). Graphical Displays: Implications for Divided Attention, Focused Attention, and Problem Solving. *Hum. Factors J. Hum. Factors Ergon. Soc.* 34, 513–533. doi:10.1177/001872089203400502.

Bennett, K. B., and Flach, J. M. (2011). *Display and interface design: Subtle science, exact art.* CRC Press.

Chechlacz, M., Mantini, D., Gillebert, C. R., and Humphreys, G. W. (2015). Asymmetrical white matter networks for attending to global versus local features. *Cortex* 72, 54–64. doi:10.1016/j.cortex.2015.01.022.

Chiou, R., and Lambon Ralph, M. A. (2016). Task-Related Dynamic Division of Labor Between Anterior Temporal and Lateral Occipital Cortices in Representing Object Size. *J. Neurosci.* 36, 4662–4668. doi:10.1523/JNEUROSCI.2829-15.2016.

Croner, L. J., and Albright, T. D. (1999). Seeing the big picture: integration of image

cues in the primate visual system. *Neuron* 24, 777–89. doi:10.1016/s0896-6273(00)81026-0.

DiCarlo, J.J., Zoccolan, D., & Rust, N.C. (2012). How does the brain solve visual object recognition?. *Neuron, 73*, 415-434.

Harel, A., and Bentin, S. (2009). Stimulus type, level of categorization, and spatial-frequencies utilization: Implications for perceptual categorization hierarchies. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1264–1273. doi:10.1037/a0013621.Stimulus.

Harel, A., Kravitz, D. J., and Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proc. Natl. Acad. Sci. U. S. A.* 111, E962–E971. doi:10.1073/pnas.1312567111.

Hasson, U., Hendler, T., Bashat, D. Ben, and Malach, R. (2001). Vase or Face ? A Neural Correlate of Shape-Selective Grouping Processes in the Human Brain. *J. Cogn. Neurosci.* 13, 744–753.

Kapadia, M. K., Ito, M., Gilbert, C. D., and Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15, 843–856. doi:10.1016/0896-6273(95)90175-2.

Kellman, P. J., and Shipley, T. F. (1991). A Theory of Visual Interpolation in Object Perception. *Cogn. Psychol.* 23, 141–221.

Koffka, K. (1935). *Principles of Gestalt Psychology*. London, England: Lund Humphries.

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., and Logothetis, N. K. (2003). Integration of Local Features into Global ShapesMonkey and Human fMRI Studies.

Neuron 37, 333–346. doi:10.1016/S0896-6273(02)01174-1.

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., and Mishkin, M. (2013).
The ventral visual pathway: An expanded neural framework for the processing of
object quality. *Trends Cogn. Sci.* 17, 26–49. doi:10.1016/j.tics.2012.10.011.

Kubilius, J., Wagemans, J., and Op de Beeck, H. P. (2011). Emergence of perceptual
Gestalts in the human visual cortex: the case of the configural-superiority effect.
Psychol. Sci. 22, 1296–303. doi:10.1177/0956797611417000.

Lamme, V. A. F., and Roelfsema, P. R. (2000). The distinct modes of vision offered by
feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579.
doi:10.1016/S0166-2236(00)01657-X.

Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., and Malach, R. (2001). A hierarchical
axis of object processing stages in the human visual cortex. *Cereb. Cortex* 11, 287–
297. doi:10.1093/cercor/11.4.287.

van Lier, R., and Gerbino, W. (2015). "Perceptual Completions," in *Oxford Handbook of
Perceptual Organization* (New York: Oxford University Press).

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. a, et al.
(1995). Object-related activity revealed by functional magnetic resonance imaging
in human occipital cortex. *Proc. Natl. Acad. Sci. U. S. A.* 92, 8135–8139.
doi:10.1073/pnas.92.18.8135.

Mersch, E. M. (2014). Linking neural states and performance: The configural superiority
effect and perception-action model (unpublished masters thesis). Wright State
University, Dayton, OH.

O'Craven, K. M., Downing, P. E., and Kanwisher, N. (1999). fMRI evidence for objects

as the units of attentional selection. *Nature* 401, 584–7. doi:10.1038/44134.

Ostwald, D., Lam, J. M., Li, S., and Kourtzi, Z. (2008). Neural Coding of Global Form in the Human Visual Cortex. 2456–2469. doi:10.1152/jn.01307.2007.

Pomerantz, J. R. (1986). Visual Form Perception: An overview. *Pattern Recognit. by humans Mach. Percept.*, 1–30.

Pomerantz, J. R., and Cragin, A. I. (2015). "Emergent features and feature combination," in *Oxford Handbook of Perceptual Organization*, ed. J. Wagemans (Oxford: Oxford University Press), 88–107.

Pomerantz, J. R., and Garner, W. R. (1973). Stimules configuration in selective attention tasks. *Percept. Psychophys.* 14, 565–569. doi:10.3758/bf03211198.

Pomerantz, J. R., and Portillo, M. C. (2011). Grouping and emergent features in vision: toward a theory of basic Gestalts. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1331–1349. doi:10.1037/a0024330.

Pomerantz, J. R., Sager, L. C., and Stoever, R. J. (1977). Perception of wholes and of their component parts: some configural superiority effects. *J. Exp. Psychol. Hum. Percept. Perform.* 3, 422–435. doi:10.1037/0096-1523.3.3.422.

Pomerantz, J. R., and Schwaitzberg, S. D. (1975). Grouping by proximity: Selective attention measures. *Percept. Psychophys.* 18, 355–361. doi:10.3758/BF03211212.

Schwaninger, A., and Mast, F. W. (2005). The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism. *Jpn. Psychol. Res.* 47, 216–222. doi:10.1111/j.1468-5884.2005.00290.x.

Schyns, P. G. (1998). Diagnostic recognition: task constraints, object information, and their interactions. *Cognition* 67, 147–179. doi:10.1016/S0010-0277(98)00016-X.

Sigman, M., and Gilbert, C. D. (2000). Learning to find a shape. *Nat. Neurosci.* 3, 264–269. doi:10.1038/72979.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012). A Century of Gestalt Psychology in Visual Perception : I . Perceptual Grouping and Figure – Ground Organization. *Psychol. Bull.* 138, 1172–1217. doi:10.1037/a0029333.

Ward, E. J., and Chun, M. M. (2015). Neural Discriminability of Object Features Predicts Perceptual Organization. *Psychol. Sci.* doi:10.1177/0956797615598317.

Woods, D. (1991). The cognitive engineering of problem representations. 169–188.

Yovel, G., and Kanwisher, N. (2004). Face Perception : Domain Specific , Not Process Specific. 44, 889–898.