

Perceived Size Modulates Cortical Processing of Objects

James M. Brown, B. A.

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science

in

Psychology

Dr. Anthony D. Cate, Chair

Dr. Kirby Deater-Deckard

Dr. Michael T. Braun

December 18, 2015

Blacksburg, VA

Keywords: object recognition, perceived size, fMRI, visual perception

Perceived Size Modulates Cortical Processing of Objects

James M. Brown

ABSTRACT

Empirical object recognition research indicates that objects are represented and perceived as hierarchical part-whole arrangements that vary according to bottom-up and top-down biases. An ongoing debate within object recognition research concerns whether local or global image properties are more fundamental for the perception of objects. Similarly, there is also disagreement about whether the visual system is guided by holistic or analytical processes. Neuroimaging findings have revealed functional distinctions between low and higher-level visual processes across lateral occipital-temporal cortex (LOC), primary visual cortices (V1/V2) and ventral occipital-temporal cortex. Recent studies suggest activations in these object recognition areas and others, such as the fusiform face area (FFA) and extra-striate body area (EBA), are collinear with activations associated with the perception scenes and buildings. Together, this information warrants the focus of the proposed study: to investigate the neural correlates of object recognition and perceived size. During the experiment subjects tracked a fixation stimulus while simultaneously being presented with images of shape contours and faces. Contours and face stimuli subtended small, medium and large visual angles in order to evaluate variance in neural activation across perceived size. In the present study visual areas were hypothesized to modulate as a function of visual angle, meaning that the part-whole relationships of objects vary with their perceived size.

Table of Contents

Abstract.....	ii
Table of Contents.....	iii
List of Figures.....	iv
List of Tables.....	vi
Perceived Size and Object Recognition.....	1
Neural Correlates of Perceived Size and Object Recognition.....	1
Hierarchical Part-Whole Relationships.....	2
Object Recognition and Contour Integration.....	3
Neural Correlates of Contour Integration and Shape Detection.....	4
Object Perception: Holistic and Analytical.....	5
Visual Processing and Holistic Perception.....	6
Face Perception.....	6
Neural Correlates of Face Perception.....	8
Hypotheses.....	10
Methods.....	10
Design.....	10
Participants.....	11
Informed Consent.....	11
Stimuli.....	11
Apparatus.....	13
Experimental and Control Conditions.....	13
Physical Size, Location and Retinal Size Conditions.....	13
Task: All Conditions.....	15
Neuroimaging Data.....	16
Analysis Plan.....	17
Results.....	17
Discussion.....	21
Conclusion.....	22
References.....	23

Appendix A.....	27
Appendix B.....	29

List of Figures

Figure 1. Example Images of Snakes and Ladders.....	4
Figure 2. Example Images of Composite Faces.....	7
Figure 3. Upright and Inverted Face Stimuli.....	12
Figure 4. GL and GO Contour Stimuli.....	12
Figure 5. Upright and Inverted Face Activations.....	18
Figure 6. GL and GO Contour Activations.....	18

List of Tables

Table 1. Face-Specific Peak Voxel MNI Coordinates.....	19
Table 2. Contour-Specific Peak Voxel MNI Coordinates.....	20

Perceived Size Modulates Cortical Processing of Objects

Perceived Size and Object Recognition

After years of empirical study, a complete explanation of the human visual experience remains elusive. While investigations into the mediating mechanisms of visual sensation and perception are ongoing, recent converging evidence regarding the spatial factors related to objects, specifically their physical sizes and distances they're viewed at, indicates exploration in this area may help make object recognition research become a bit more lucid.

Recent studies manipulating the intrinsic and extrinsic spatial cues of objects (e.g. their physical size and viewing distance) suggest these overlooked factors profoundly effect how objects are recognized and perceived. In a study by Tan (2004) examining the efficacy of physical display size and mental rotation performance, subjects were superior on a larger display compared to a smaller display, even when the stimuli subtended equal visual angles across different display sizes. In a similar study by Bakdash, Augstyn, & Proffitt (2006) investigating performance differences across small and large displays, subjects were required to navigate through a virtual environment and mentally track the locations of objects when they were out of sight. Like the previous study, subjects' performances were superior on the large display. These performance differences are hypothesized to be the product of an increased egocentric bias, wherein subjects feel more immersed and mentally interactive with the task at hand. Given that objects are inherently physical, it may seem intuitive that spatial reasoning abilities interact as a function of perceived size. However, these results are particularly striking as they challenge the long held tenet that object perception is invariant of real-world size (Kaufman & Rock, 1962).

Neural Correlates of Perceived Size and Object Recognition

Recent neuroimaging studies have found the presence of topographically defined regions of the brain that show specificity for objects' real-world sizes. In a study by Konkle & Oliva (2012), subjects were asked to mentally visualize inanimate objects at different sizes and were also presented with images of big and small real-world objects that subtended varying visual angles (e.g. trashcan, refrigerator and coffee mug, leaf respectively). Results of this study revealed that large and small objects preferentially activated along the lateral surface of the brain. These activations appeared to be organized according to object's real-world sizes, navigating respective object activations, big object were associated with parahippocampal activity and small objects were associated inferior temporal activity. Additionally, lateral

occipital, inferotemporal, and parahippocampal brain regions displayed characteristic size-related activations for objects that were invariant across changes in both visual angle and mentally visualized size. Although these results suggest that knowledge about an object's real-world size influences how an object is cortically represented, these findings do not answer whether or not an object's perceived size alters its mental representation. In a subsequent study Konkle & Caramazza (2013), they investigated the presence of cortical differentiation between the perception of animate and inanimate objects and perceptual size.

Agreeing with previous research, areas selective for images of objects were found within the ventral and lateral surface of the occipital-temporal cortex. Moreover, viewing animate objects was strongly associated with activity in the FFA, fusiform body area (FBA) and the extrastriate body area (EBA), while inanimate objects were strongly associated with PPA and transverse occipital cortex (TOS) activity. Interestingly, activations that were associated with inanimate objects differed across subtended visual angle, but activations associated with perceiving animate objects did not. While more research is needed in order better understand these findings, at face value it can be posited that an object's perceived size is a determining factor in how inanimate objects are cortically represented. This raises the possibility that altering an object's perceived size can also alter how the object is mentally represented. If perceptual size can alter object representations and object perception, by extension the structural image properties of object would also be susceptible to these effects.

Hierarchical Part-Whole Relationships

Although research on object recognition and spatial context is relatively nascent, spatial factors regarding objects structure have long been studied. Specifically, part-whole relationships are thought to describe the structural image properties of objects. In this way, an objects structure is hierarchically based. For years there has been much debate concerning how the hierarchical part-whole relationships of objects interact in such away that influences their recognition and subsequent perception. Further, it appears that visual system uses bottom-up and top-down processes to cortically represent and visually mentalize objects (Wagemans, 2012b). Under the framework of hierarchical part-whole relationships, the perceptual recognition of objects posits that objects are made of local elements (i.e. parts) that comprise a global form (i.e. whole) (Wagemans, 2012b). Moreover, global and local percepts are low-level visual processes, as they are the building block of edge detection. Considering the visual perceptual awareness of objects

(e.g. being consciously aware that an object has been viewed), it's thought that the visual system either uses either a holistic process or an analytical process (Wagemans, 2012b; Rossion, 2013). Additionally, this level of visual processing is higher-level, and implies that the global and local structures of objects have already been detected by the visual system. Importantly, these lower and higher-level processes are fundamentally different; however that's not say they're not interdependent and modulate one another. Although the processes underlying edge detection (e.g. global and local image processing) and the processes underlying the visual awareness of objects (e.g. holistic and/or analytical perception) research suggest that these processes may share conceptually similar perceptual biases (Wagemans, 2012b).

Object Recognition and Contour Integration

Studies since the 1970s object recognition studies have demonstrated consistent global-local relationships, such as “global precedence” and “global-local interference”, perceptual phenomena demonstrating rapid global form awareness and local feature obscuring in typical subjects (Navon, 1997). Current studies assessing the interactions between global-local image properties use Gabor stimuli to test part-whole relationships by equating individual Gabor wavelets as local elements and spatially consistent arrays of Gabor wavelets as global forms. Gabor wavelets, which are often called Gabor patches in vision studies, are 2D sinusoidal contrast bandpass functions filtered by a spatially small Gaussian contrast window. The narrow frequency and spatial domains of Gabor patches mirror the response specificities of V1 receptive fields, including specific 2D orientation and scale. Given the fundamentality of spatial frequency information, Gabor stimuli allow tightly controlled investigation of the relationship between contour detection and global-local image processing across a wide variety of abstractions (Jain, Ratha, & Lakshmanan, 1997).

Edge detection and contour integration studies typically assess how quickly subjects perceive form and how accurately subjects make figural judgments using Gabo stimuli that include both simple line segments and also closed shape contours. Line segments and shape contours, congruent figures are defined by high local phase and orientation uniformity across a global form, whereas incongruent figures are defined by global-local phase and orientation orthogonality. Performance studies evaluating global-local contour integration of line figures commonly refer to congruent forms as “snakes” and incongruent forms as “ladders” (Bex, Simmers, & Dakin, 2001). Across a wide variety of psychophysical tasks, such as how long it

takes participants to perceive a figure and how accurately participants make global-local spatial judgments, snake performance is consistently superior. While it may seem trivial that global-local congruent figures would be easier to judge than incongruent figures, these findings are important because they emphasize that congruence between local and global features is important for successful object perception, and that neither global nor local image properties are more or less important on their own.

Obfuscating these findings are recent studies that have indicated that snakes and ladders may not necessarily be processed by the same perceptual mechanisms. In a contour detection study by Bellacosa Marotti, Pavan, & Casco (2012) investigating the duration it took participants to perceive snakes and ladders, researchers found that altering figural global-local phase constancy and local element scale could improve subjects' ability to detect snakes, but not ladders. Additionally, other studies have found that global-local spatial judgment accuracy is differentially influenced by the curvature of the paths defining snakes and ladders (Pettet, 1999). Given the performance differences demonstrated across these studies, it appears that the processes regulating global-local object perception of line forms may not be absolute.

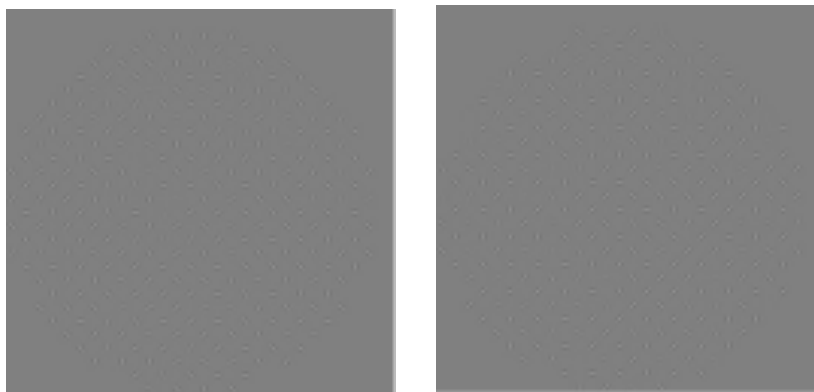


Figure 1. “Snake” and “Ladder” forms depicted from left to right respectively.

Neural Correlates of Contour Integration and Shape Detection

In a study by Mijovic et al. (2014), EEG and fMRI were used to investigate the neural correlates of contour detection using congruent and incongruent shape stimuli. Global-local congruent shapes were operationally defined as figures with perfect local alignment across an embedded global contour and incongruent shapes were defined as figures whose global structures were characterized by orthogonal local alignment. For the neuroimaging study, participants engaged in a passive viewing paradigm in which they were presented congruent and

incongruent shape stimuli while attending to a fixation stimulus that changed its screen location every trial. Results revealed significant activations in the lateral occipital cortex (LOC) and V1/V2. Importantly, these results agree with previous work emphasizing the relationship between the high- and low-level visual areas. It is believed that the LOC and V1/V2, recurrently process spatial frequency information in order to integrate local elements into closed contours. However, LOC to V1/V2 activations were not particularly specific to either congruent or incongruent shape stimuli. Paralleling contour detection studies involving line segments, these findings and others seem to indicate that the mechanisms used to perceive congruent and incongruent shape stimuli are likely to be independent.

Thinking back to the cognitive findings regarding perceived size and mental rotation performance and the neuroimaging findings regarding showing that the real-world physical sizes of objects are topographically represented on the brain, it's reasonable to posit that these effects may be robust across different conditions (Tan, 2004; Bakdash, Augstyn, & Proffitt, 2006; Konkle & Oliva, 2012; Konkle & Caramazza, 2013). Moreover, if these effects were found during the visual perception of contours, this would demonstrate that object recognition is modulated by spatial factors at the lowest level of visual processing. Based on this information, the following predictions can be made:

- Contours that subtend large retinal sizes will activate areas of the brain associated with the visual perception of scenes and buildings (PPA).
- Contours with congruent global forms that are described by complete local alignment will be cortically represented in the LOC, while contours with global forms that are described by incomplete local alignment will be cortically represented in brain areas outside of the LOC.

Object Perception: Holistic and Analytical

The classic definition of holistic perception is perceptual awareness of an object as a whole in a way that cannot be explained as a mental concatenation of that object's parts. Conversely, analytical perception occurs when the perception of an object as a whole is dependent on the summation of an object's parts (Wagemans, 2012b). Through the lens of global-local object recognition, holistic processing can be conceptualized as occurring when the global form of an object is perceived independently from the awareness of that object's local parts and analytical processing occurs when global and local figure information are integrated as

needed to perceive an object as a whole. Similar to the debate regarding global-local image processing, there is disagreement as to whether holistic or analytical perception is more fundamental for object recognition.

Visual Processing and Holistic Perception

The classic definition of holistic perception is perceptual awareness of an object as a whole in a way that cannot be explained as a mental concatenation of that object's parts. While contour detection paradigms are appropriate for studying the low-level visual mechanisms involved in global-local image processing, they are not adequate for evaluating higher level visual processes like holistic and analytical perception. Presently, holistic and analytical perception studies are focused on human face perception. As human faces are complex figures that can be analyzed as hierarchical object arrangements, they confer the ability to study holistic and analytical processes at a level that is tightly controlled and ecologically valid (Rossion, 2013).

Face Perception

In face perception research, part-whole relationships are described by a global form, which is a Gestalt representing a whole face, and the local elements that comprise a face. Local elements in this sense are features, such as mouth shape and eye color, and also configuration, which pertains spatial relativity between distal features. Face perception researchers state that holistic mental representation corresponds to a visual percept that's founded by internal knowledge, like a template (Rossion, 2013). If a face is processed holistically, the visual percept is sensorily unitary and face parts are not perceived independently from the whole face. In this way, if one part of the face is changed, then the perceptual awareness of that face as a Gestalt also changes. Among the most widely used paradigms used to evaluate the presence of holistic and analytical processing is the composite face paradigm. The composite face paradigm is a delayed matching task in which subjects are presented a set of composite faces that are comprised of top and bottom face halves. Furthermore, these face halves are either aligned or misaligned and top face halves always belong to the same face, but the bottom face halves change (See Figure 2 for composite face example). In a given trial, subjects are tasked to decide whether the two top face halves belong to the same face. Studies have continually shown that observers mistake the top face halves to be different from each other when they are spatially aligned compared to misaligned faces. These performance differences indicate that

discriminating individual features becomes harder when face halves are aligned compared to being misaligned.

These findings have been used as evidence to support that notion that when faces are aligned, they are processed holistically. If upright faces are being processed holistically, local face cue differentiation is complicated because they cannot be perceived independently from the Gestalt representing the whole face. Moreover, differentiating local face cues for misaligned faces is easier because holistic perception is disrupted, as the face is no longer a Gestalt biasing the perception of facial features and facial feature based configurations (Rossion, 2013).

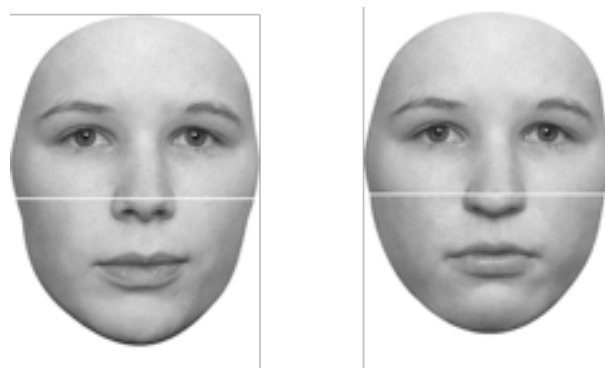


Figure 2. Two composite faces of identities

Adding another dimension to differential face perception processes is a phenomenon known as the “face inversion effect”. The face inversion effect is evinced by deficits in facial recognition ability whenever faces are presented in an inverted orientation, compared to faces that are upright. Importantly, studies using images of inverted non-face objects, like cars and houses, do not demonstrate the same performance differences typified by the face inversion effect (Rossion, 2013). The dissociations in object recognition performance have been hypothesized to imply that the perception of inverted faces is qualitatively distinct from the perception of upright faces and non-face objects. Building on these findings, studies have examined the effects of using inverted and upright face stimuli in the composite face paradigm. These studies have revealed that, compared to upright faces, the composite face effect is absent when faces are inverted. Face perception researchers posit that the absence of this effect indicates that inverted faces are processed analytical, much like what happens when upright faces that misaligned are being processed. These findings have been hypothesized to suggest that holistic face processing only occurs for upright faces that are aligned, and that faces that are

inverted are processed analytical because local face cues are free from the bias of a holistic template. Put another way, these hypotheses say that the visual system may have distinct modes for processing face information that utilize either local or global/Gestalt features depending on the perceptual encoding of the stimulus (Rossion, 2013).

While there is much support for the existence of holistic face perception, other studies have found contradicting results. Contextual factors of face stimuli such as demographic information, stimulus construction, and task design specificities have been found to nullify the composite face and face inversion effects where they would normally be present. These findings have been used to indicate holistic perception is rarely, if ever, the process by which we perceive faces (e.g. Blau, Maurer, Tottenham, & McCandliss, 2007; Nisbett & Miyamoto, 2005; Thierry, Martin, Downing, & Pegna, 2007). However valid these findings are, they fail to rule out holistic perception as a process as their effects can be explained by the existences of multiple holistic face templates defined by different part based saliences.

Neural Correlates of Face Perception

In light of the disagreement over behavior findings, recent neuroimaging studies have revealed areas of convergence regarding where and when face selectivity begins in the visual system. In general, three regions of the cerebral cortex have continually been implicated to be involved face perception, the occipital face area (OFA), lateral middle fusiform gyrus (FG) and the superior temporal sulcus (STS). Within the right hemisphere, the strongest differentiations between faces and objects have been found in an area of the lateral middle FG known as the “fusiform face area” (“FFA”) (e.g. Gauthier, Tarr, Moylan, Skudlarski & Gore , 2000; Kanwisher, McDermott & Chun, 1997; Rossion, Caldara, Seghier, Schuller, Lazeyras, & Mayer, 2003; Sergent, Ohta, & MacDonald, 1992). In a study by Schlitz, Dricot, Goebel, & Rossion (2010) using an fMRI adaptation design to evaluate the composite face effect differential activations were found between aligned and misaligned stimuli. Although both elicited activation in the FFA, it was demonstrated that compared to aligned faces with different bottom face halves, misaligned stimuli produced less fMRI adaptation. That is, FFA activation was reduced by presenting two displays with aligned faces in a row more than by presenting two unaligned face displays in a row. This indicates misaligned faces remained more novel than aligned face stimuli, meaning differential perceptual processing. Interestingly, in neuroimaging studies testing the face inversion effect compared to upright faces, FFA activation diminishes significantly

when subjects are presented with inverted faces (Yovel & Kanwisher, 2005). Importantly, other studies have not found significant activation decreases when subjects are presented with inverted non-face objects like cars and household appliances, thus implying that the FFA is not associated with perceiving inverted objects in general (Haxby et al., 1999). Studies of the inferior occipital gyrus (infOG), another area associated with face perception indicate that this region may be involved in processing facial feature-based input, subsequently relaying this information to the FFA and STS. This does not indicate that faces are encoded by features, in fact the opposite has been demonstrated, where encoding of the face as whole has been found. Given that inverted faces tend to elicit decreased activations in the FFA, it is reasonable to conclude that greater activations normally implicated in nonface object perception would be more active to inverted faces than upright faces. Multiple studies have demonstrated the absence of significant activations in the FFA when subjects are presented with nonface objects (Kanwisher, Stanley and Harris, 1999; Kanwisher, N. & Yovel, G., 2006). However, areas traditionally associated with the perception of nonface objects and scenes, such as the parahippocampal gyrus and the parahippocampal place area, have been shown to be significantly more active in the presence of inverted face stimuli compared to upright face stimuli. Moreover, in these same studies, the FFA still remains significantly activated in the presence of upright and inverted faces, but is typified by inverted face activations that are significantly diminished compared to upright face activations. That is, face selective regions are activated less and object selective regions are activated more in the presence of inverted face stimuli (Schlitz, Dricot, Goebel, & Rossion, 2010). Altogether, the findings of these studies lend more support to the notion that upright faces are processed holistically, whereas inverted and misaligned upright faces are processed analytical.

Once again, thinking back to the cognitive findings regarding perceived size and mental rotation performance and the neuroimaging findings regarding showing that the real-world physical sizes of objects are topographically represented on the brain, it's reasonable to posit that these effects may be robust across different conditions (Tan, 2004; Bakdash, Augstyn, & Proffitt, 2006; Konkle & Oliva, 2012; Konkle & Caramazza, 2013). Also, if these effects were found during the visual perception of contours, these effects would likely also manifest in high level visual processing (e.g. such as face perception). Based on this information, the following predictions can be made:

Whether up-right or inverted, face stimuli that subtend large retinal sizes will activate areas of the brain associated with the visual perception of scenes and buildings (PPA).

The visual perception up-right faces will activate more brain areas related to the perception of faces (e.g. the middle fusiform gyrus, FFA) than inverted faces.

Hypotheses

The present study evaluated recent findings suggesting that perceived size alters the perception of objects. Investigation into the mediating mechanisms of these processes were framed through the lens of hierarchical object recognition theories. Global and local processes were evaluated with shape contour global-local and global-only congruent stimuli (GL contours and GO contours). Holistic and analytical processes were evaluated with faces presented in upright and inverted spatial orientations. The hypotheses relevant to the present study were as follows:

- Perceived size will alter the perception of the local and global image properties of shape contours such that shape contours subtending large visual angles will elicit activations that will be positively correlated with increased activations in areas traditionally associated with the perception of scenes and buildings (e.g. PPA).
- GO shape contours will elicit greater activations in areas outside of the LOC than GL shape contours.
- Perceived size will alter the perception of upright and inverted faces such that faces subtending large visual angles will elicit activations that will be positively correlated with increased activations in areas traditionally associated with the perception of scenes and buildings (e.g. PPA).
- Inverted faces will elicit greater activations in areas outside of the FFA than upright faces.

Methods

Design

The present study entailed collecting neuroimaging data on subjects engaged in a passive-viewing paradigm. Five distinct conditions were employed to test for differences in holistic, analytical and global-local image processing as functions of perceived image size. Data were gathered utilizing a five-way factorial design (2x2x2x2x3) and a box-car blocking procedure. Within-subjects factorial analysis were used to conduct inferential and descriptive statistics on all

fMRI and behavioral measurements.

Participants

Participants were recruited from the Virginia Tech student and surrounding community populations. In accordance with the compensation practices of the VT Visual Neuroscience Lab, all participants were compensated at least 25 dollars per hour of participation that they completed in the study. In order to participate, subjects must have been between the ages of 18 and 50, and have had normal or corrected 20/20 vision. Participants who required corrective-eyewear were afforded MRI-safe glasses that were adjustable to individual prescription strength needs. In order to meet the statistical power requirements evinced by comparable visual perception studies, the present study involved 16 participants.

Informed Consent

Prior to participation subjects were required to read and sign an informed consent form. The consent form provided comprehensive information on experimental procedures, rights of the participant, possible risks and benefits, and subject compensation. Further, the informed consent form made clear that no subject was obliged to complete the study, subjects who withdrew would not be penalized in anyway, and that full compensation would be awarded for any study participation. After the participant had a sufficient amount of time to read and sign the consent form, the experimenter verbally went over the information provided in the document and made sure the participant had no further questions regarding the study.

Stimuli

Faces. In the face conditions subjects viewed images of faces that were either upright (Up Faces) or inverted (Inv Faces). Images of faces were provided by the *Yale Face Database B* (Georghiadis, Belhumeur, and Kriegman, 2001). The face images were high-resolution, real-life photographs of adult Caucasian males and females. All images were altered to remove discrete gender identifying descriptors, like hair, and were also gray-scale luminance rendered to make skin tone uniform between face identities (see Figure 3A and B).

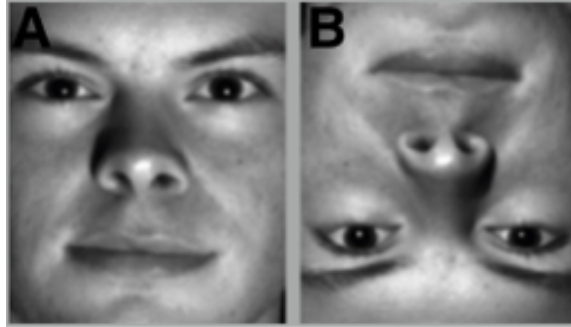


Figure 3. Up Faces (A) and Inv Faces (B) respectively.

Contours. In the contour conditions subjects viewed two types contours, global-local contours (GL contours) and global-only contours (GO contours). GL contours had global forms comprised of elements that were in complete local alignment, whereas GO contours had global forms comprised of elements that were not in complete local alignment (see Figure 4A and B). Contours were created by replicating the methods used by Mijovis et al. (2014), as they not only investigated research questions similar to the proposed study, but they also managed to obtain strong cortical activations as well. Using GERT, the Grouping Elements Rendering Toolbox, shape contours were generated by constructing arrays of non-overlapping Gabor elements.

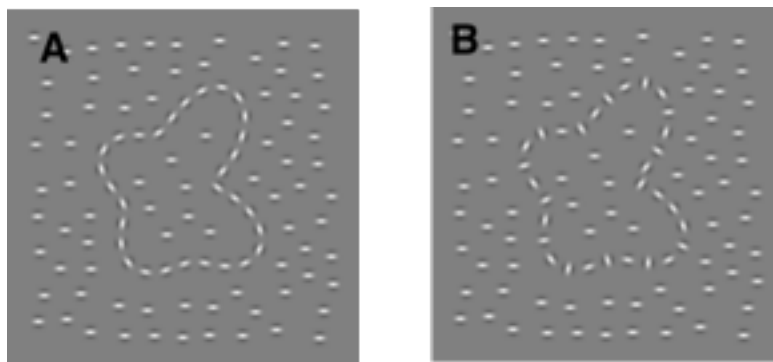


Figure 4. GL contours (A) and GO contours (B) respectively.

Gabor elements formed contours around globally defined shapes embedded in a uniform gray background. Global outlines of shapes were rendered by plotting radial frequency components in a polar coordinate plane; further, each component having a random phase angle and amplitude. This process constructed moderately complex shapes whose features were balanced for convexities, protrusions and indentations. GL contour elements were locally aligned along the global outline of background embedded shapes. For GO contour elements, half were aligned locally and the other half aligned perpendicularly along the global outline of background embedded shapes (Mijovis et al., 2014).

Apparatus

A 3T Siemens scanner at Virginia Tech was used to collect all neuroimaging data. A 3D T1-weighted data set encompassing the whole brain was acquired for every subject (120 slices, 1 mm slice thickness, matrix size = $344 \times 344 \times 344$). All visual stimuli were presented using a PC running PsychToolbox through a projector. PsychToolbox will also be used to track behavioral responses. A compound microscope (“SwiperScope”) was also built by the experimenter for the purposes of being able to show images at varying distances away from the scanner.

Experimental and Control Conditions

Over the course of the experiment all participants engaged in eight runs of blocked trials, eight blocks of experimental trials and nine blocks of control trials. All blocks of experimental trials were eight trials in length and all blocks of control trials were six trials in length. During the blocks of experimental trials all participants viewed images from all four experimental image conditions, which included two contour conditions (GL contours and GO contours) and two face conditions (Up Faces and Inv Faces). During the blocks of control trials subjects viewed a white fixation cross set against a black background. Blocks of experimental trials and control trials were counterbalanced with respect to each other, both within a given run and across all other runs in the experiment. The presentation order of the images within a given block of experimental trials or a given block of control trials was pseudorandom.

Physical Size, Location and Retinal Size Conditions

Detecting changes in brain activation as a function of perceived size was a unique challenge, as factors related to how large or small objects are visually perceived are naturally confounded. In particular, these factors include: physical size of an object, physical distance of an object from an observer and the size of visual angle subtended by an object on the retina of an observer. Given that the third factor is a linear combination of the first two factors, we reasoned that we would have to assess changes in brain activation as a function of perceived size by accounting for these factors as individual, but co-occurring, experimental conditions. The real-world size of the images presented, the real-world physical distance of the images from the participant and the retinal image size subtended by the object on the participant’s retina, were individually represented as Physical Size, Location and Retinal Size conditions respectively.

Physical Size Condition. This condition represented the physical size of the images viewed by the participants. There were two qualitatively distinct levels of physical image size, images that

were small (200x200 pixels) and images that were large (600x600 pixels). Both small and large images occurred equally often during the experiment.

Location Condition. This condition represented the physical distance between the participants and the screen they viewed the images on. Participants were exposed to two qualitatively distinct levels of screen location, where participants viewed images that were either physically near or far away from them, shown at 28 cm and 185 cm distances respectively. The presentation of images at near and far distances occurred equally often during the experiment. This experimental condition was reliant on two apparatus manipulations. One of those manipulations involved the SwiperScope we built for this experiment. During the far screen conditions, but not the near screen conditions, the SwiperScope was utilized. The other manipulation involved physically moving the screen we build for this experiment, which was placed at a closer distance to the scanner during the near screen condition than where the screen was placed during the far screen condition.

Retinal Size Condition. This condition represented the retinal image size of the images viewed by the participants. There were three qualitatively distinct levels of retinal image size, small (1°), medium (1.7°) and large (3.1°) respectively. All retinal image sizes were presented equally often in terms of co-occurring with all of the other experimental conditions (e.g. no retinal images sizes were overrepresented by either levels of the Physical Size and Location conditions). However medium retinal image sizes were overrepresented with respect to the Retinal Size conditions (e.g. images more often subtended the medium retinal image size than either of the other retinal image size conditions).

The primary reason for the presence of three retinal images sizes, and not just two, was for the purpose of being able to validly test our hypotheses related to the effects of perceived size and cortical processing. Specifically, we rationalized that in order to detect changes in brain activation as a function of perceived size, we would have to account for any effects related to qualitatively differences in retinal image size. Apart from the possible cognitive effects induced by differences perceived size, it's conceivable that differences significant differences in cortical processing could also be due to attentional effects, such as novelty and visual salience. Further, effects unrelated to perceived size would manifest as a function of either the Physical Size or Location conditions independently. To address this possibility we added retinal size conditions where there were objective differences in retinal image size (e.g. the small and large retinal size

conditions). In this manner, differences as a function of perceived size would be found by holding the effects of the Physical Size and Location conditions constant (e.g. the medium retinal size condition). If after controlling for any effects related to differences between small and large retinal image sizes, as indexed by their independent interactions across the Physical Size and Location conditions, then significant differences found between either the Physical Size or Location conditions would reflect effects not intrinsic to the medium retinal image and likely represent effects of perceived size.

While having three levels of retinal size condition was critical to being able to test our experimental hypotheses, both the confounding nature of the Physical Size and Location conditions, and our desire to reduce participant fatigue contributed, to our allowing the medium retinal image size condition to be overrepresented. In terms of design, having three qualitatively distinct levels of retinal image size represented by dichotomous manipulations of the Physical Size and Location conditions forced the structure of the experiment to either be uneven with respect the representation of the retinal image size or uneven with respect to the levels of the Physical Size and Location conditions. Because finding significant differences between overrepresented levels of either of the Physical Size or Location conditions would not answer any questions relevant to the goals of our experiment, we allowed these levels of these conditions to occur equally often. In this way, having all levels of the Retinal Size condition would have required us to not only increase the number of scanning runs, but also the number of blocked experimental and control trials in the experiment. Had we added instituted those increases, participants would have had to engage in the experiment significantly longer, and become more susceptible to fatigue and practice effect. In addition, these increases would have significantly increased the total cost of the experiment. Although this aspect of our design allowed for the possibility that our results would be biased in the direction of the medium retinal image size condition, we rationalized that our balanced and careful control of all other experimental conditions allowed for being able to account for potential biases through statistical analysis.

Task: All Conditions

While being scanned, all participants engaged in a task that required them to visually attend to a white fixation cross that appeared in the center of the screen. Prior to being scanned, all subjects instructed that they were to track whenever the fixation cross changed colors, which was were either light yellow or dark yellow; subjects indicated color changes in the fixation

cross with button box responses. The purpose of this task was to prevent participants from habituating to the images presented to them during the experiment, as well as prevent practice effects. This task was the same throughout all conditions of the experiment and occurred over the full duration of all experimental and control block of trials.

Neuroimaging Data

Procedure. Following their provided informed consent, subjects were briefed on the experimental tasks that they would perform. Before entering the scanner, the experimenter would show the subjects images of the stimuli they would be presented, which would be GL and GO images during the contour trials of the experiment and upright and inverted face images during the face trials experiment. Next the subjects were briefed on MRI safety protocol. In accordance with standard safety MRI procedures, all participants were asked if they had any free metal on or in their body, such as metal implants or piercings. Before loading the subject into the MRI scanner, the experimenter screened subjects with a metal detecting wand to maximally reduce any potential hazards posed by the scanner's magnetic field. Next subjects were loaded into the scanner, where they engaged in the experiment. Participants wore a head coil equipped with a vertically rotatable mirror to view visual stimuli. During all experiments participants were asked to attend to a fixation cross displayed at the center of the screen.

All visual stimuli were displayed once every TR ($TR = 2.26$ s) and had fixation durations of 1.5 s (allowing a 1 s interstimulus interval). The temporal scanning parameters were chosen to prevent subjects from becoming fatigued, while also maintaining a rigorous criterion for functional data acquisition. In order to induce passive processing and reduce the potential for confounding attentional biases to visual stimuli, subjects were asked to make fine feature judgments about the fixation cross and to ignore all other images. Total scanning time per subject lasted approximately 45 minutes, divided into 8 runs approximately 6 minutes in length. In order to avoid subject fatigue. This scan length provided adequate time to complete screening and set-up procedures as well as the 5-min MPRAGE anatomical scan while ensuring that the entire experiment lasts no longer than 1 hr. Counterbalancing of the blocks of the experimental and control trials was carried out by ensuring that within a given run the order of a given set of four blocks of experimental trials, which represented the four experimental image conditions, never occurred twice in a given run. Specifically, for a given run, the ordered set of the four blocks that defined a given second set of four blocks of experimental trials was the reverse order of the first

set of four blocks of experimental trials. Further, every run in the experiment began and ended with a control block of trials and every experimental block of trials was preceded and/or followed by a control block of trials.

Analysis Plan

The main analysis consisted of a group analysis using random effects to identify between-subjects relationships. Relationships examined were between subtended visual angle of stimuli and activation in cortical visual processing regions. The random effects analysis was based on the results of a within-subjects analysis for each participant. This consisted of a regression in which voxels related to blood oxygenation level dependent (BOLD) signal were correlated with the subtended visual angles of shape contours and faces across all conditions. Time courses of indicator variables were represented according to a dummy coded design matrix: 1 indicated the image appeared in the trial in question and 0 otherwise. A general linear model (GLM) analysis performed in Statistical Parametric Mapping (SPM) software was used to assign beta weights to every independent variable. This GLM indicated how much the BOLD response changed when each image appeared. These values were used to evaluate the regions expected to correlate with object recognition and perceived size.

Results

Statistical maps representing peak activation were projected onto the MNI brain, a standardized T1-weighted MRI anatomical template. Labeled activations by brain area are based on to their MNI coordinate positions. Collapsing across the effects of retinotopic size, distance and physical size, significant bilateral and unilateral activations were found for all visual stimuli (e.g. GL contours, GO contours, Up Faces, Inv Faces). For face stimuli, both the left and right fusiform gyrus the visual perception of upright faces and inverted faces (see Figure 5A, B and C, and Table1). For contour stimuli, both the left and right calcarine sulci and the left middle occipital gyrus during the visual perception of GL and GO contours (see Figure 6A, B and C, and Table2). Planned contrasts testing our hypotheses did not reveal any statistical differences in brain activation.

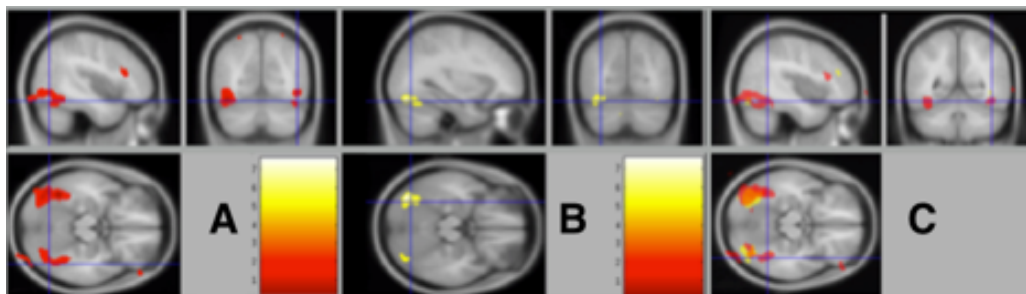


Figure 5. Up Faces (A), Inv Faces (B), and the overlay of Up Faces and Inv Faces (C) respectively. Contrasts between the individual activities of upright and inverted faces in both the left and right fusiform gyri were significant relative to their fixation baselines. Activations were significant at $p < .001$ (uncorrected for multiple comparisons).

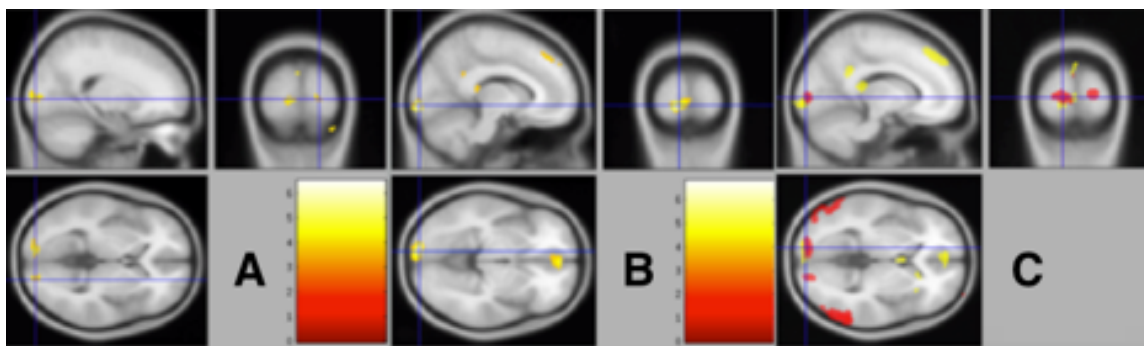


Figure 6. GL contours (A), GO contours (B), and the overlay of GL and GO contours (C) respectively. Contrasts between the individual activities of GL and GO contours in both the left and right calcarine sulci were significant relative to their fixation baselines. Activations were significant at $p < .001$ (uncorrected for multiple comparisons).

Table 1

Face-Specific Activations by Regions of interest (ROIs) in MNI coordinates

ROIs	X	Y	Z
Right fusiform gyri (Up Faces)	41	-49	-17
Left fusiform gyri (Up Faces)	-34	-48	-17
Right fusiform gyri (Inv Faces)	37	-58	-18
Left fusiform gyri (Inv Faces)	-33	-56	-20

Notes. Contrasts between the face-specific activations and their fixation baselines.

Table 2

Contour-Specific Activations by Regions of interest (ROIs) in MNI coordinates

ROIs	X	Y	Z
Right calcarine sulcus (GL contours)	19	-91	0
Left calcarine sulcus (GL contours)	-10	-90	0
Left calcarine sulcus (GO contours)	-5	-102	-6

Notes. Contrasts between the contour-specific activations and their fixation baselines.

Discussion

The purpose of our study was to investigate whether or not the neural processes underlying object recognition modulate as a function of visual angle, such that the part-whole relationships comprising objects covary with their perceived sizes. In order to answer this question we had our subjects observe images of contours and faces that subtended small, medium and large visual angles, and were also viewed at varying distances. Further, contours and faces viewed by the subjects were differed with respect to their part-whole relationships. Contours and faces were chosen because a wealth of empirical studies the previously indicated that the visual processing of contours and faces is susceptible to distortion, depending on the part-whole relationships describing them (Rossion, 2013; Wagemans et al., 2012a). Specifically, the effects of global-local precedence and FIE were chosen because they have been well established by previous object contour integration and face perception studies respectively (Rossion, 2013; Wagemans et al., 2012a). To test cortical processing effects related to global local precedence would covary with perceived size, contours viewed by subjects were either defined by global forms comprised of locally elements, or global forms defined by partially aligned local elements.

None of our hypotheses were supported. In the first and third hypotheses we predicted that the PPA would become activated during the visual perception of contours and faces that subtended large visual angles. Moreover, we predicted this activation would vary in magnitude, depending on the global and local structures comprising the contours, as well as the orientations of the faces. The lack of significant effects could have been due to a number of factors involving the design of the study, as well as our statistical analyses. In terms of design, it's possible neither contours or faces were presented often enough to elicit activations in the areas of interest. In order to address this issue instead of focusing on both contours and face, we could have chosen to only show our subjects only one of those stimulus types. In terms of statistical analyses, it's possible that there were effects present, however the signal filtration procedures we used were too strict to see the effects. To be more specific, it's possible that the interaction between the stimulus effects, which were presented at a high frequency, and spatial context effects (e.g. physical size, viewing distance, retinal size) were missed by the spatial context effects were presented at a lower frequency. For example, whereas a given stimulus type changed 16 times in a given scanner run, the alterations in spatial context variables changed no more than four times over the whole course of the experiment. This presents a unique problem evaluating our data,

and may mean the conventional analysis procedures used for fMRI block design experiments won't suffice here. With regard to our second and fourth hypotheses, where we predicted there would be significant differences in activation within the stimulus type conditions (e.g. GL contours versus GO contours, and Up Faces versus Inv Faces), our inability to find significant effects here could also stem from the same problems that influenced results related to our first and third hypotheses.

A part from methodological issues, another possibility is that the effects we hypothesized are absent. While our predictions flowed logically from the research discussed, it's reasonable that the effects related to perceived size and mental rotation performance and the real-world physical sizes of objects being topographically represented on the brain are truly separate. In this way, these effects would stem from process specific interactions that do not share a significant overlap in neural network representation. If this is the case, this would go against recent research suggesting the brain is less modal than we've come to believe and instead bolster existing literature suggesting the brain is rigidly modal in function.

Conclusion

The present study investigated cortical differentiation in the visual system of contours and faces as a function of their part-whole relationships and perceived sizes. Our study was the first of its kind, as no other fMRI studies have had subjects view images on displays that were featured at different distances while being scanned. In addition, our study was also among the first contour integration studies conducted in an fMRI setting. None of our hypotheses were supported, and so at this time there is no evidence to suggest that cortical activations elicited across different part-whole configurations, in the context of contours and faces, modulate as a function of their perceived sizes. However, given the potential methodological flaws in our study and all of the previous research strongly suggesting a functional link between the intrinsic and extrinsic spatial contexts of objects, their recognition and perception, it would be remiss to say that research in this area complete.

References

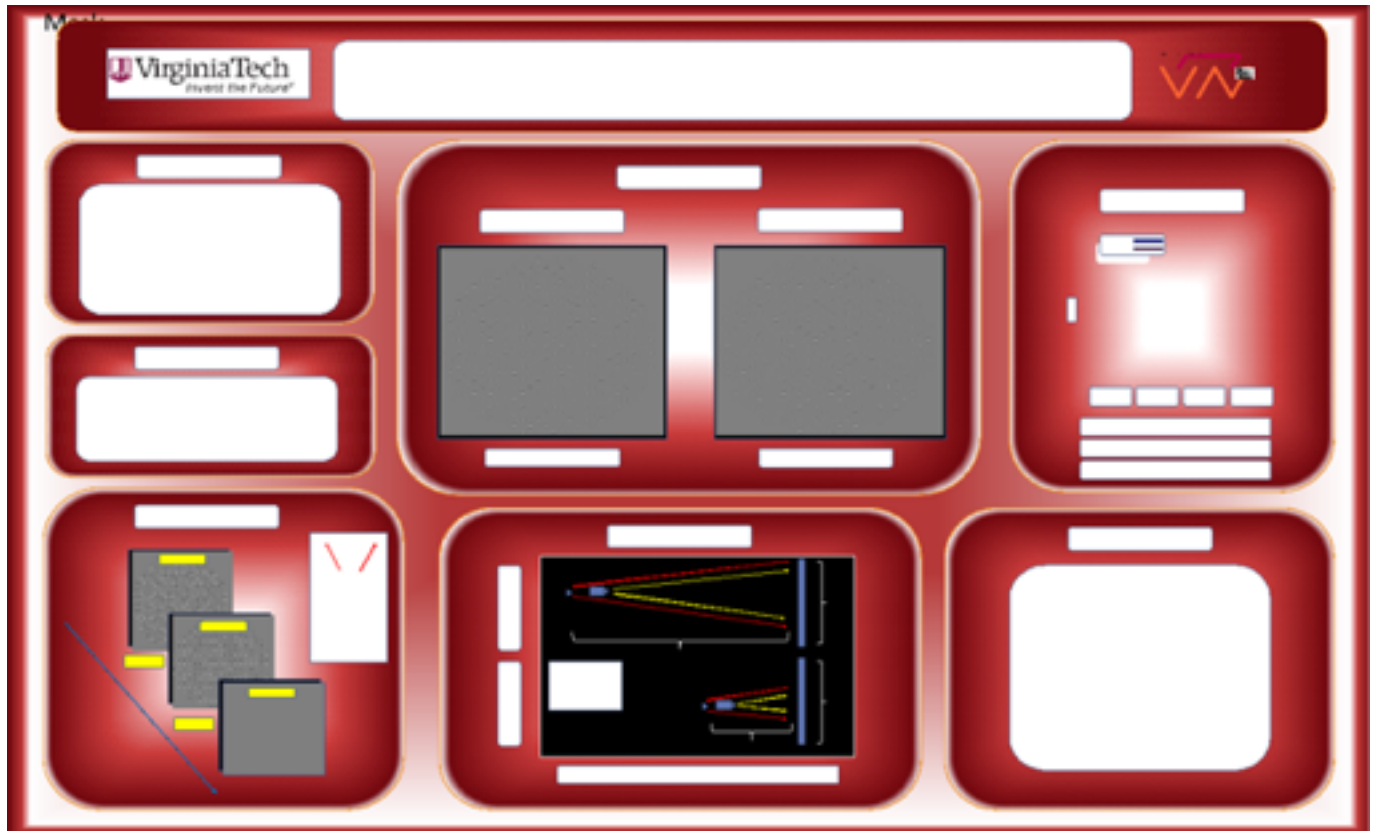
- Aguirre, G. K., Zarahn, E., & D'esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron*, *21*(2), 373-383.
- Bakdash, J. Z., Augstyn, J. S., & Proffitt, D. R. (2006). Large Displays Enhance Spatial Knowledge of a Virtual Environment, APGV '06 Proceedings of the 3rd symposium on Applied perception in graphics and visualization, pp 59-62.
doi:[10.1145/1140491.1140503](https://doi.org/10.1145/1140491.1140503)
- Bellacosa Marotti, R., Pavan, A., & Casco, C. (2012). The integration of straight contours (snakes and ladders): The role of spatial arrangement, spatial frequency and spatial phase. *Vision research*, *71*, 44-52.
- Bakdash, J. Z., Augstyn, J. S., & Proffitt, D. R. (2006). Large Displays Enhance Spatial Knowledge of a Virtual Environment, APGV '06 Proceedings of the 3rd symposium on Applied perception in graphics and visualization, pp 59-62.
doi:[10.1145/1140491.1140503](https://doi.org/10.1145/1140491.1140503)
- Bellacosa Marotti, R., Pavan, A., & Casco, C. (2012). The integration of straight contours (snakes and ladders): The role of spatial arrangement, spatial frequency and spatial phase. *Vision research*, *71*, 44-52.
- Bex, P. J., Simmers, A. J., & Dakin, S. C. (2001). Snakes and ladders: The role of temporal modulation in visual contour integration. *Vision research*, *41*(27), 3775-3782.
- Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, *3*(7), 1-13.
- Cho, S., Metcalfe, A. W., Young, C. B., Ryali, S., Geary, D. C., & Menon, V. (2012). Hippocampal–prefrontal engagement and dynamic causal interactions in the maturation of children's fact retrieval. *Journal of cognitive neuroscience*, *24*(9), 1849-1866.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598-601.
- Davidoff, J., Fonteneau, E., & Fagot, J. (2008). Local and global processing: Observations from a remote culture. *Cognition*, *108*(3), 702-709.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision research*, *33*(2), 173-193.

- Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of cognitive neuroscience*, *12*(3), 495-504.
- Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*(1), 189-199.
- Heath, S., McMahon, K. L., Nickels, L., Angwin, A., MacDonald, A. D., van Hees, S., & Copland, D. A. (2012). Neural mechanisms underlying the facilitation of naming in aphasia using a semantic task: an fMRI study. *BMC neuroscience*, *13*(1), 98.
- Hess, R. F., & Dakin, S. C. (1999). Contour integration in the peripheral field. *Vision research*, *39*(5), 947-959.
- Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences*, *96*(16), 9379-9384.
- Jain, A. K., Ratha, N. K., & Lakshmanan, S. (1997). Object detection using Gabor filters. *Pattern Recognition*, *30*(2), 295-309.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302-4311.
- Kanwisher, N., Stanley, D. & Harris, A. (1999). The fusiform face area is selective for faces not animals, *Neuroreport*, *10*, 183-187.
- Kanwisher, N. & Yovel, G. (2006) The fusiform face area: a cortical region specialized for the perception of faces, *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476):2109-28.
- Kaufman, L. & Rock, I. (1962), The Moon Illusion, I: Explanation of this phenomenon was sought through the use of artificial moons seen on the sky. *Science*, *136*, pp. 953–961.
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), 1114-1124.
- Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *The Journal of Neuroscience*, *33*(25), 10235-10242.
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Computer vision*,

1999. *The proceedings of the seventh IEEE international conference on* (Vol. 2, pp. 1150-1157). Ieee.
- Machielsen, W., Rombouts, S. A., Barkhof, F., Scheltens, P., & Witter, M. P. (2000). FMRI of visual encoding: reproducibility of activation. *Human brain mapping*, 9(3), 156-164.
- Mijović, B., De Vos, M., Vanderperren, K., Machilsen, B., Sunaert, S., Van Huffel, S., & Wagemans, J. (2014). The dynamics of contour integration: A simultaneous EEG–fMRI study. *NeuroImage*, 88, 10-21.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive psychology*, 9(3), 353-383.
- Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *Trends in cognitive sciences*, 9(10), 467-473.
- Pettet, M. W. (1999). Shape and contour detection. *Vision Research*, 39(3), 551-557.
- Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal-neocortical functional reorganization underlies children's cognitive development. *Nature Neuroscience*, 17(9), 1263-1269.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, 126(11), 2381-2395.
- Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, 21(2), 139-253.
- Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision*, 10(2), 25.
- Sergent, J., OHTA, S., & MACDONALD, B. (1992). Functional neuroanatomy of face and object processing A positron emission tomography study. *Brain*, 115(1), 15-36.
- Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature neuroscience*, 10(4), 505-511.
- Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological bulletin*, 138(6), 1218.

- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological bulletin*, 138(6), 1172.
- Yovel, G. & Kanwisher, N. (2005) The Neural Basis of the Behavioral Face-Inversion Effect, *Current Biology*, 15, pp. 2256-2262.

Appendix A



Appendix B

Curriculum Vitae

James M. Brown**Contact Information**

Address: Department of Psychology
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0002

Phone: (540) 838-1552

Email: bjames1@vt.edu

Education

2013-present **Virginia Polytechnic Institute & State University**, Blacksburg, VA.
Ph.D. candidate in the Biological Psychology program.
Cumulative GPA: 3.75. Expected Graduation: May 2018

2007-2011 **Coastal Carolina University**, Conway, SC.
B.A. in Psychology.
Cumulative GPA: 3.48. Graduated: May 2011

Conference Poster Presentations

Brown, J. M., Cate, A. D. (2014, Nov) *Effects of Physical Size on Visual Contour Integration and Global-Local Judgments of Hierarchical Forms*. Society for Neuroscience, Washington, DC.

Cate, A. D, **Brown, J. M.**, Roldan, S. (2014, Nov). *Visual Grouping and Visual Enumeration Performance: Signal Detection Analysis*. Society for Neuroscience, Washington, DC.

Research & Related Experience

2013-present **Graduate Teaching Assistant**, Virginia Tech Visual Neuroscience Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA.
Dr. Anthony Cate, advisor