Does face inversion qualitatively change face processing: An eye movement study using a face change detection task

Buyun Xu Department of Psychology, University of Victoria, Victoria, Canada

James W. Tanaka Department of Psychology, University of Victoria, Victoria, Canada

Understanding the Face Inversion Effect is important for the study of face processing. Some researchers believe that the processing of inverted faces is qualitatively different from the processing of upright faces because inversion leads to a disproportionate performance decrement on the processing of different kinds of face information. Other researchers believe that the difference is quantitative because the processing of all kinds of facial information is less efficient due to the change in orientation and thus, the performance decrement is not disproportionate. To address the *Qualitative* and *Quantitative* debate, the current study employed a response-contingent, change detection paradigm to study eye movement during the processing of upright and inverted faces. In this study, configural and featural information were parametrically and independently manipulated in the eye and mouth region of the face. The manipulations for configural information involved changing the interocular distance between the eyes or the distance between the mouth and the nose. The manipulations for featural information involved changing the size of the eyes or the size of the mouth. The main results showed that change detection was more difficult in inverted than upright faces. Specifically, performance declined when the manipulated change occurred in the mouth region, despite the greater efforts allocated to the mouth region. Moreover, compared to upright faces where fixations were concentrated on the eyes and nose regions, inversion produced a higher concentration of fixations on the nose and mouth regions. Finally, change detection performance was better when the last fixation prior to response was located on the region of change, and the relationship between last fixation location and accuracy was stronger for inverted than upright faces. These findings reinforce the connection between eye movements and face processing strategies, and suggest that face inversion produces a qualitative disruption of looking behavior in the mouth region.

Introduction

It has been known for decades that inversion impairs the recognition of faces more than any other object category (e.g., airplanes, stick figures, houses), as reported by the landmark study by Yin (1969). The robustness of the effect has been demonstrated in old/new recognition tasks (Leder & Bruce, 2000; Leder & Carbon, 2006; Rhodes, Brake, & Atkinson, 1993), same/different discrimination tasks (Goffaux & Rossion, 2007; Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Tanaka, Kaiser, Bub, & Pierce, 2009; Yovel & Duchaine, 2006; Yovel & Kanwisher, 2004) and delayed forced-choice matching tasks (Boutet & Faubert, 2006; Freire, Lee, & Symons, 2000; Pellicano, Rhodes, & Peters, 2006; Rhodes, Hayward, & Winkler, 2006; Tanaka & Farah, 1993; Tanaka & Sengco, 1997). It has been suggested that the face inversion effect is one of the most compelling arguments that faces are processed by distinct cognitive mechanisms (e.g., Rossion, 2008; but see Valentine, 1988).

The qualitative versus quantitative debate

Two diverse opinions about whether inversion produces a qualitative or quantitative change in face processing can be found in the literature. Researchers holding the *Qualitative* view argued that inversion differentially impaired one kind of information more than another (Rossion, 2008). For example, there are two kinds of cues that one can derive from an individual face, namely the featural information and the configural information. Featural information refers to the properties of the individual parts of a face, such as the shape of the mouth, the size of the eyes, etc. Configural

Citation: Xu, B. & Tanaka, K. W. (2013). Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. *Journal of Vision*, *13*(2):22, 1–16, http://www.journalofvision.org/content/13/2/22, doi:10.1167/13.2.22.

doi: 10.1167/13.2.22 Received August 16, 2011; published February 18, 2013 ISSN 1534-7362 © 2013 ARVO

information refers to the metric distances between features on the face, such as the distance between eyes and eyebrows, the distance between two eyes, the distance between the nose and the mouth, etc. In a holistic face representation, the two sources of featural and configural information are combined in an integrated perceptual representation (Rossion, 2008; Tanaka & Farah, 1993). According to the *Qualitative* view, featural and configural information are decoupled when a face is inverted, such that inversion is more disruptive to the processing of configural relations between features than the features themselves (Barton, Keenan, & Bass, 2001; Cabeza & Kato, 2000; Freire et al., 2000; Leder & Bruce, 2000; Leder, Candrian, Huber, & Bruce, 2001; Leder & Carbon, 2006; Rhodes et al., 1993). For example, Leder and Carbon (2006) studied the recognition of three different sets of faces in both upright and inverted orientations. One set of faces (color set) differed from each other only in color, the second set of faces (relational set) had identical shaped local features but the spatial relation between these features were different, and the third set of faces (component set) only differed in two of the three (eyes, mouth or nose) components. The results showed that the relational set revealed a strong inversion effect; the component set revealed a moderate inversion effect and the color set revealed no inversion effect. According to the configural/featural interpretation, inversion qualitatively impairs the perception of configural information in a face more than the perception of featural information.

Other evidence suggests that inversion qualitatively disrupts a different kind of information processing related to the region of a face (Malcolm, Leung, & Barton, 2004; Tanaka et al., 2009). To test this claim, Tanaka and colleagues designed the Face Dimensions Task (Bukach, Le Grand, Kaiser, Bub, & Tanaka, 2008; Rossion, Kaiser, Bub, & Tanaka, 2009; Wolf et al., 2008) in which configural and featural information are independently and parametrically manipulated in the upper and lower regions of the face. Configural information was manipulated by changing the distance between the eyes or the distance between the mouth and the nose. Featural information was manipulated by changing the size of the eyes or the size of the mouth. Tanaka and colleagues (2009) found whereas inversion had relatively little effect on the discrimination of featural and configural differences in the eye region, it severely disrupted the perception of changes in the lower region of the face. According to the *Regional* view then, inversion qualitatively impairs featural and configural information in the lower mouth region of the face while preserving featural and configural information in the upper eye region.

In contrast, advocates of *Quantitative* view have claimed that inversion impairs the processing of featural information[1] as much as (e.g., Riesenhuber et al., 2004; Yovel & Duchaine, 2006; Yovel & Kanwisher, 2004), and in some cases even more than the processing of configural information (e.g., Rhodes et al., 1993). For example, Yovel and Kanwisher (2004) tested the Face Inversion Effect in a sequential same/different discrimination task. When performance on configural and featural trials was equalized in the upright orientation, recognition of the featural changes was as difficult as the recognition of configural changes when the faces were inverted. Critically, in the Yovel and Kanwisher study, featural and configural changes included manipulations to both the eye and mouth regions. Sekuler, Gaspar, Gold, and Bennett (2004) also found that in a perceptual matching task, subjects attended more to information in the eye region of a face regardless of whether it was presented in an upright or inverted orientation. Compatible with the *Quantitative* view, they argued that upright and inverted faces are processed equivalently, but that information is extracted more efficiently in an upright face than an inverted face.

The role of eye movements in face processing

A potentially useful method to examine the source of the inversion effect is to monitor eye movements while participants are looking at upright and inverted faces. Although viewers can allocate attention independent of eye position in simple tasks (Posner, 1980), Rayner (2009) argued that eye location (overt attention) and covert attention are highly associated in more complex tasks and therefore this method is a useful tool for understanding mediating cognitive operations. In the face literature, eye-tracking techniques has been employed to study holistic face processing (Bombari, Mast, & Lobmaier, 2009; de Heering, Rossion, Turati, & Simion, 2008; van Belle, de Graef, Verfaillie, Rossion, & Lefevre, 2010a), face recognition (Henderson, Williams, & Falk, 2005; Hsiao & Cottrell, 2008), the perception of facial expressions (Aviezer et al., 2008; Wong, Cronin-Golomb, & Neargarder, 2005), the processing of faces in different views (Bindemann, Scheepers, & Burton, 2009), and the recognition of familiar and unfamiliar faces (Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006; Heisz & Shore, 2008; van Belle, Ramon, Lefèvre, & Rossion, 2010b). In terms of the direct comparison between the processing of upright and inverted faces, existing eye movement studies do not solve the *Qualitative* versus *Quantitative* debate. Consistent with the *Regional* view, Barton et al. (2006) found that participants had more fixations to the eye region in an upright face and more fixations to mouth and lower face in an inverted face. However, in another recognition study, Williams and Henderson (2007) recorded eye movements during both the learning and recognition phase, and found that eye movement did not differ whether faces were presented upright or inverted, suggesting that the face inversion effect was not a consequence of distinct patterns of eye movement. In sum, the two studies yielded inconsistent results regarding the utility of eye movement behaviors to explain whether inversion influences the *Qualitative* versus *Quantitative* processing of faces.

The aim of the current study is to investigate whether the pattern of eye movement differs between the processing of upright and inverted faces. We employed a change detection task (Rensink, 2002) in which participants were asked to decide whether two alternating face stimuli were the "same" or "different" using the stimulus set from the Face Dimensions Task. This task has been used to study the featural and configural processes of healthy adults (Tanaka et al., 2009) and infants (Quinn & Tanaka, 2009), individuals with autism (Wolf et al., 2008) and patients with prosopagnosia (Bukach et al., 2008; Rossion et al., 2009). One of the strengths of the Face Dimensions Task is that it decouples the processing of configural and featural information from the processing of information in eye and mouth regions. Therefore, it provides a rigorous test of the *Configural/Featural* versus *Regional* qualitative views as well as the *Quantitative* view of the face inversion effect.

In the current study, the *Qualitative* and *Quantitative* views are examined by linking the participant's eye movements to their ability to detect featural and configural changes in different regions on upright and inverted faces. According to the *Quantitative* perspective, the same cues will be used during the processing of upright and inverted faces. Thus, inversion will cause a uniform performance decrement in face processing (i.e., equal amount of decrement of performance in featural and configural processing in both the eye and mouth regions). In contrast, the *Qualitative* view argues that inversion should lead to disproportionate decrements of performance on judgments of different kind of face information (such as configural and featural, or eye region and mouth region) as indicated by performance and eye tracking behaviors. According to the *Regional view*, inversion should result in impaired performance and different eye movement behavior when detecting changes in the eye versus mouth region. According to the *Configural/Featural* view, inversion should result in impaired performance and different eye movement behaviors when detecting featural versus configural changes.

Method

Participants

Twenty-two (nine female; 13 male) undergraduate students at the University of Victoria volunteered for the study. All participants had normal or corrected-to-normal visual acuity and were naïve to the purpose of the study.

Stimuli

Each face picture was 300 × 400 pixels in size and subtended a visual angle of 7.7° horizontally, and 10.3° vertically at a viewing distance of 78 cm. Two faces from the picture database of the Face Dimension Task were selected. The stimuli were created using high quality, gray-scale digitized photographs of six children's faces (three male, three female). Images were cropped at each side of the head. No jewelry, glasses, or makeup were present in those pictures, and facial markings such as freckles, moles, and blemishes were removed digitally. Using Adobe Photoshop, the size of the eyes or mouth of each original face, and the distance between the inner edges between the two eyes and the inner edges between the nose and the mouth were modified. Thus, four dimensions of change were created: configural eyes, configural mouth, featural eyes, and featural mouth. Each dimension of change consisted of five faces along a continuum: the original (primary) face and four incrementally varied (secondary) face images. This process created a total of 20 variations per face. In the featural condition, the location and the shape of the eye or the mouth are kept unchanged, and the size of the eyes or mouth was manipulated by resizing the original feature by 80%, 90%, 110%, or 120%. Due to the nature of the manipulations in the featural condition, changing the size of the eyes or the mouth while maintaining their original positions necessarily induces some configural changes. The magnitude of these changes was as follows: Within the eye condition, the interocular distance varied in increments of 4 pixels between each level of change; in the mouth condition, the distance from the philtrum varied in increments of 2 pixels. In the configural condition, the distance between the features was modified. Within the configural eye condition, the interocular distance was modified by increasing and decreasing this measure by 10 (approximately 16% of the original distance) and 20 pixels relative to the primary face. Configural mouth modifications involved shifting the mouth upwards and downwards vertically by 5 pixels (approximately 16% of the original distance) and 10 pixels. The size and shape of the features were held constant. Sample stimuli can be found in Figure 1. For every dimension along the five-step continua, the differences between faces that are separated by three steps in the continuum should be relatively "easy" to detect, faces separated by two steps should be "intermediate" and differences between faces separated by only one step should be

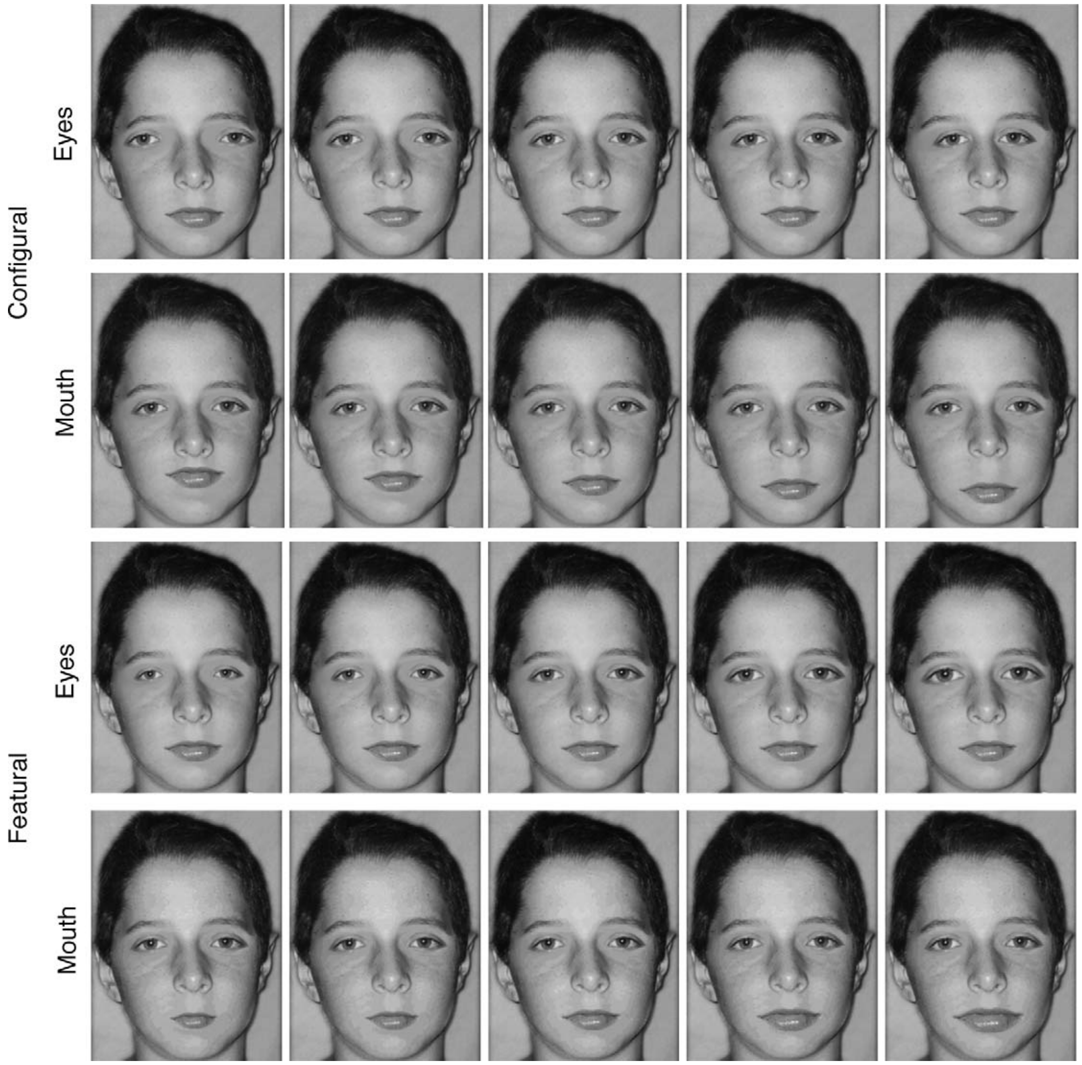

Figure 1. Example stimuli from the Dimension Tasks. From left to right, the difference between any two adjacent images refers to one step of change. The changes could be the distance between the eyes (Configural eye), the distance between the nose and the mouth (Configural mouth), the size of the eyes (Featural eye) and the size of the mouth (Featural mouth).

"difficult" to detect. In the current study, only two of the pictures of Caucasian boys at the intermediate level of difficulty were chosen as the stimuli.

Apparatus

Stimuli were displayed on a white background on an 18-inch CRT monitor (ViewSonic, Walnut, CA) controlled by a Macintosh desktop computer (Apple, Cupertino, CA). The viewing distance was 78 cm. Subjects responded by pressing one of two keys on a keyboard using the left and right index fingers. Eye movements were recorded at a 1000-Hz sampling rate using the tower mount configuration of an SR Research EyeLink 1000 system (SR Research, Osgoode, ON). This configuration provides an average fixation location accuracy between 0.25° and 0.50°. The pupil, using the centroid detection model, and corneal reflection of each subject's left eye was tracked under binocular viewing conditions. The participant's head was set on a chin rest, and fixed by a forehead rest. Eye tracking data were recorded using a Dell desktop computer (Dell, Round Rock, TX).

Procedure

There are 96 trials in all, with half of the trials presented in the upright orientation, and the half of the trials presented in the inverted orientation. Each trial consisted of two pictures of the same person's face, with or without changes to it. Among the 48 upright trials, 16 of the trails were regarded as catch trials since the two pictures were exactly the same. Among the remaining 32 trials, eight trials used images that differed in the size of

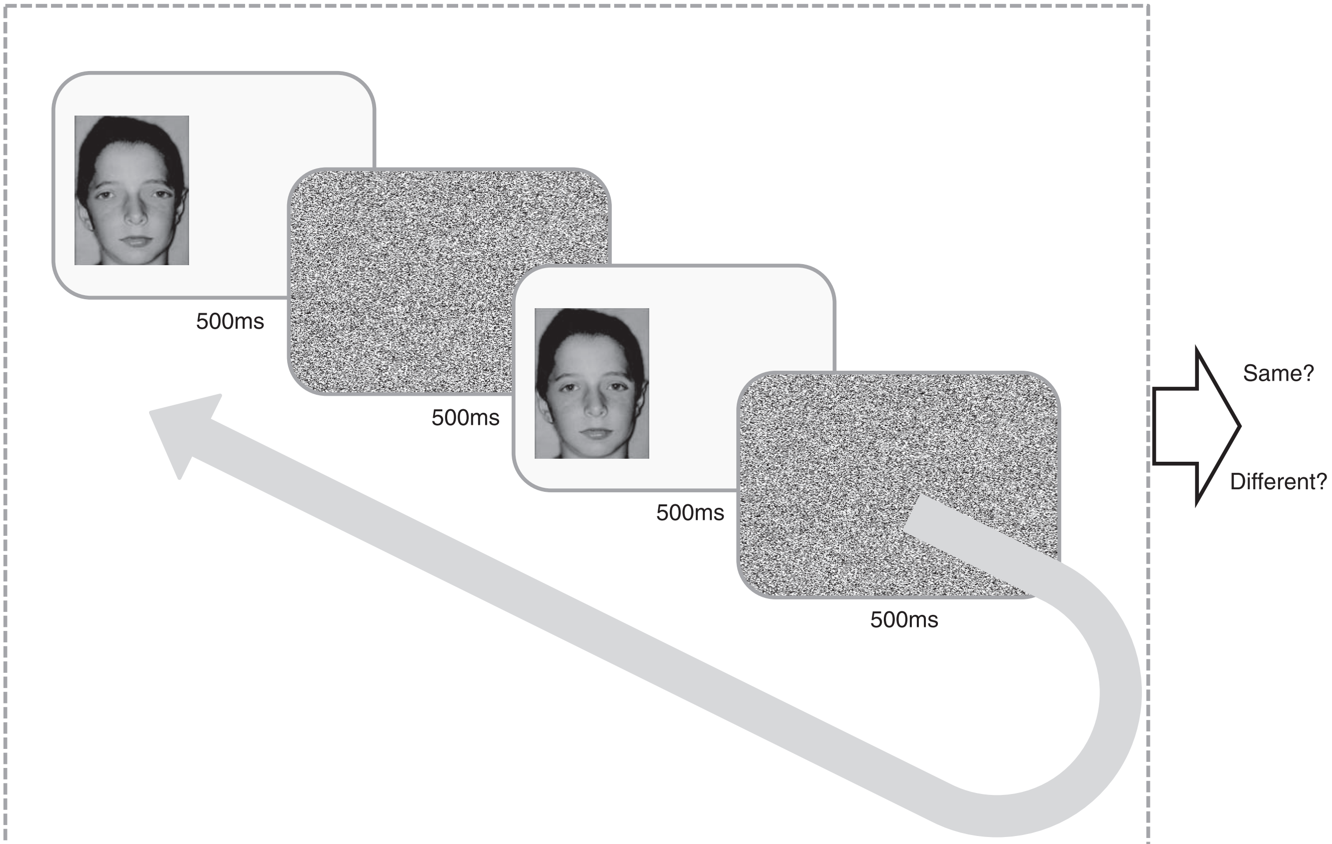

Figure 2. An illustration of the change detection task. Participants are required to fixate at the center of the screen and faces are presented either on the left or right randomly. Two faces of the same identity with or without configural or featural differences are presented sequentially with white noise masks between them. The presentation sequence is terminated by the participants' key response.

the eyes, eight trials used images that differed in the size of the mouth, eight trials used images that differed in the distance between the eyes and finally, the last eight trials that used pictures differing in the distance between the nose and the mouth. Therefore, the study was a 2 (Orientation: upright or inverted) × 2 (Region: eyes or mouth) × 2 (Change type: configural or featural) within-subjects design.

Before the eye movement data was recorded, a nine-point calibration process was conducted. Participants were given instructions and trained on the experimental process during a practice block of trials. Trials were presented in two randomized experimental blocks, separated by a short break. All eight conditions were mixed, and counterbalanced across the two blocks. Before each experiment block, the calibration process was conducted in which a center dot was presented on the screen for 2000 ms as the fixation point for drift correction, and subjects were instructed to always fixate on the dot when they see this screen. After that, participants then saw alternating images of faces displayed for 500 ms, separated by a 500 ms white noise mask (Figure 2). The alternating face images were pictures of the same person with identical or different facial features, but in the same orientation. To be specific, they could be exactly the same or different in size of the eyes (Featural eye difference), the size of the mouth (Featural mouth difference), the distance between two eyes (Configural eye difference) or the distance between the nose and the mouth (Configural mouth difference). In order to measure the first fixation landing on the face, images were randomly presented either to the left or right of the center fixation point. The noise masks were used to prevent two sequential presentations of similar images to create some perception of motion at the location where two images differ (e.g., Zelinsky, 2001). The repeating sequence was terminated either by pressing the response keys or after 30 seconds. Participants were required to press one key if the two pictures were the same or another key if they were different.

Results

Preprocessing of data

The recording of eye movement data begins with the onset of the face picture, and ends with the partici-

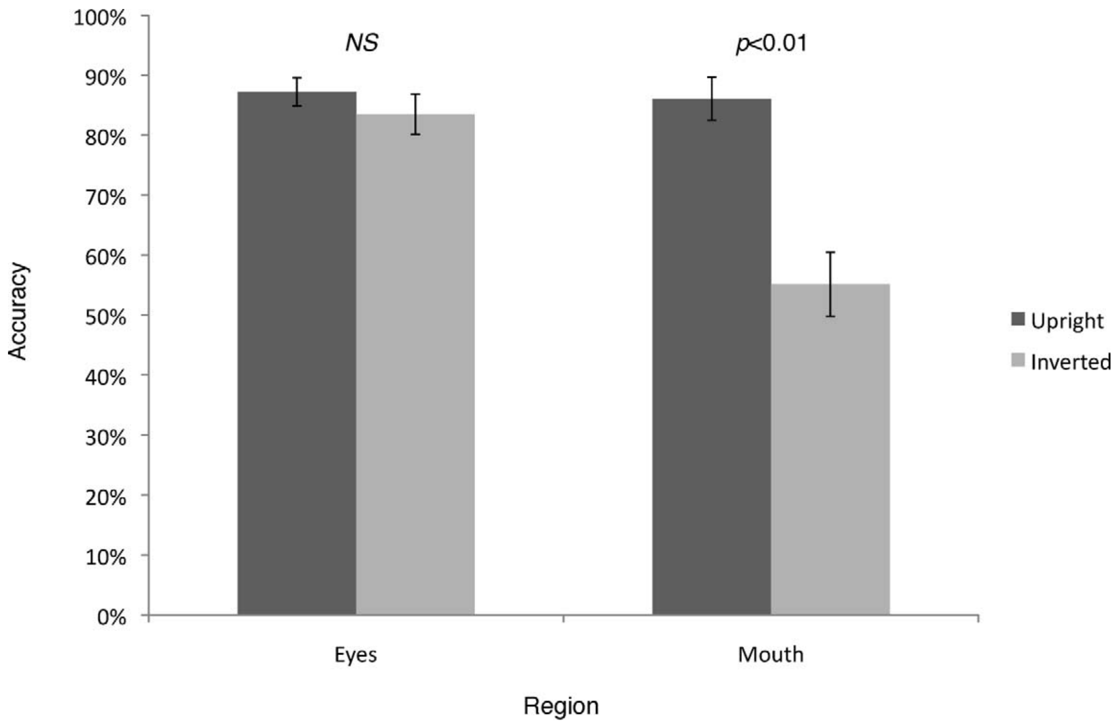

Figure 3. Change detection accuracy by region and orientation. Error bars refer to standard error.

pant's response. Because the face pictures are presented either to the left or right of the center fixation, there are fixations that are not on the picture after the picture's onset. Those fixations are not meaningful and are thus excluded. Moreover, fixations with the duration shorter than 50 ms were merged with nearby fixations. The nearby fixation was defined as either the preceding or the following fixation that is less than 0.5° away from the original fixation.

Behavioral measures

Response bias

Because different trials (67%) outnumbered same trials (33%), participants might develop a tendency to make "different" responses more often than "same." The overall response bias c ($M = 0.25$, $SE = 0.08$) was significantly larger than 0 ($t_{21} = 3.03$, $p < 0.01$), indicating that participants were more inclined to make a "different" response.

Accuracy

ANOVA was conducted with the main factors of Orientation (upright or inverted), Change Type (configural or featural) and Region (eyes or mouth). The results (Figure 3) showed a significant main effect of Orientation, $F(1, 21) = 24.35$, $p < 0.01$, and Region, $F(1, 21) = 16.97$, $p < 0.01$. Change detection in upright faces ($M = 0.86$, $SE = 0.02$) was significantly better than change detection in inverted faces ($M = 0.69$, $SE = 0.04$) and the change detection accuracy for the eye region ($M = 0.85$, $SE = 0.03$) was significantly better than that for the mouth region ($M = 0.71$, $SE = 0.04$). The two-way interaction between Orientation and Region, $F(1, 21) = 26.41$, $p < 0.01$, was also reliable. Performance in detecting changes in the eye region on upright faces ($M = 0.87$, $SE = 0.02$) was not significantly different ($p > 0.05$) than that of inverted faces ($M = 0.84$, $SE = 0.03$). However, the performance of detecting changes in the mouth region on upright faces ($M = 0.86$, $SE = 0.04$) was significantly better ($p < 0.01$) than that of inverted faces ($M = 0.55$, $SE = 0.05$). The two-way interaction between Change Type and Region was also found to be significant, $F(1, 21) = 8.25$, $p < 0.05$, indicating that configural changes in the mouth region ($M = 0.78$, $SE = 0.05$) were more difficult ($p < 0.01$) to detect than the eye region ($M = 0.92$, $SE = 0.02$), but no such effects were found ($p > 0.05$) for the featural changes. However, the two-way interaction between Orientation and Change Type was not significant, $F(1, 21) = 0.71$, $p > 0.05$), indicating that inversion effects did not differ between configural and featural conditions.

Total response time[2]

The analysis of response time (Figure 4) on correct trials showed a significant main effect of Orientation, $F(1, 19) = 24.11$, $p < 0.01$, and Region, $F(1, 19) = 15.69$, $p < 0.01$. Participants required significantly more time to detect changes in inverted faces ($M = 3638$ ms, $SE =$

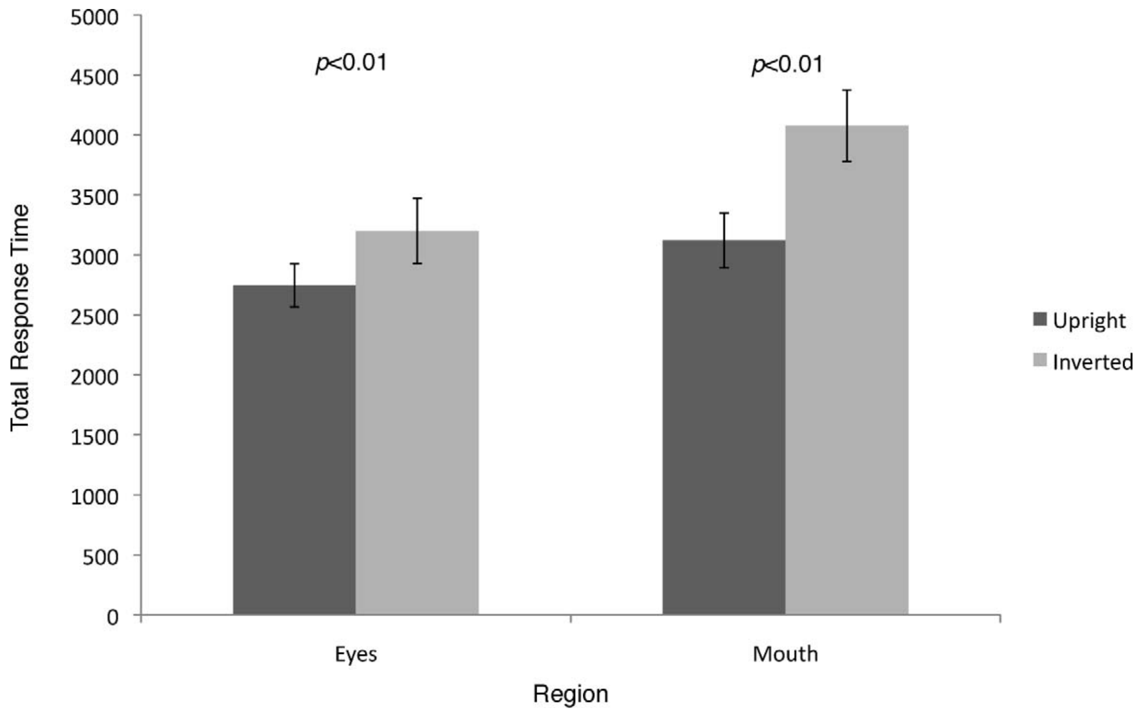

Figure 4. Total response time for correct trials by region and orientation. Error bars refer to standard error. Only data from correct trials were included.

260 ms) than upright faces ($M = 2934$ ms, $SE = 188$ ms). In addition, participants took more time to detect changes in the mouth region ($M = 3599$ ms, $SE = 241$ ms) than the eye region ($M = 2973$ ms, $SE = 217$ ms). The interaction between Change Type and Region was also significant, $F(1, 19) = 13.03$, $p < 0.01$). For configural trials, participants took longer time ($p < 0.01$) to detect changes in the mouth region ($M = 3721$ ms, $SE = 253$ ms) than eye region ($M = 2808$ ms, $SE = 177$ ms), however this effect was not present ($p > 0.05$) for featural trials. Moreover, the interaction between Orientation and Region was marginally significant, $F(1, 19) = 4.29$, $p = 0.052$. Participants spent significantly more time ($p < 0.01$) detecting changes in the mouth region on inverted faces ($M = 4077$ ms, $SE = 297$ ms) than in upright faces ($M = 3121.1$ms, $SE = 229$ ms). This effect was also present ($p < 0.01$) when detecting changes in the eye region, but the magnitude of the effect was smaller. The interaction between Orientation and Change Type was not significant, $F(1, 19) = 1.37$, $p > 0.05$, indicating that inversion effects did not differ between configural and featural conditions.

Eye movement measures

Number of saccades[3]

This measurement provides information of how many saccades were made in one trial. The significant main effect of Orientation, $F(1, 18) = 18.24$, $p < 0.01$ (Figure 5), showed that more saccades were executed in the inverted ($M = 11.29$, $SE = 0.7$) than the upright orientation ($M = 10.04$, $SE = 0.7$). The main effect of Region was also significant, $F(1, 18) = 9.55$, $p < 0.01$, indicating that participants needed more saccades to detect changes in the mouth ($M = 11.19$, $SE = 0.7$) than in the eyes ($M = 10.15$, $SE = 0.7$). The interaction between Change Type and Region was significant, $F(1, 18) = 5.33$, $p < 0.05$, indicating that detecting configural changes in the eye region ($M = 9.9$, $SE = 0.6$) requires less ($p < 0.01$) saccades than mouth region ($M = 11.4$, $SE = 0.7$), but the effect was not present for the detection of featural changes ($p > 0.05$). Moreover, the Orientation by Region interaction was marginally significant, $F(1, 18) = 4.01$, $p = 0.06$, participants needed to execute significantly more ($p < 0.01$) saccades to detect changes in the mouth region on inverted faces ($M = 12.1$, $SE = 0.7$) than upright faces ($M = 10.5$, $SE = 0.7$), whereas this effect was smaller for the eye region ($p = 0.06$). The interaction between Orientation and Change Type was not significant, $F(1, 18) = 1.37$, $p > 0.05$, indicating that inversion effects did not differ between configural and featural conditions.

Saccade distance

The saccade distance was measured by calculating the distance (visual degree) between two successive fixations, which reflects the average size of every saccade within each trial. Combined with the mea-

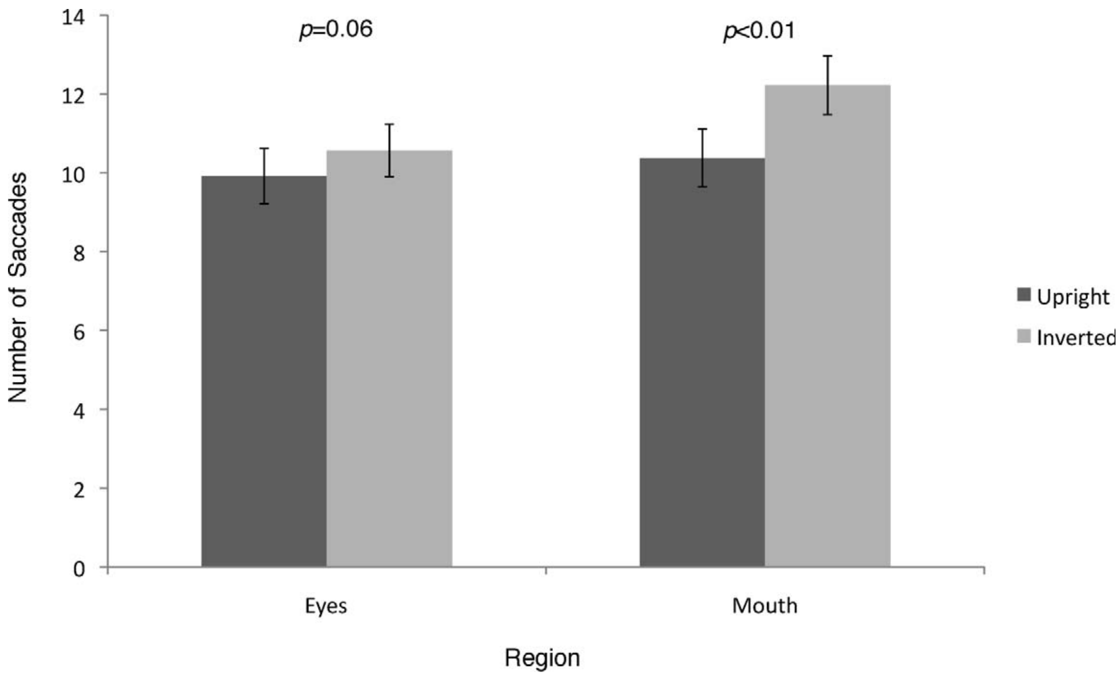

Figure 5. Number of saccades by region and orientation. Error bars refer to standard error.

surement of number of saccades, a general eye movement strategy could be inferred to reflect whether the participant was making a detailed scan of the picture or not. For the average saccade distance (Figure 6), Orientation showed a significant main effect, $F(1, 18) = 6.58$, $p < 0.05$. The average saccade distance was smaller when faces were inverted ($M = 2.6$, $SE = 0.1$) than when they were upright ($M = 2.8$, $SE = 0.1$). Region also showed a main effect, $F(1, 18) = 5.52$, $p < 0.05$. The average saccade distance was smaller when detecting changes in the mouth region ($M = 2.6$, $SE = 0.1$) than eye region ($M = 2.8$, $SE = 0.1$). The interaction between Orientation and Region was significant, $F(1, 18) = 4.39$, $p = 0.05$. When detecting changes in the mouth region on inverted faces ($M = 2.5$, $SE = 0.1$), the average saccade distance was smaller ($p < 0.01$) than that of upright faces ($M = 2.8$, $SE = 0.1$). However, this effect was not present when detecting changes in the eye region ($p > 0.05$). The interaction between Orientation and Change Type was not significant, $F(1, 18) = 2.76$, $p > 0.05$, indicating that inversion effects did not differ between configural and featural conditions.

Region of interest analysis

In order to quantify the eye movements observed over different face regions during the change detection process, 10 areas of interest were created on the face. They were left eye (area 1), right eye (area 2), nose (area 3), mouth (area 4), chin (area 5), left cheek (area 6), right cheek (area 7), forehead (area 8), left periphery (area 9) and right periphery (area 10) (Williams & Henderson, 2007). The 10 areas of interest were further collapsed into four key areas: the eyes (areas 1 and 2), which covers an area of 5.6 × 1.7 visual degrees, the nose (area 3), which covers an area of 3.2 × 1.8 visual degrees, the mouth (area 4) which covers an area of 3.2 × 1.4 visual degrees, and other region (areas 5, 6, 7, 8, 9, and 10; Figure 5).

First fixation

The first fixation location indicates where the participants first look at after the onset of the visual presentations. It acts like an anchor to the followed fixations. Locations of first fixation were coded into the four key areas of interest that included the eye region, nose region, mouth region, and "other." Numbers of first fixations that landed on the key facial feature areas of the eyes, nose and mouth, but not the "other" area were scored and analyzed (Figure 8). An Orientation (Upright, Inverted) by Area of Interest (eyes, mouth, nose) ANOVA was conducted. The main effect of Orientation was significant, $F(1, 20) = 5.54$, $p < 0.05$, indicating that more first fixations landed on the key features (eyes, nose and mouth, rather than other regions) of upright faces ($M = 14.9$, $SE = 0.3$) than inverted faces ($M = 13.4$, $SE = 0.6$). The main effect of Area of Interest was also significant, $F(2, 19) = 64.47$, $p < 0.01$, indicating that regardless of orientation, most of the first fixations landed in the eyes ($M = 24.2$, $SE = 1.5$) and nose region ($M = 17.7$, $SE = 1.6$), rather than

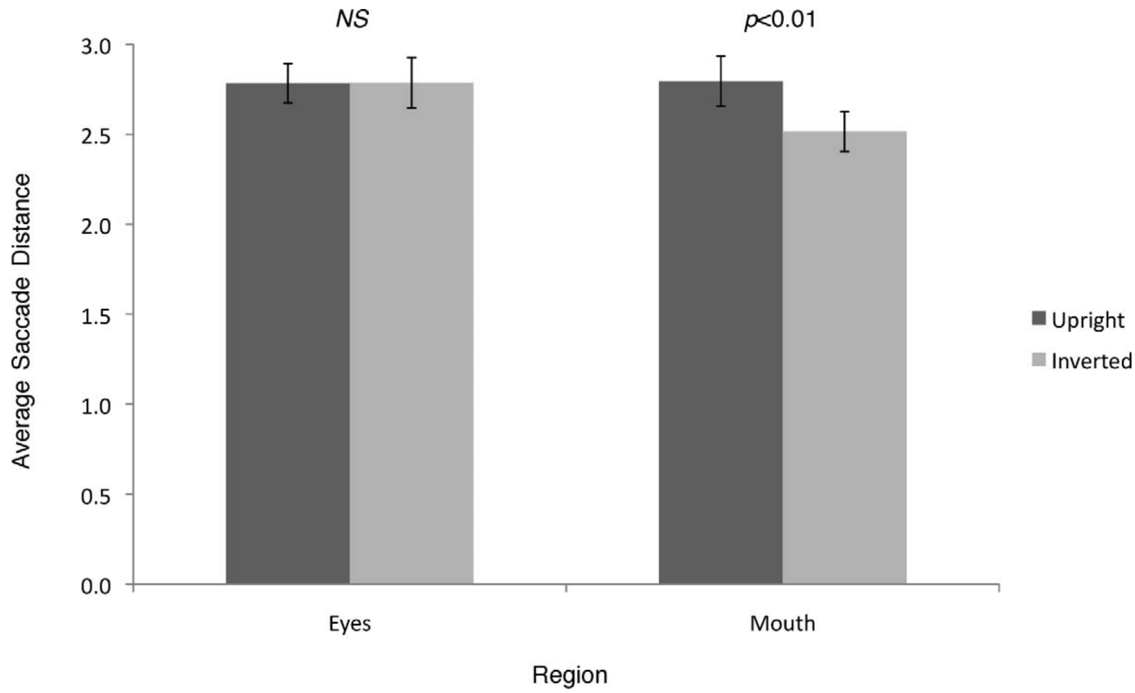

Figure 6. Average saccade distance by region and orientation. Error bars refer to standard error.

the mouth region ($M = 0.7$, $SE = 0.2$). More importantly, the interaction between Orientation and Area of Interest was also significant, $F(2, 19) = 7.35$, $p < 0.05$. Number of first fixations landed in the eye region on upright faces ($M = 31.9$, $SE = 3.4$) was significantly larger than ($p < 0.01$) that on inverted faces ($M = 16.4$, $SE = 2.6$). However, number of first fixations in the mouth ($M = 0.2$, $SE = 0.2$) and nose ($M = 12.7$, $SE = 2.9$) region increased significantly ($p < 0.01$ and $p < 0.05$, respectively) after faces were inverted ($M = 1.2$, $SE = 0.3$, and $M = 22.7$, $SE = 2.5$, respectively). This interaction could also be explained by the fact that, for upright faces, first fixations landed mostly on the eye region, rather than the mouth and nose region. However, when faces were inverted, the number of the first fixations landed on the eye region became equivalent ($p > 0.05$) with that on the nose region. No other two-way interactions were significant.

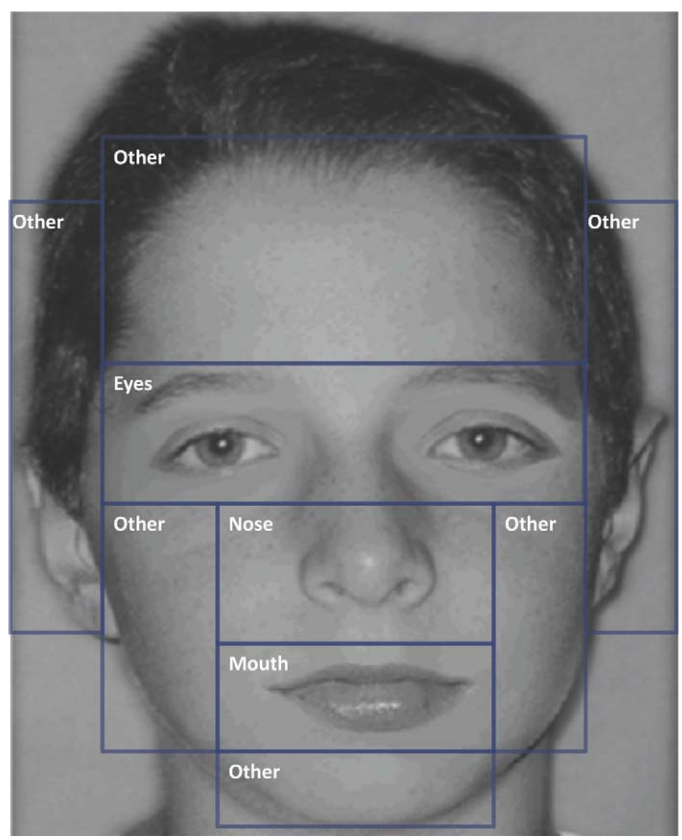

Figure 7. Areas of interest (see Henderson et al., 2005, as a reference). Left and right eyes are coded together into the eyes area, and every area except for eyes, nose and mouth are collapsed into the other area.

Viewing time proportion

The location and duration of every fixation were recorded during the change detection process. Heat maps were generated based on the average fixation time at each location (Figure 9). Heat maps provide a direct visualization of the looking behaviors of the participants. The hotter color on the heat map indicates a larger proportion of time spent in the specific location. Visual inspection of the heat maps indicated that, when making discriminations in upright faces, participants spent most of their time viewing the eye region of the face. However, when making discriminations in an

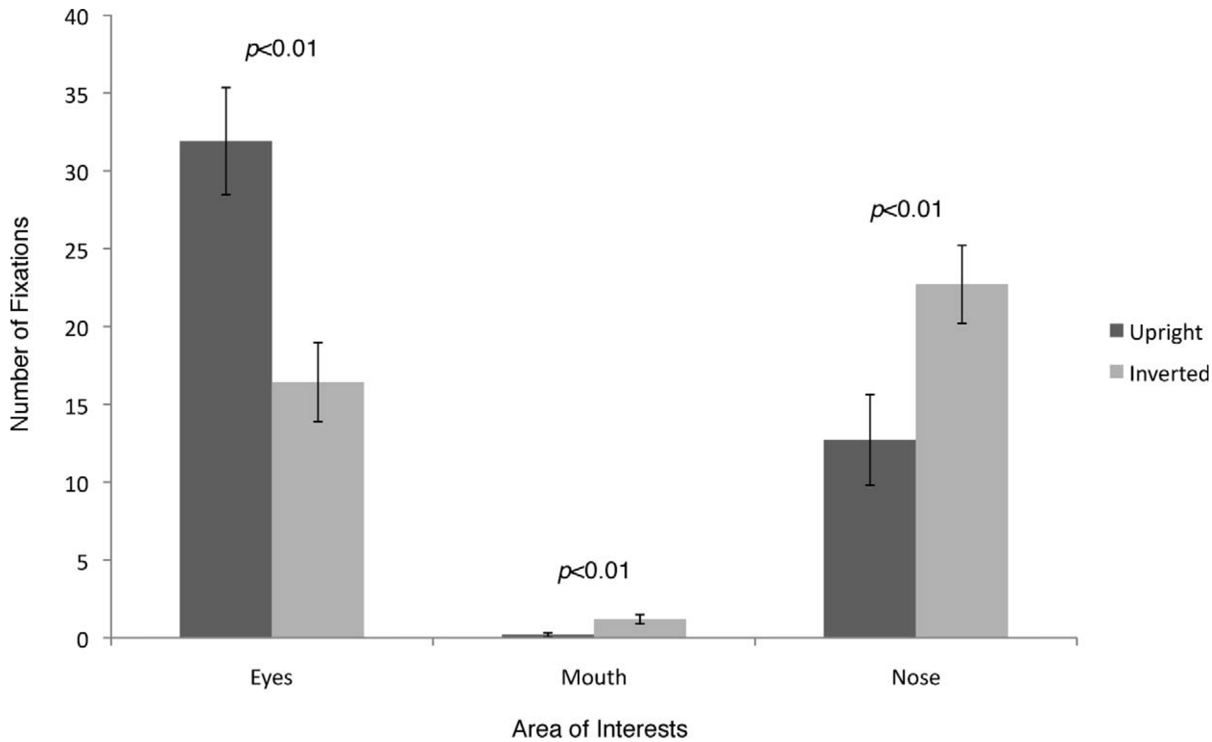

Figure 8. First fixation location distribution when processing upright and inverted faces. All the locations were coded into four areas of interest, but only the first fixations landing on the eyes, nose and mouth regions were analyzed.

inverted face, their eye fixations were distributed across a wider area, including the mouth region of the face. Due to the fact that different amounts of time were used from trial to trial to detect changes in the face pictures, fixation time in each region was calculated as a proportion of total fixation time for that trial. The Viewing Time Proportion in the areas of interest of eyes, nose and mouth was analyzed across Orientation, Change Type and Region. The results showed that, overall, the main effect of the distribution of Viewing Time Proportion was significant, $F(2, 17) = 30.31$, $p < 0.01$, indicating that participants spent most of the time looking at the eye region ($M = 0.42$, $SE = 0.03$). The time spent looking at the mouth ($M = 0.14$, $SE = 0.02$) and nose region ($M = 0.18$, $SE = 0.03$) did not reliably differ. Moreover, the interaction between the distribution of Viewing Time Proportion and Orientation was significant, $F(2, 17) = 17.65$, $p < 0.01$. Participants spent significantly more time ($p < 0.01$) looking at the eyes for upright faces ($M = 0.49$, $SE = 0.04$) than inverted faces ($M = 0.34$, $SE = 0.03$), but spent less time ($p < 0.01$) looking at the mouth ($M = 0.10$, $SE = 0.02$) for upright faces than inverted faces ($M = 0.19$, $SE = 0.02$). The difference was also significant for the viewing time for nose ($p < 0.01$). Participants spent significantly more time looking at the nose on inverted faces ($M = 0.21$, $SE = 0.03$) than upright faces ($M = 0.14$, $SE = 0.03$). In addition, the interaction between the distribution of Viewing Time Proportion and Region was also significant, $F(2, 17) = 96.46$, $p < 0.01$. Participants spent significantly more time ($p < 0.01$) looking at the eyes to detect changes in the eye region ($M = 0.55$, $SE = 0.03$) than mouth region ($M = 0.30$, $SE = 0.03$), but spent less time ($p < 0.01$) looking at the mouth to detect changes in the eye region ($M = 0.05$, $SE = 0.01$) than mouth region ($M = 0.23$, $SE = 0.03$). The difference was also significant for the viewing time for nose, $F(1, 18) = 33.25$, $p < 0.01$, with participants spending less time looking at the nose to detect changes in the eye region ($M = 0.14$, $SE = 0.03$) than mouth region ($M = 0.21$, $SE = 0.03$).

Accuracy and location of the last fixation prior to response

For all the trials, the location of the last fixation was collected and coded into the four areas of interest. This measurement provides information of which location on the picture is processed before a response is made. It's informative especially under the response-contingent paradigm because the relationship between the last fixation location and performance could be tested. The change detection accuracy was conditionalized according to whether the last fixation was located on the region of change (on-target) or off the region of change (off-target). Specifically, for eye trials, if the last fixation was on the eye region, this would be considered an on-target trial and if the last fixation landed on the

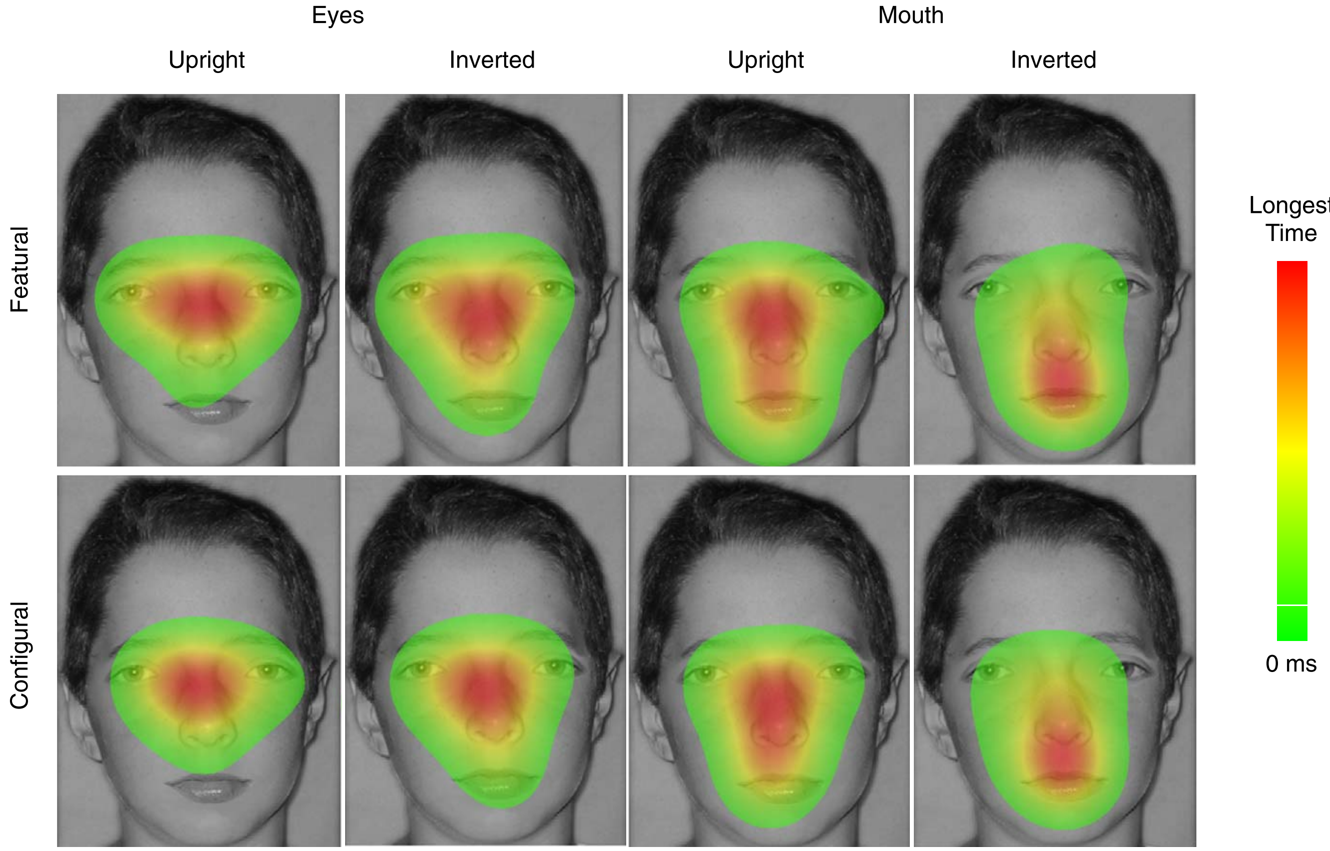

Figure 9. Heat maps for relative averaged fixation time within each trial across all eight conditions. All heat maps share the same scale ranging from 0 ms to the longest time within a certain condition. Hotter colors mean longer fixating time at this location. Only data from correct trials were included. The heat maps for the inverted faces were vertically flipped for the convenience of comparison.

nose, mouth or "other" regions, this would be considered an off-target trial. Similarly, for mouth trials, if the last fixation was on the mouth region, this would be considered an on-target trial and if the last fixation landed on the eyes, nose or "other" regions, this would be considered an off-target trial. The analysis showed a significant On/Off-target main effect, $F(1, 20) = 47.32$, $p < 0.01$, such that if the last fixation was on-target ($M = 0.88$, $SE = 0.02$), the performance was significantly higher than if the last fixation is off-target ($M = 0.63$, $SE = 0.04$). More importantly, the interaction between Orientation and On/Off-target was also significant, $F(1, 20) = 7.13$, $p < 0.05$. As shown in Figure 10, the difference in accuracy between on-target ($M = 0.83$, $SE = 0.03$) and off-target fixations ($M = 0.49$, $SE = 0.06$) in inverted faces was greater than the difference between on- ($M = 0.92$, $SE = 0.02$) and off-target ($M = 0.76$, $SE = 0.05$) fixations in upright faces.

Discussion

The purpose of this study was to investigate whether inverted faces elicited a qualitatively or quantitatively different mode of processing as indicated by eye movement patterns. The *Qualitative* view holds that inversion will lead to disproportionate decrements of performance for the processing of different kinds of face information. Whereas the *Configural/Featural* qualitative view maintains that inversion will differentially impair configural information in a face relative to its featural information, the *Regional* qualitative view argues that information in the mouth area will be differentially compromised compared to information in the eye region. In contrast, the *Quantitative* view proposes that inversion will impair the processing of different kinds of information in the same way. Moreover, same sets of cues will be processed when viewing upright and inverted faces.

The results from the current study supported the *Regional* qualitative view of inversion. Both the behavioral and eye movement evidence indicated that the qualitative distinction was *not* between featural versus configural information, but between information contained in the eye region versus information in the mouth region. First, task performance showed a larger inversion effect for detecting changes in the mouth region than in the eye region. However, no disproportional inversion effect was found for configural and

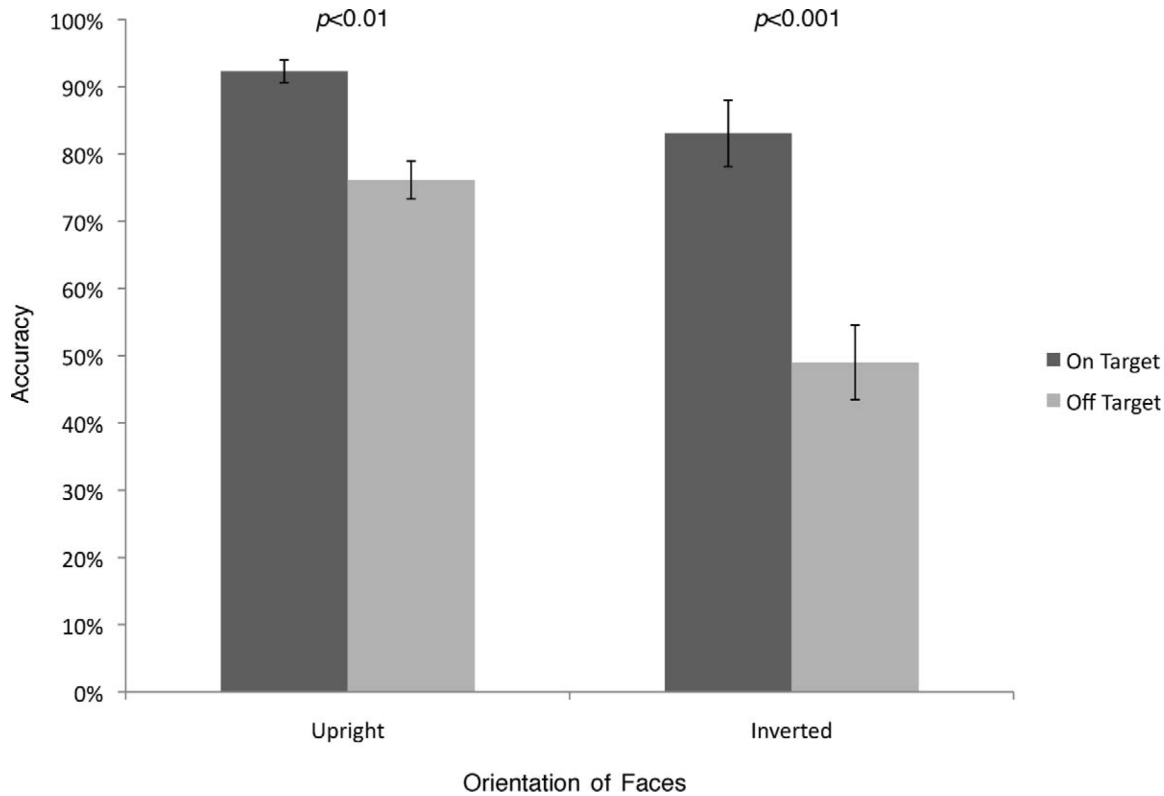

Figure 10. Last fixation location and performance. On-target designates the last fixation landed on the eyes (mouth) on the eye (mouth) trial, while off-target designates the opposite. Error bars refer to standard error.

featural changes. Second, more detailed visual analysis was required in the inverted than upright mouth trials as indicated by the greater number of saccades and smaller saccadic distances, which was not the case for eye trials. However, no such difference was found when detecting configural and featural changes in faces. Third, analysis of the first fixation location and the distribution of time in the four areas of interest showed that more processing was devoted to the mouth and nose region in the inverted face relative to the upright face. Finally, in both orientations, detection was more accurate when the last fixation was in the region of change. This effect was more pronounced for inverted than upright faces.

It should be noted that a bias of making "different" responses was present in this study due to the larger number of "different" trials than "same" trials. The reason for this type of design was that, for the investigation of eye movement in change detection task, "different" trials provide richer information than same trials. "Same" trials were used only as catch trials and were not entered into the analysis. Despite the tendency to make a "different" response, participants nevertheless failed to detect changes in the inverted mouth condition on 45% of the trials. The disproportionate inversion effect replicated the results from the study by Tanaka et al. (2009) in which the same number of "same" and "different" trials were used. Therefore, the findings from the current study should not be undermined by the response bias.

The disproportionate inversion effect

One of the arguments between the *Qualitative* and *Quantitative* view lies in whether inversion leads to disproportionate performance decrement of the processing of different kinds of face information. In the current study, we found that the orientation of the face interacted with the location of change regardless of the change type. Changes in the mouth region were more difficult to detect in the inverted than upright orientation. This pattern of performance decrement was not observed for the eye region. The finding that Orientation interacted with Region rather than Change Type replicates the study by Tanaka et al. (2009). In that study, the authors argued that when eye and mouth spacing are independently manipulated and equated for difficulty with featural eye and mouth changes, information in the mouth region suffers disproportionately during inversion than information in the eye region. McKone and Yovel (2009) similarly argued that the manipulation of the size or shape of the local facial features (instead of color change, feature substitution, etc.) yielded an inversion effect of equal magnitude when compared to configural changes. The manipulation of this study was strictly with respect to the size of the features, therefore it was not surprising that the interaction was significant between Region and Orientation, but not between Change Type and Orientation. However, it should be noted that only the horizontal distance between the eyes was manipulated. According to the literature, while the processing of the horizontal distances between the eyes are relatively unaffected by inversion, inversion effect was found when vertical displacement of the eyes and eyebrows were manipulated (Crookes & Hayward, 2012; Goffaux & Dakin, 2010; Goffaux & Rossion, 2007; Sekunova & Barton, 2008). Therefore, it could partly be the reason that the current study did not find an inversion effect for the configural eye condition. In short, the interaction between Region and Orientation, which reflects the disproportionate performance decrement (rather than a global performance decrement) between the information processing of the eye and mouth region, indicates the difference is qualitative. However, due to the insignificant interactions between Change Type and Orientation, only the *Regional* view was supported by the current study, and the *Configural*/*Featural* view was not supported.

The functional role of eye movement in face perception

The eye movement data confirmed that upright and inverted faces are processed differently. From the first fixation, visual saccades were directed to different locations when participants were viewing upright and inverted faces. When processing upright faces, the largest number of the initial fixations was directed to eye region, and the second greatest percentage to the nose region. However, when processing inverted faces, although most of the initial fixations were still directed to the eye and nose region, the numbers became equivalent. This evidence shows that even in the early period of the processing, participants started to use different initial cues to process upright and inverted faces.

After the first fixation landed on the face, more time was spent in processing the information in the eye region when viewing upright faces than when viewing inverted faces. For the information processing in the mouth region, however, the opposite trend was found, with more time spent in processing the information in the mouth region of inverted faces than upright faces. The pattern clearly shows a shift to the mouth region brought by inversion. Barton et al. (2006) also found similar results, with fixations redistributed to the mouth and lower face region when participants were viewing an inverted face. They attributed this effect to the disabled processing mechanism with which structural information is usually extracted globally and efficiently. With this mechanism impaired when processing inverted faces, fixations must be deployed to each region specifically, especially to the less-salient lower face region. This is also a plausible explanation with respect to results of the current study where inversion forced a redirection of the fixations from the more-salient eye region to the less-salient mouth region.

The response-contingent method employed in this study allowed us to meaningfully interpret the functional value of last fixation before response. The last fixation was the most important fixation for change detection. For all the correct trials, responses should be made right after the change was detected, so it was reasonable to deduce that the last fixation should be the one fixation that spotted the change. The last fixation location was further categorized according to whether it landed on the region of change (on-target) versus off the region of change (off-target). The results of the last fixation showed that when viewing upright faces, measures of task performance were less sensitive to whether the last fixation was "on" or "off" the target. However when the faces were inverted, measures of task performance were more sensitive to whether last fixation location was "on" or "off" target. The last fixation was used as a marker of the end of visual processing. Therefore, the information extracted by this last fixation should include relevant information for decision-making. The results suggested that, for upright faces, participants were still able to detect changes outside the area of foveated vision, but for inverted faces, if the last fixation landed on the area of change, there was a high probability of detecting the change. However, if the last fixation landed outside the area of change, the probability of detecting the changes was reliably less.

Why does the last fixation predict successful change detection more for inverted faces than upright faces? According to the *Perceptual Field* theory by Rossion (2009), when processing upright faces, humans have the relatively large "perceptual field" in which facial information can be extracted in both the foveal and parafoveal regions of fixation. With an expanded perceptual field, changes in the periphery can be detected even when they are not foveated and hence, eye fixations would not necessarily be predictive of performance. However, when a face is inverted, the perceptual field shrinks in size and only information in the fovea is processed. With a reduced perceptual field, the likelihood of detecting a change is significantly increased if the area of change is foveated on and hence, eye tracking would be more strongly correlated with performance. This logic supports the view that inversion would lead to qualitatively, rather than quantitatively, different face processing strategies.

According to the perceptual field theory, it is plausible to infer that when the face is upright, the location of the last fixation is not critical because the perceptual field is broad and encompassing the entire face. However, when the face is inverted, the perceptual field shrinks and its span only extends to a limited area, probably restricted to the processing of single features. Hence, if participants are not fixated on the critical region, the probability of detection is quite low.

Conclusion

The face inversion effect is important for the understanding of face processing. The *Qualitative* and *Quantitative* debate is one of the open topics in this field that requires more research efforts. The current study attempted to answer this question by revealing the different eye movements used during the processing of upright and inverted faces. Existing eye movement studies on the face inversion effect (e.g., Barton et al., 2006; Williams & Henderson, 2007) could not resolve the *Qualitative* and *Quantitative* debate. The current study employed the change detection paradigm to study eye movement during the processing of upright and inverted face. The results showed that: (a) inversion impaired information processing in the mouth region more than the eye region, (b) inversion led to a more deliberate scanning pattern characterized by a longer response time, a larger number of saccades, and smaller saccade distance when detecting changes in the mouth region, (c) different sets of cues were used when processing upright and inverted faces. For inverted faces, more cues in the mouth and nose region of the face were processed and (d) when detecting changes in inverted faces, if the last fixation before the response landed on the region where the changes occurred, the changes were more likely to be detected. This effect was smaller for upright faces. All this evidence supports the view that face inversion led to a qualitatively different type of face processing that selectively disrupts information in the mouth region.

Keywords: face inversion, qualitative, change detection, eye movement

Acknowledgments

This research was supported by grants from the Chinese Scholarship Council, the Temporal Dynamics of Learning Center (NSF Grant #SBE-0542013) and the National Sciences and Engineering Research Council of Canada. We would also like to thank Professor Michael Masson in Department of Psychology of University of Victoria for his coordination on the eye-tracking laboratory and Marnie Jedynak for her technical assistance.

Commercial relationships: none.
Corresponding author: Buyun Xu.
Email: xubuyun@uvic.ca.
Address: Department of Psychology, University of Victoria, Victoria, British Columbia, Canada.

Footnotes

[1] McKone and Yovel (2009) pointed out that the magnitude of featural and configural inversion effect is also determined by how a "feature" is defined. While a feature defined by the size and shape properties produces inversion effects comparable to configural changes, a feature defined by the color or luminance of a face part are orientation invariant and therefore produce weak inversion effects (Barton et al., 2001; Leder & Bruce, 2000).

[2] Due to the failure of two participants in detecting changes in all the trials of certain conditions, their data was not entered into the repeated-measures ANOVA where only correct trials were included. This applied to the analysis of Total Response Time, Number of Saccades, Saccade Distance and Viewing Time Distributions.

[3] The eye movement data for one participant was excluded from analysis because most of the fixations were off the face. However, the behavioral data for this participant (i.e., accuracy and total response time) was retained.

References

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724–732.

Barton, J. J., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology, 92,* 527–549.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35,* 1089–1105.

Bindemann, M., Scheepers, C., & Burton, A. M.

(2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision, 9*(2): 7, 1–16, http://www.journalofvision.org/content/9/2/7, doi:10.1167/9.2.7. [PubMed] [Article].

Bombari, D., Mast, F. W., & Lobmaier, J. S. (2009). Featural, configural, and holistic face-processing strategies evoke different scan patterns. *Perception, 38,* 1508–1521.

Boutet, I., & Faubert, J. (2006). Recognition of faces and complex objects in younger and older adults. *Memory & Cognition, 34,* 854–864.

Bukach, C. M., Le Grand, R., Kaiser, M., Bub, D., & Tanaka, J. W. (2008). Preservation of mouth region processing in two cases of prosopagnosia. *Journal of Neuropsychology, 2,* 227–244.

Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science, 11,* 419–433.

Crookes, K., & Hayward, W. G. (2012). Face inversion disproportionately disrupts sensitivity to vertical over horizontal changes in eye position. *Journal of Experimental Psychology: Human Perception & Performance.*

de Heering, A., Rossion, B., Turati, C., & Simion, F. (2008). Holistic face processing can be independent of gaze behavior: Evidence from the composite face illusion. *Journal of Neuropsychology, 2,* 183–195.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29,* 159–170.

Goffaux, V., & Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Psychology, 1.*

Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 995–1002.

Heisz, J. J., & Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision, 8*(1): 9, 1–10, http://www.journalofvision.org/content/8/1/9, doi:10.1167/8.1.9. [PubMed] [Article].

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33,* 98–106.

Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science, 19,* 998–1006.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology: Section A, 53,* 513–536.

Leder, H., & Carbon, C. C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97,* 19–29.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30,* 73–83.

Malcolm, G., Leung, C., & Barton, J. (2004). Regional variation in the inversion effect for faces: Different patterns for feature shape, spatial relations, and external contour. *Perception, 33,* 1221–1231.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16,* 778–797.

Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces? *Developmental Science, 9,* 270–277.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Quinn, P. C., & Tanaka, J. W. (2009). Infants' processing of featural and configural information in the upper and lower halves of the face. *Infancy, 14,* 474–487.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62, 8,* 1457–1506.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53,* 245–277.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47,* 25–57.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13,* 499–505.

Riesenhuber, M., Jarudi, I., Gilad, S., & Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B, 271,* 448–450.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128,* 274–289.

Rossion, B. (2009). Distinguishing the cause and consequences of face inversion: The perceptual field hypothesis. *Acta Psychologica, 132,* 300–312.

Rossion, B., Kaiser, M. D., Bub, D., & Tanaka, J. W. (2009). Is the loss of diagnosticity of the eye region of the face a common aspect of acquired prosopagnosia? *Journal of Neuropsychology, 3,* 69–78.

Sekunova, A., & Barton, J. J. (2008). The effects of face inversion on the perception of long-range and local spatial relations in eye and mouth configuration. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1129.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46A,* 225–245.

Tanaka, J. W., Kaiser, M. D., Bub, D., & Pierce, L. J. (2009). Generalized impairment of featural and configural information in the lower region of the face through inversion. *Journal of vision, 9*(8):531, http://www.journalofvision.org/content/9/8/531, doi:10.1167/9.8.531. [Abstract].

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25,* 583–592.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471–491.

van Belle, G., de Graef, P., Verfaillie, K., Rossion, B., & Lefevre, P. (2010a). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision, 10*(5):10, 1–13, http://www.journalofvision.org/content/10/5/10, doi:10.1167/10.5.10. [PubMed] [Article].

van Belle, G., Ramon, M., Lefèvre. P., & Rossion, B. (2010b). Fixation patterns during recognition of personally familiar and unfamiliar faces. *Frontiers in Psychology, 1,* 1–8.

Williams, C. C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory & Cognition, 35,* 1977–1985.

Wolf, J. M., Tanaka, J. W., Klaiman, C., Klaiman, C., Cockburn, J., Herlihy, L., et al. (2008). Specific impairment of face-processing abilities in children with autism spectrum disorder: Using the Let's Face It! skills battery. *Autism Research, 6,* 329–340.

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology, 19,* 739–749.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Yovel, G., & Duchaine, C. B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience, 18,* 580–593.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron, 44,* 889–898.

Zelinsky, G. J. (2001). Eye movements during change detection: Implications for search constraints, memory limitations, and scanning strategies. *Attention, Perception, & Psychophysics, 63,* 209–225.