Vision Research 138 (2017) 78–85

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

3D faces are recognized more accurately and faster than 2D faces, but with similar inversion effects

Z.H.D. Eng [a,*], Y.Y. Yick [a], Y Guo [b], H. Xu [a], M. Reiner [c], T.J. Cham [d], S.H.A. Chen [a,e,f,*]

[a] Psychology, School of Social Sciences, Nanyang Technological University, Singapore
[b] Institute for Media Innovation, Nanyang Technological University, Singapore
[c] The Virtual Reality and Neurocognition lab, Faculty of Education in Science and Technology, Technion, Israel Institute of Technology, Israel
[d] School of Computer Science and Engineering, Nanyang Technological University, Singapore
[e] Centre for Research and Development in Learning (CRADLE), Nanyang Technological University, Singapore
[f] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 3 August 2016
Received in revised form 26 April 2017
Accepted 1 June 2017
Available online 4 August 2017

Number of Reviews = 2

Keywords:
Face recognition
Holistic processing
Featural processing
2D
3D
Face inversion effect
Stereoscopic images

ABSTRACT

Recognition of faces typically occurs via holistic processing where individual features are combined to provide an overall facial representation. However, when faces are inverted, there is greater reliance on featural processing where faces are recognized based on their individual features. These findings are based on a substantial number of studies using 2-dimensional (2D) faces and it is unknown whether these results can be extended to 3-dimensional (3D) faces, which have more depth information that is absent in the typical 2D stimuli used in face recognition literature. The current study used the face inversion paradigm as a means to investigate how holistic and featural processing are differentially influenced by 2D and 3D faces. Twenty-five participants completed a delayed face-matching task consisting of upright and inverted faces that were presented as both 2D and 3D stereoscopic images. Recognition accuracy was significantly higher for 3D upright faces compared to 2D upright faces, providing support that the enriched visual information in 3D stereoscopic images facilitates holistic processing that is essential for the recognition of upright faces. Typical face inversion effects were also obtained, regardless of whether the faces were presented in 2D or 3D. Moreover, recognition performances for 2D inverted and 3D inverted faces did not differ. Taken together, these results demonstrated that 3D stereoscopic effects influence face recognition during holistic processing but not during featural processing. Our findings therefore provide a novel perspective that furthers our understanding of face recognition mechanisms, shedding light on how the integration of stereoscopic information in 3D faces influences face recognition processes.

1. Introduction

Face recognition is an innate ability that is essential to our daily social interactions. Most of us are able to recognize and distinguish faces instantly, suggesting that faces are a special category of our visual expertise (Heisz, Watter, & Shedden, 2006; Maurer, Grand, & Mondloch, 2002; Richler, Mack, Gauthier, & Palmeri, 2009; Taubert, Apthorp, Aagten-Murphy, & Alais, 2011). Research largely supports that a face is processed holistically such that individual features are integrated and represented as a whole (Behrmann, Richler, Avidan, & Kimchi, 2014; Diamond & Carey, 1986; Maurer et al., 2002; Tanaka & Farah, 1993). Such holistic processing has been shown to be more important for recognition of faces than for other objects, as the resulting Gestalt representations help us make sense of the visual information and perceive different identities (Behrmann et al., 2014). This is in contrast to featural processing, where the visual stimulus is recognized based on its individual components (e.g., eyes, mouth, nose, face contour, colour, brightness, etc.) rather than as a whole (Diamond & Carey, 1986; Tanaka & Farah, 1993). Featural processing occurs when faces are inverted and subsequently processed more similarly to objects based on their individual features instead, hence leading to upside-down faces not being recognized as Gestalt representations with inherent identities (Farah, Wilson, Drain, & Tanaka, 1995; Rossion & Gauthier, 2002; Tanaka & Farah, 1993; Tanaka & Sengco, 1997). This is referred to as the “face inversion effect”,

* Corresponding authors at: Psychology, School of Social Sciences, Nanyang Technological University, 14 Nanyang Drive, HSS-04-19, Singapore 637332, Singapore (S.H.A. Chen).

E-mail addresses: engz0001@e.ntu.edu.sg (Z.H.D. Eng), annabelchen@ntu.edu.sg (S.H.A. Chen).

http://dx.doi.org/10.1016/j.visres.2017.06.004
0042-6989/

whereby inversion deteriorates our face recognition ability drastically compared to the recognition of non-face stimuli (Yin, 1969).

The existing knowledge about face processing, however, is based on studies that examined 2-dimensional (2D) faces presented on computer screens and lacking the visual depth information inherent in real life faces. In contrast, 3-dimensional (3D) images provide greater depth and visual details (Häkkinen et al., 2008; Lambooij, IJsselsteijn, Bouwhuis, & Heynderickx, 2011), thereby leading to richer information of both individual features as well as the spatial interrelationship between them (configural information) (Schwaninger, Ryf, & Hofer, 2003), and thus a more "comprehensive" Gestalt representation. These enhanced featural and spatial details in 3D faces also provide additional visual information that could help to make the stimuli more closely resemble the real-life perceptions that our visual systems are attuned to. Based on these premises, it is expected that 3D details would provide an advantage over 2D images during face recognition for both holistic and featural processing.

To date, no published studies have examined the manner in which 3D stereoscopic faces influence the mechanisms of holistic and featural processing in face recognition. This study therefore aims to address this research gap by comparing the recognition of 3D stereoscopic and 2D faces in a classic face inversion paradigm (Tanaka & Farah, 1993; Yin, 1969). We examined to what extent the findings of the face inversion effect on 2D faces could be generalized to 3D faces, and aimed to understand processing that is involved more heavily for 3D faces. It is hoped that the findings from this study not only extend our current understanding of face recognition but also provide novel perspectives for research ideas that are enabled by the advancement of 3D technology.

It has yet to be established how the proposed advantages of enhanced visual details in 3D influence the mechanisms underpinning the holistic and featural processing involved in face recognition. Existing literature suggests that faces are processed over three different stages (Maurer et al., 2002). During the first stage, first-order processing occurs based on the general organization of the face's features (i.e. two eyes, above a nose, above a mouth), for the initial face detection. Subsequently, holistic processing occurs in the second stage where facial features are integrated to form a Gestalt representation. At the third stage, second-order processing takes place in which the variance between faces is analyzed, such as the distance between the eyes, to form accurate and distinct face representations (Diamond & Carey, 1986; Freire, Lee, & Symons, 2000; Taubert et al., 2011). Holistic face recognition is therefore based on how basic attributes are spatially arranged to form the prototypical representation of a face (Diamond & Carey, 1986; Taubert et al., 2011). When faces are presented upside-down, however, it disrupts the spatial relationship among the facial features (the first-order information), slowing down face detection and impairing holistic face processing consequently (Farah, Wilson, Drain, & Tanaka, 1998; Farah et al., 1995; Maurer et al., 2002; Sekuler, Gaspar, Gold, & Bennett, 2004; Tanaka & Farah, 1993; Yin, 1969). As a result, inverted faces are recognized as an amalgamation of facial parts rather than as a congruent face.

A substantial number of studies have manipulated the upright or inverted orientation of faces as a reliable method of eliciting holistic or featural processing (Itier & Taylor, 2002; Leder & Carbon, 2006; Rossion et al., 1999; Sekuler et al., 2004; Tanaka & Farah, 1993; Tanaka & Sengco, 1997; Taubert et al., 2011). It is typically shown that holistic processing contributes to greater face recognition accuracy and faster response time, as it facilitates the formation of a coherent representation of a face (Itier & Taylor, 2002; Jacques, D'Arripe, & Rossion, 2007; Rossion et al., 1999; Tanaka & Farah, 1993; Taubert et al., 2011). On the other hand, inversion leads to an increase in cognitive demand due to the disruption of first-order information (Behrmann et al., 2014; Maurer et al., 2002; Rock, 1974), resulting in lower accuracy and slower reaction times during featural processing (Itier & Taylor, 2002; Jacques et al., 2007; Jiang, Dricot, Blanz, Goebel, & Rossion, 2009; Maurer et al., 2002; Rossion et al., 1999; Tanaka & Farah, 1993; Tanaka & Sengco, 1997; Taubert et al., 2011).

Results from event-related potential (ERP) studies focusing on the face-sensitive N170 modulation have corroborated the behavioral findings. In particular, the amplitude and latency of N170 are thought to index the degree and onset of early structural encoding of faces (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Bentin, Deouell, & Soroker, 1999; Eimer, Kiss, & Nicholas, 2010; Heisz et al., 2006; Itier & Taylor, 2002; Jacques et al., 2007; Maurer et al., 2002). Studies have consistently shown that N170 has a later onset for inverted faces compared to upright faces, supporting the notion of delayed processing speed for inverted faces (Heisz et al., 2006; Itier & Taylor, 2002, 2004; Jacques et al., 2007; Rossion & Gauthier, 2002; Rossion et al., 1999, 2000; Sadeh & Yovel, 2010). Moreover, these studies have also shown a larger N170 amplitude (more negative) for inverted faces compared to upright ones, suggesting more complex structural encoding for inverted relative to upright faces (Bentin et al., 1996; Eimer et al., 2010; Heisz et al., 2006; Itier & Taylor, 2002; Maurer et al., 2002; Rossion et al., 1999, 2000; Sekuler et al., 2004). Taken together, results from both behavioral and ERP studies provide convincing evidence that upright faces are associated with higher accuracy and shorter processing time due to less complex structural encoding compared to inverted faces (Eimer et al., 2010; Itier & Taylor, 2002; Jacques et al., 2007; Rossion et al., 1999, 2000). Here, we employed the face inversion paradigm to investigate whether such holistic and featural processes in 2D face recognition are similarly engaged during processing of 3D faces.

As aforementioned, 3D stereoscopic images provide greater depth and visual details compared to their 2D counterparts (Häkkinen et al., 2008; Lambooij et al., 2011). Therefore, it is expected that 3D faces would provide (i) enriched configural information between facial features which could be beneficial to holistic processing, and (ii) richer visual details of the individual facial parts which could be beneficial to featural processing (Liu, Collin, & Chaudhuri, 2000). In order to test for these differences and their effects on holistic and featural processing during face recognition, participants completed a delayed face-matching task for 2D and 3D faces that were presented upright or inverted.

Given the premise that 3D provides greater visual depth information, it is expected that the information providing first-order structure would thus be enhanced, thereby facilitating holistic processing. We hypothesized that 3D upright faces would be recognized faster and with greater accuracy than 2D upright faces. Similarly, it is expected that 3D would facilitate featural processing due to the increased richness of visual details in local information. Therefore, individual facial parts are surmised to be more easily discernible, leading to the prediction that 3D inverted faces would also be recognized faster and with greater accuracy than 2D inverted faces. For inverted faces, since first-order information is disrupted when faces are presented upside-down, we hypothesized that the inversion effects would be observed regardless of whether the faces are shown in 2D or 3D.

2. Methodology

2.1. Participants

Twenty-four undergraduates and two recent graduates, with normal or corrected-to-normal vision, were recruited from four local universities in Singapore – Nanyang Technological University, National University of Singapore, Singapore Management

University, and Singapore Institute of Management. One female participant was excluded from the study due to accuracy results falling below chance level (0.5) across all conditions. The remaining participants had a mean age of 24.5 years (11 males). Informed consent was obtained from all participants prior to the experiment. The study obtained ethics approval from the Institutional Review Board at Nanyang Technological University, and was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2. Stimuli

Fifty-two male front-view Chinese faces were used to form the 2D and 3D stimuli in the experiment. These faces were taken from the 3D face database in the Center of Signal Processing (CSP) in the school of Electrical and Electronic Engineering in Nanyang Technological University (© 2010 Gede Putra Kusuma, NTU). Every face was presented in each of the 4 conditions once (i.e. 2D upright, 2D inverted, 3D upright, 3D inverted), except in 'match' trials where the study face was repeated as the test face (see Fig. 3). Each face also appeared in an equal number of 'match' and 'nonmatch' trials, randomized among the four conditions. Hence, all faces appeared six times throughout the entire experiment – four times as the study face in each condition, and twice as the matching test face in two random conditions.

The 3D faces were created following the setup illustrated in Fig. 1. A face model was first translated close to the "Screen Z plane" during rendering. Following which, the face was rendered between two different planes – the "Near Z plane" and "Far Z plane" – to the rendering plane. In the 2D mode, only one image is projected; whereas in the 3D model, the monitor displays images from the left and right alternately. Due to the high refresh rate (200 Hz) of the monitor, viewers perceived the left and right images appearing simultaneously as one stimulus. More importantly, although the perception of faces in 2D is based on a symmetrical viewpoint, the perception of 3D faces is facilitated by left and right eye views being non-parallel (see Fig. 1b). In the 3D display mode, the left eye looks towards the "centre-right" direction while the right eye looks towards the "centre-left" direction instead. All stimuli were presented on a black background displayed using an Alienware laptop and viewed with a pair of 3D Vision wireless active shutter glasses (see Fig. 2).

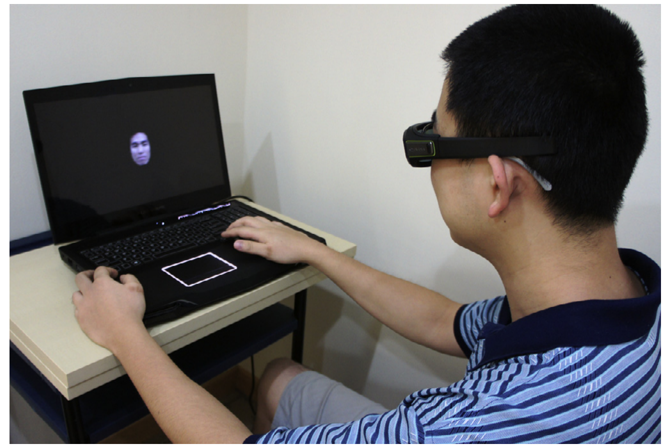


Fig. 2. Image of experimental set-up depicting participant equipped with 3D Vision wireless active shutter glasses. The experiment was conducted in a dark room to minimize reflections of surrounding objects.

Stimuli were set according to the following parameters, where image height and width measured 10 cm by 10 cm. The "Near Z", "Far Z" and "Screen Z" planes were rendered within the Normalized Devices Coordinate (NDC) space, which is a reference frame used within the application programme interface, OpenGL. The

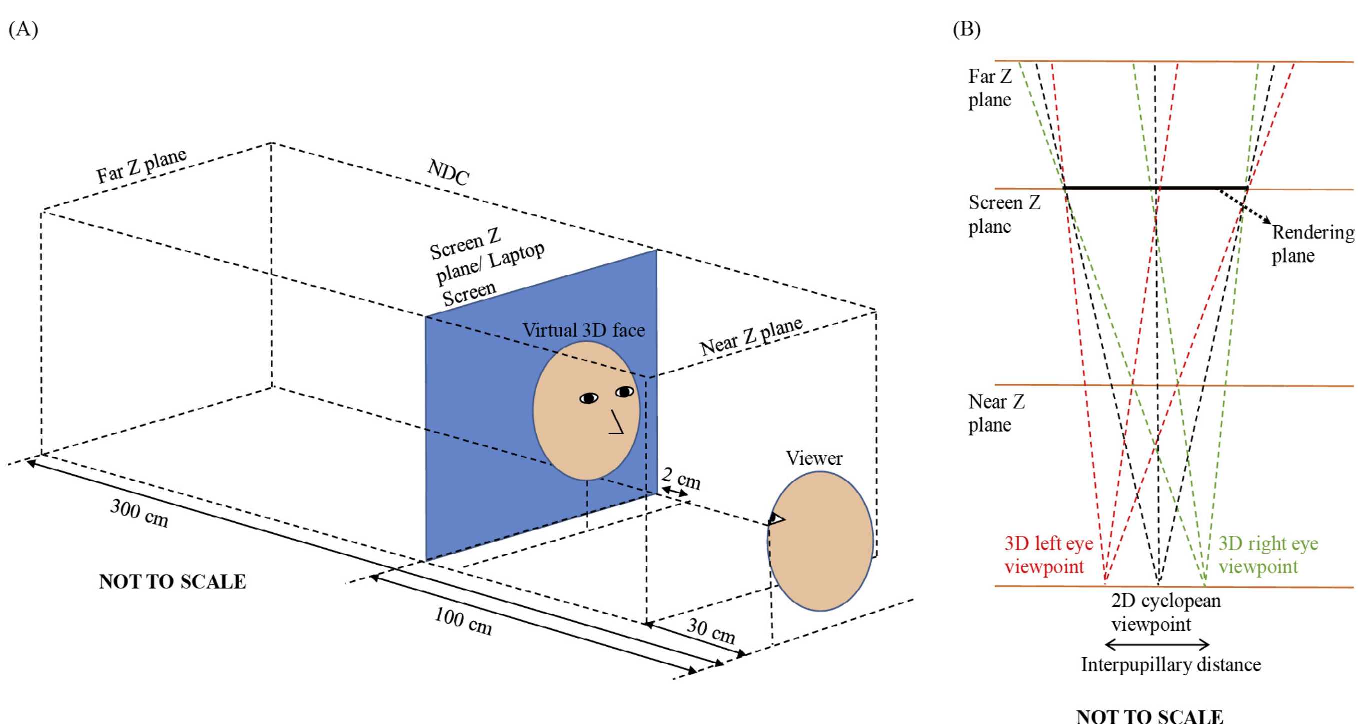


Fig. 1. Illustrations of projection of visual stimuli. (A) Image of set-up depicting coordinates of the Far Z, Screen Z and Near Z planes in the projection of the 3D face stimuli. (B) Top view of projection of visual stimuli presented in 2D and 3D. The solid line denotes the rendering plane. Dashed lines represent the cyclopean view in 2D display mode, while red dashed lines and green dashed lines represent the left and right eye views respectively in 3D display mode. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

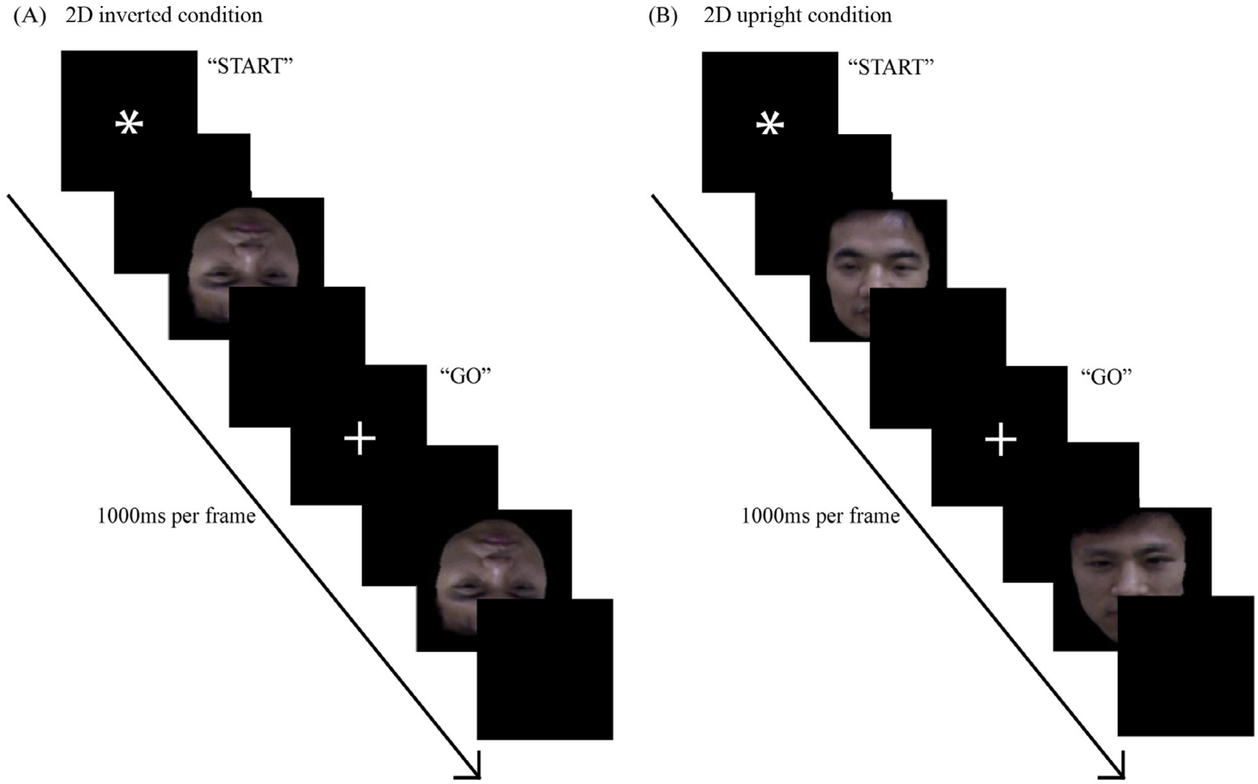


Fig. 3. Illustrations of trial sequences in the delayed face-matching task. The study image is presented following a fixation point accompanied by an auditory "START" cue. A second fixation point is presented next with an auditory "GO" cue, followed by the presentation of the test image where the participant is to make a binary "same/different" response as quickly and accurately as possible. (A) Example of trial sequence for 2D inverted condition with matching study and test faces. (B) Example of trial sequence for 2D upright condition with nonmatching study and test faces.

NDC space refers to a coordinate system that depicts the positions of plotted virtual points, wherein the visual coordinates of a stimulus seen from a real-world eye view may be mapped into this space to render 2D and 3D vector graphics (Hughes et al., 2013). Within the cube-shaped NDC space, the "Near Z-plane" was located away from the participant at a distance of 30 cm, while the "Far Z-plane" was located further away at 300 cm. The "Screen Z-plane" was placed at a distance of 100 cm, and the interpupillary distance was maintained at 5 cm. As a result, participants were able to experience the stereo effect that led to the perception of faces in 3D faces appearing stereoscopically at a distance of approximately 2 cm in front of the monitor screen. 2D faces were also viewed through the same set up, albeit appearing flat on the monitor screen (see Fig. 1).

2.3. Procedure

The experiment was conducted in a dark room in order to minimize reflections of surrounding objects that might interfere with the perception of the visual stimuli. Participants were seated approximately 100 cm away from the monitor screen and fitted with the 3D eyewear before the commencement of the experiment (see Fig. 2). Participants were given a short practice run which served as an opportunity to test the equipment, ensure that participants understood instructions, and receive verbal confirmation from participants that they were able to distinguish 2D and 3D facial stimuli prior to the start of the experiment. In each trial, participants were presented with two faces in succession that were matched in orientation (upright or inverted) and modality (2D or 3D). They were asked to indicate as quickly and accurately as possible if the test (second) face was a "match" or "non-match" with the study (first) face via a binary button press. Participants were not exposed to any of the face stimuli presented in the experimental conditions prior to the commencement of the test trials.

Each trial commenced with an audio signal "START" along with the presentation of a fixation cross in the center of the screen for 1000 ms. This was followed by a 1000 ms blank screen prior to the onset of the study face for 1000 ms. A blank screen was then presented for 1000 ms, followed by an audio signal "GO" along with a 1000 ms fixation cross. The fixation cross was then offset for 1000 ms (i.e. blank screen) before the onset of the test face for 1000 ms during which participants were required to make the binary "match" or "non-match" responses by pressing the left and right key respectively. A 1000 ms inter-stimulus blank screen is then presented before the onset of the next trial. Illustrations of the trial sequences are presented in Fig. 3.

Participants completed four blocks, each consisting of 13 trials for each of the four conditions (2D upright, 2D inverted, 3D upright, 3D inverted). All 52 trials in each block were fully randomized to avoid possible habituation (in particular becoming less perceptive to the 3D effect over time) and learning effects (e.g. developing a fixed strategy for the task). Each block lasted approximately seven minutes, with rest periods of 2 min between blocks to minimize fatigue.

2.4. Data analysis

Incorrect trials and correct trials that exceeded the 1000 ms response window were coded as outliers, as participants were able to execute delayed face recognition tasks successfully within similar time frames in previous studies (Caharel, Jiang, Blanz, & Rossion, 2009; Rossion et al., 1999; Taubert et al., 2011). In the current study, the highest mean RT for the current study was 705 ms (S.E. = 12) obtained from the 2D inverted face condition. Analyses of accuracy rates and RT were restricted to trials in which participants provided correct responses. A three-step analysis was conducted: (1) A 2-way repeated measure ANOVA of modality (2D/3D) by orientation (upright/inverted) was conducted for accuracy and reaction time separately. In the analysis of accuracy outcomes, a Friedman two-way analysis of variance of ranks was also conducted due to the non-normal distribution in at least one condition (see below); (2) Follow-up tests were conducted through planned contrasts with Bonferroni corrections to compare the effects of interest. In this analysis, paired-comparison *t*-tests were conducted to examine (i) the performance differences between 2D and 3D stimuli of the same orientation (2D upright vs 3D upright, 2D inverted vs 3D inverted) and, (ii) inversion effects (2D upright vs 2D inverted, 3D upright vs 3D inverted). Wilcoxon signed-rank tests were also run for the above comparisons pertaining to accuracy results in view of their non-normal distribution; and (3) to compare the extent of inversion effects between 2D and 3D conditions, subtraction scores were obtained by measuring differences in accuracy and reaction time separately (e.g. 2D upright accuracy score minus 2D inverted accuracy score). Paired-comparison *t*-tests were then conducted between the subtraction scores of 2D and 3D conditions for both accuracy and reaction time.

3. Results

The behavioral performances are displayed in Fig. 4. Paired-samples *t*-tests showed that participants performed significantly above chance level for all 4 conditions [$ts > 52.52$, $ps < 0.001$].

Accuracy results were non-normally distributed in two of the four conditions (2D upright and 3D upright conditions), therefore we conducted a non-parametric Friedman test in addition to the initial planned ANOVA. The pattern of results for both tests were similar – the Friedman test indicated a significant difference among the accuracy scores across the four conditions [χ^2 (3) = 50.72, $p < 0.001$, Kendall's W = 0.68]. This result was confirmed by ANOVA which revealed significant main effects of orientation [$F(1,\ 24) = 58.56$, $p < 0.001$, $\eta_p^2 = 0.71$] and modality [$F(1,\ 24) = 4.71$, $p = 0.040$, $\eta_p^2 = 0.16$] on accuracy. There was no significant interaction between the effects of modality and orientation on accuracy [$F(1, 24) = 1.28$, $p = 0.269$, $\eta_p^2 = 0.05$].

Post hoc multiple comparisons were conducted following ANOVA. To account for the non-normal distributions of accuracy data, Wilcoxon signed-rank tests with multiple comparisons were conducted (Bonferroni corrected p-values = 0.0125). These results were consistent with the results from Bonferroni-corrected paired-samples *t*-tests. Both parametric and non-parametric results are reported in Table 1. As predicted, 3D upright faces were recognized with significantly greater accuracy than 3D inverted faces. Consistent with previous literature, 2D upright faces elicited significantly greater accuracy than 2D inverted faces. Interestingly, it was found that 3D upright faces were recognized with significantly greater accuracy than 2D upright faces. However, no differences were found between 3D inverted and 2D inverted face conditions.

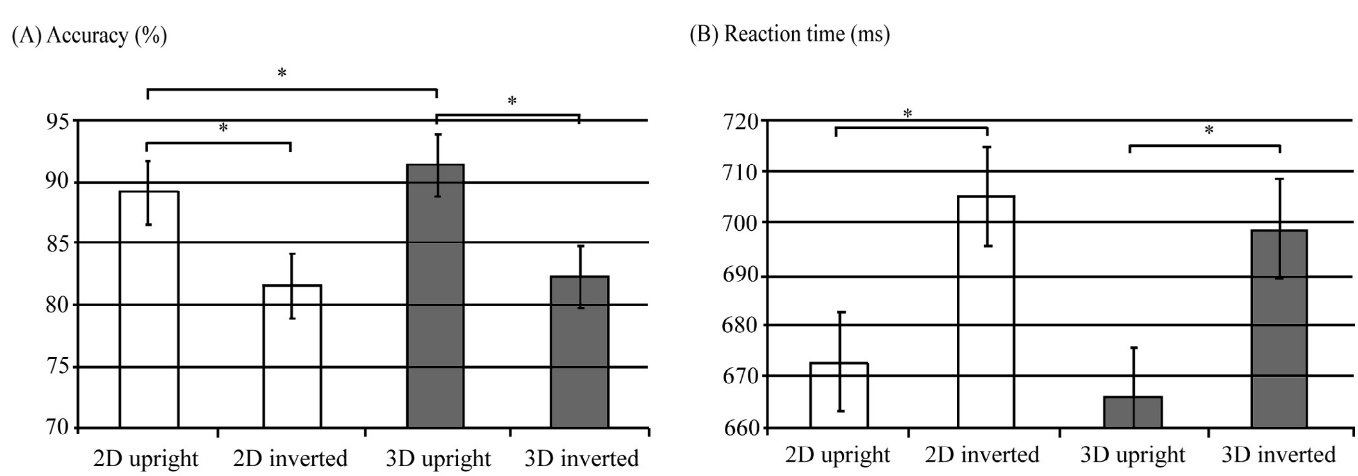
Reaction time data for all conditions was found to be normally distributed. The ANOVA revealed significant main effects of orientation [$F(1, 24) = 79.34$, $p < 0.001$, $\eta_p^2 = 0.77$], and modality [$F(1, 24) = 4.40$, $p = 0.047$, $\eta_p^2 = 0.16$] on reaction time performance. There was no significant interaction between the effects of modality and orientation on reaction time [$F(1,\ 24) = 0.01$, $p = 0.933$, $\eta_p^2 = 0.00$].

Paired-samples *t*-tests (Bonferroni corrected p-values = 0.0125) were conducted to compare reaction time performances between the different conditions (see Table 1). It was found that the reaction time performance for 3D upright faces was significantly faster than 3D inverted faces. The same pattern was observed for the comparison between 2D upright and 2D inverted faces. It was also found that there were no significant differences between the 3D upright and 2D upright face conditions, as well as between the 3D inverted and 2D inverted face conditions.

Finally, to determine the extent of inversion effects in 2D and 3D stimuli, we compared the differences in performances between upright and inverted conditions of 2D and 3D stimuli [Accuracy: (3D upright – 3D inverted) vs. (2D upright – 2D inverted); RT: (3D upright – 3D inverted) vs. (2D upright – 2D inverted)]. No significant differences were found, suggesting that the inversion effects did not differ reliably between 2D and 3D stimuli for both accuracy [$t(24) = 1.13$, $p = 0.269$, $d = 0.23$] and reaction time [$t(24) = 0.08$, $p = 0.933$, $d = 0.02$].

4. Discussion

The current study examined the differential processes between 2D and 3D face recognition using the face inversion paradigm. Our



	2D upright	2D inverted	3D upright	3D inverted
Mean (S.E.)	89.23 (0.84)	81.62 (0.75)	91.38 (0.81)	82.23 (0.80)

	2D upright	2D inverted	3D upright	3D inverted
Mean (S.E.)	672.78 (12.59)	705.39 (12.28)	665.77 (12.39)	698.83 (12.13)

Fig. 4. Behavioral results. (A) Mean accuracy for correct responses (±standard error) (B) Mean reaction time for correct responses (±standard error). *$p < 0.010$.

Table 1
Results of parametric and non-parametric pairwise comparisons for accuracy and reaction time.

Pairwise comparisons	Paired-sample *t*-tests#			Wilcoxon signed-rank tests#		
	t	*p*	*d*	*Z*	*p*	*r*
Accuracy:						
3D upright vs 3D inverted	7.37	<0.001	1.47	−4.21	<0.001	−0.60
2D upright vs 2D inverted	5.71	<0.001	1.14	−3.89	<0.001	−0.55
2D upright vs 3D upright	−3.31	0.003	0.66	−2.73	0.006	−0.39
2D inverted vs 3D inverted	−0.54	0.597	0.11	−0.54	0.592	−0.08
Reaction Time:						
3D upright vs 3D inverted	−7.04	<0.001	1.41	–	–	–
2D upright vs 2D inverted	−7.53	<0.001	1.51	–	–	–
2D upright vs 3D upright	1.62	0.118	0.32	–	–	–
2D inverted vs 3D inverted	1.65	0.113	0.33	–	–	–

Paired-sample *t*-tests: df = 24.
Bonferroni corrected *p*-values = 0.0125.

results showed that the mechanisms involved in holistic and featural processing likely underlie these differences. We found 3D upright faces were recognized with greater accuracy than 2D upright faces with no difference in reaction time. This is consistent with the hypothesis that enriching the contours and visual details via 3D increases the prominence of a face's first-order information, hence leading to enhanced accuracy during holistic processing. Simultaneously, it was demonstrated that regardless of whether the stimuli were presented in 2D or 3D, inverted faces were associated with slower reaction times and lower accuracy compared to upright faces. This provides evidence for the first time that face inversion effects that are commonly obtained for 2D faces (Itier & Taylor, 2002; Jacques et al., 2007; Rossion et al., 1999; Tanaka & Farah, 1993; Taubert et al., 2011) could also be observed for 3D faces. The unexpected finding was that 3D inverted faces did not yield any advantages in accuracy or reaction time over 2D inverted faces, and the size of the face inversion effect did not differ between 2D and 3D faces. Collectively, the results of the current study suggest that (i) the influence of 3D is more likely to manifest in conditions that require holistic processing where face stimuli retain intact first-order information, and (ii) less likely in conditions that require featural processing where first-order information is disrupted. The significance of the current findings will be discussed hereafter.

The key finding of this study was that 3D upright faces were recognized with greater accuracy than 2D upright faces, albeit with no difference in reaction time. At first glance it appears that the enriched depth information in 3D does not help to enhance processing speed; however, it is crucial to note that within this similar time frame, 3D faces provide a greater volume of visual information compared to those presented in 2D. This provides evidence that the additional visual information provided by 3D faces do not lead to an increase in information processing load and instead result in an improvement of accuracy. This is counter-intuitive to the idea of information computational speed, where larger amounts of information would lead to longer processing time (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; Fink & Neubauer, 2001). In sum, the additional information provided by 3D faces benefits accuracy at no additional cost of processing speed. Interestingly, this result indicates that processing speed or cognitive load do not necessarily correlate with the volume of information but are instead influenced by the *resolution* of information. The current result thus suggests that enriching the depth information between facial features and facial details via 3D enhances face recognition performance during holistic processing. This might be attributed to 3D details helping to make the stimuli appear more alike the real-life face perceptions that our visual systems are attuned to. In addition, this is in line with the notion that faces are distinguished from one another based on second-order information (Diamond & Carey, 1986), the variation in the spatial relationships between local features (e.g. distance between the eyes) among different faces. Accordingly, the more prominent second-order information provided by 3D upright faces make individual faces appear more distinctive from each other, leading to higher accuracy. Moreover, this is consistent with the hierarchical order of visual information processing, where the extraction of second-order information is not only facilitated by, but also necessitates a Gestalt representation where first-order information remains intact (Behrmann et al., 2014; Maurer et al., 2002).

Second, there was no reliable difference in either accuracy or reaction time between 2D inverted and 3D inverted faces. This suggests that inverted faces were processed similarly regardless of the image modality, and that enhancing visual information via 3D does not aid in featural processing in the current study. This is in line with the evidence that similar inversion effects occur in both 2D and 3D stimuli (i.e. 2D upright vs. 2D inverted, 3D upright vs. 3D inverted), indicating that the inversion effect did not change as a function of face modality. Although 3D stimuli are typically associated with enhanced visual information, the current results suggest that such information does not offer any behavioral advantage for inverted face recognition supported by featural processing (Maurer et al., 2002; Sadeh & Yovel, 2010; Sekuler et al., 2004; Tanaka & Farah, 1993). Indeed, although 3D enriches visual details of facial parts, it does not enhance the visual information that facilitates the construction of a cogent face representation (i.e. first-order information) during inversion. This study hence suggests that the benefits of 3D are abated when first-order information is disrupted and do not aid in featural processing, as reflected by 3D consistently failing to improve accuracy or reaction time when face stimuli were inverted.

Finally, while previous literature has established that inversion disrupts holistic processing by distorting first-order information, these studies have only been based on 2D face stimuli. The current study, however, not only replicated these findings but also demonstrated for the first time that there was a significant difference in accuracy and reaction time between 3D upright and 3D inverted face conditions. Our results thus reinforce the importance of first-order information in creating a meaningful Gestalt representation by which we are able to perceive an inherent identity that facilitates face recognition. Furthermore, these findings also suggest that during inversion, the incorporation of 3D only provides a "superficial" benefit of refining visual details and does not address the incoherence brought about by the disruption of first-order information. These results appear to be consistent with existing literature. That is, face recognition relies on global information – derived from the face's shape and physical structure – to a greater extent than local information of facial features (Caharel et al., 2009; Jiang, Blanz, & Rossion, 2011; Jiang et al., 2009), and such global diagnostic cues constitute second-order information crucial in distinguishing faces (Russell, Biederman, Nederhouser, & Sinha, 2007; Russell, Sinha, Biederman, & Nederhouser, 2006). However, the disruption of first-order information during inversion impairs the extraction of second-order information (Behrmann et al., 2014; Maurer et al., 2002). As such, inversion effects remain un-mitigated during 3D conditions as the enhancement of individual features fail to ameliorate the disruption of first-order information necessary for the initial construction of a holistic face representation and ensuing face recognition processes. Hence, even though the current study did not support the hypothesis that 3D would facilitate featural processing, the findings have furnished important preliminary results in this previously unexplored area of research. More specifically, the study has helped to further elucidate that first-order information serves as a "catalyst" for holistic processing, while providing evidence suggesting that the benefits of 3D during face processing also appear to be dependent on the validity of first-order information.

Nevertheless, a limitation of our study is that the stimuli used comprised only male faces, and it is unclear whether our results can be generalized to female face stimuli. This is particularly relevant given previous findings of gender perception to be largely supported by holistic processing (Yokoyama, Noguchi, Tachibana, Mukaida, & Kita, 2014; Zhao & Hayward, 2010), while studies have also consistently found that females demonstrate own-gender bias during face recognition (Herlitz & Lovén, 2013; Lewin & Herlitz, 2002; Palmer, Brewer, & Horry, 2013).

The current study has provided a good starting point for the understanding of face recognition processes in 3D, although the use of behavioral measures has limited the interpretation of the underlying mechanisms behind the observed effects. This is especially pertinent given that our study was indicative of enhanced processing speed for 3D upright faces but could not pinpoint the causality behind such a change. Future research could look into using neuroimaging techniques, such as electroencephology (EEG), to detect possible underlying differences in the processing of 2D and 3D faces that are not captured by the behavioral approach. Furthermore, a large number of ERP studies have previously demonstrated that face inversion elicits delayed latency and increased amplitude in N170. Future studies could thus explore how the amplitude and latency of N170 deviate when 3D is incorporated during holistic processing, and also examine whether 3D inverted faces elicit similar object-associated N170 responses as 2D inverted faces during featural processing. Other potential studies could also explore different methods of observing holistic and featural processing other than the face inversion effect. Indeed, a large number of studies (e.g. Aguirre, Singh, & D'Esposito, 1999; Farah et al., 1998; Tanaka & Farah, 1993; Taubert et al., 2011; Zhao & Hayward, 2010) have also employed different methods, such as scrambled faces, as a marker of disrupted holistic processing. While suggesting that the use of inverted faces might offer greater external validity than scrambled faces due to the former being more "natural", Tanaka and Farah (1993) also underlined that the distinction between holistic and featural processing occurs along a continuum rather than as a strict dichotomy. Thus, based on the evidence suggesting that scrambled faces elicit a greater degree of featural processing than inverted faces (Lê, Raufaste, & Démonet, 2003; Taubert et al., 2011), future studies could consider using scrambled faces to compare against upright and/or inverted faces, where differences in behavioral performance could indicate the *extent* to which the disruption of first-order information impacts 3D face processing. Findings from these studies would therefore establish the threshold at which 3D ameliorates or hinders featural processing. Lastly, another avenue of future research could involve exploring how the incorporation of 3D details might prove informative for existing models of face recognition (e.g. Bruce & Young, 1986), including the influence of 3D in the construction and subsequent identification of a face.

5. Conclusion

Our results provide a preliminary understanding of how 3D is (or is not) assimilated into our facial recognition processes. In particular, the current study found that 3D further increases accuracy during holistic processing while also providing evidence indicating that our facial recognition systems are capable of operating at greater efficiency – contrary to the view that a greater volume of information is associated with longer processing time. At the same time, the findings suggest that the advantages of 3D are not utilized in a manner that might ameliorate featural processing as we initially predicted, hence further emphasizing the importance of first-order information and adding to current literature on its role in facial recognition. The findings of this study have not only shed light on our current understanding of holistic face processing, but also provide practical implications. The enhanced accuracy found for 3D upright faces would presumably prove useful in telecommunication between parties, and also in forensic settings where success rates in the identification of criminal suspects may be improved using facial composites rendered in 3D. In all, there remains much room for further research into the effects of 3D on face recognition. The current study has initiated the bridging of this gap and serves as a platform for future research.

Acknowledgments

This research is supported by the BeingTogether Centre, a collaboration between Nanyang Technological University (NTU) Singapore and University of North Carolina (UNC) at Chapel Hill. The BeingTogether Centre is supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative.

References

Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *NeuroReport, 10*(1), 189–194. http://dx.doi.org/10.1097/00001756-199901180-00036.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*(3), 570–585. http://dx.doi.org/10.1037/0278-7393.33.3.570.

Behrmann, M., Richler, J., Avidan, G., & Kimchi, R. (2014). Holistic face perception. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization*. New York: Oxford University Press.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*(6), 551–565. http://dx.doi.org/10.1162/jocn.1996.8.6.551.

Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: Evidence from developmental prosopagnosia. *NeuroReport, 10*(4), 823–827.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*(Pt 3), 305–327. http://dx.doi.org/10.1111/j.2044-8295.1986.tb02199.x.

Caharel, S., Jiang, F., Blanz, V., & Rossion, B. (2009). Recognizing an individual face: 3D shape contributes earlier than 2D surface reflectance information. *NeuroImage, 47*(4), 1809–1818. http://dx.doi.org/10.1016/j.neuroimage.2009.05.065.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology. General, 115*(2), 107–117. http://dx.doi.org/10.1037/0096-3445.115.2.107.

Eimer, M., Kiss, M., & Nicholas, S. (2010). Response profile of the face-sensitive N170 component: A rapid adaptation study. *Cerebral Cortex, 20*(10), 2442–2452. http://dx.doi.org/10.1093/cercor/bhp312.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research, 35*(14), 2089–2093. http://dx.doi.org/10.1016/0042-6989(94)00273-O.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498. http://dx.doi.org/10.1037/0033-295X.105.3.482.

Fink, A., & Neubauer, A. C. (2001). Speed of information processing, psychometric intelligence: And time estimation as an index of cognitive load. *Personality and Individual Differences, 30*(6), 1009–1021. http://dx.doi.org/10.1016/S0191-8869(00)00089-1.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in encoding of configural information: Direct evidence. *Perception, 29*(2), 159–170. http://dx.doi.org/10.1068/p3012.

Häkkinen, J., Kawai, T., Takatalo, J., Leisti, T., Radun, J., Hirsaho, A., & Nyman, G. (2008). Measuring stereoscopic image quality experience with interpretation based quality methodology. *Proceedings of SPIE, 6808*. http://dx.doi.org/10.1117/12.760935. 68081B–68081B–12.

Heisz, J. J., Watter, S., & Shedden, J. M. (2006). Automatic face identity encoding at the N170. *Vision Research, 46*(28), 4604–4614. http://dx.doi.org/10.1016/j.visres.2006.09.026.

Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: A meta-analytic review. *Visual Cognition, 21*(9–10), 1306–1336. http://dx.doi.org/10.1080/13506285.2013.823140.

Hughes, J. F., Van Dam, A., McGuire, M., Sklar, D., Foley, J., Feiner, S., & Akeley, K. (2013). *Computer graphics: Principles and practice* (3rd ed.). New Jersey: Addison-Wesley Professional.

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage, 15*(2), 353–372. http://dx.doi.org/10.1006/nimg.2001.0982.

Itier, R. J., & Taylor, M. J. (2004). Source analysis of the N170 to faces and objects. *NeuroReport, 15*(8), 1261–1265. http://dx.doi.org/10.1097/01.wnr.0000127827.73576.d8.

Jacques, C., D'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision, 7*(8), 3. http://dx.doi.org/10.1167/7.8.3.

Jiang, F., Blanz, V., & Rossion, B. (2011). Holistic processing of shape cues in face identification: Evidence from face inversion, composite faces, and acquired prosopagnosia. *Visual Cognition, 19*(8), 1003–1034. http://dx.doi.org/10.1080/13506285.2011.604360.

Jiang, F., Dricot, L., Blanz, V., Goebel, R., & Rossion, B. (2009). Neural correlates of shape and surface reflectance information in individual faces. *Neuroscience, 163*(4), 1078–1091. http://dx.doi.org/10.1016/j.neuroscience.2009.07.062.

Lambooij, M., IJsselsteijn, W., Bouwhuis, D., & Heynderickx, I. (2011). Evaluation of stereoscopic images: Beyond 2D quality. *IEEE Transactions on Broadcasting, 57*(2), 432–444. http://dx.doi.org/10.1109/TBC.2011.2134590.

Lê, S., Raufaste, E., & Démonet, J. F. (2003). Processing of normal, inverted, and scrambled faces in a patient with prosopagnosia: Behavioural and eye tracking data. *Cognitive Brain Research, 17*(1), 26–35. http://dx.doi.org/10.1016/S0926-6410(03)00077-6.

Leder, H., & Carbon, C.-C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97*(Pt 1), 19–29. http://dx.doi.org/10.1348/000712605X54794.

Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition – Women's faces make the difference. *Brain and Cognition, 50*(1), 121–128.

Liu, C. H., Collin, C. a., & Chaudhuri, A. (2000). Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. *Perception, 29*(6), 729–743. http://dx.doi.org/10.1068/p3065.

Maurer, D., Grand, R. Le., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255–260. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12039607.

Palmer, M. A., Brewer, N., & Horry, R. (2013). Understanding gender bias in face recognition: Effects of divided attention at encoding. *Acta Psychologica, 142*(3), 362–369.

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research, 49*(23), 2856–2861. http://dx.doi.org/10.1016/j.visres.2009.08.025.

Rock, I. (1974). The perception of disoriented figures. *Scientific American, 230*(1), 78–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6221066.

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guérit, J. M. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology, 50*, 173–189. http://dx.doi.org/10.1016/S0301-0511(99)00013-7.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews, 1*(1), 63–75. http://dx.doi.org/10.1177/1534582302001001004.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport, 11*(1), 69–74. http://dx.doi.org/10.1097/00001756-200001170-00014.

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research, 47*(2), 157–165. http://dx.doi.org/10.1016/j.visres.2006.11.002.

Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (2006). Is pigmentation important for face recognition? Evidence from Contrast Negation. *Perception, 35*(6), 749–759. http://dx.doi.org/10.1068/p5490.

Sadeh, B., & Yovel, G. (2010). Why is the N170 enhanced for inverted faces? An ERP competition experiment. *NeuroImage, 53*(2), 782–789. http://dx.doi.org/10.1016/j.neuroimage.2010.06.029.

Schwaninger, A., Ryf, S., & Hofer, F. (2003). Configural information is processed differently in perception and recognition of faces. *Vision Research, 43*(14), 1501–1505. http://dx.doi.org/10.1016/S0042-6989(03)00171-8.

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*(5), 391–396. http://dx.doi.org/10.1016/j.cub.2004.02.028.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 46*(2), 225–245. http://dx.doi.org/10.1080/14640749308401045.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*(5), 583–592. http://dx.doi.org/10.3758/BF03211301.

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research, 51*(11), 1273–1278. http://dx.doi.org/10.1016/j.visres.2011.04.002.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Yokoyama, T., Noguchi, Y., Tachibana, R., Mukaida, S., & Kita, S. (2014). A critical role of holistic processing in face gender perception. *Frontiers in Human Neuroscience, 8*, 477. http://dx.doi.org/10.3389/fnhum.2014.00477.

Zhao, M., & Hayward, W. G. (2010). Holistic processing underlies gender judgments of faces. *Attention, Perception, & Psychophysics, 72*(3), 591–596. http://dx.doi.org/10.3758/APP.72.3.591.