

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

Eng Z. H. D.¹, Yick, Y. Y.¹, Guo, Y.², Xu, H.¹, Reiner, M.³, Cham, T. J.⁴, Chen S. H. A.^{1, 5, 6}

¹Division of Psychology, School of Humanities and Social Sciences, Nanyang Technological University, 14 Nanyang Drive, HSS-04-19, Singapore 637332, Singapore

²Institute for Media Innovation, Nanyang Technological University, 50 Nanyang Drive, Research Techno Plaza, XFrontiers Block, Level 03-01, Singapore 637553, Singapore

³Technion, Israel Institute of Technology, Technion International, Mauerberger Building, Technion City, Haifa 3200003, Israel

⁴Computer Engineering, School of Computer Engineering, Nanyang Technological University, Block N4 #02a-32, Nanyang Avenue, Singapore 639798, Singapore

⁵Centre for Research and Development in Learning (CRADLE), Nanyang Technological University, 62 Nanyang Drive, Block N1.2-B1-02a, Singapore 637459, Singapore

⁶Lee Kong Chian School of Medicine, Nanyang Technological University,
Experimental Medicine Building, 59 Nanyang Drive, Singapore 636921, Singapore

Corresponding authors:

SH Annabel Chen, PhD

Division of Psychology, School of Humanities and Social Sciences, Nanyang
Technological University, 14 Nanyang Drive, HSS-04-19, Singapore 637332,
Singapore

annabelchen@ntu.edu.sg; +65-6316-8836

and

Derric Eng ZH

Phone (+6597562530)

Email (engz0001@e.ntu.edu.sg; derric.engzh@gmail.com)

1 **Abstract**

2 Recognition of faces typically occurs via holistic processing where individual features
3 are combined to provide an overall facial representation. However, when faces are
4 inverted, there is greater reliance on featural processing where faces are recognized
5 based on their individual features. These findings are based on a substantial number
6 of studies using 2-dimensional (2D) faces and it is unknown whether these results can
7 be extended to 3-dimensional (3D) faces, which have more depth information that is
8 absent in the typical 2D stimuli used in face recognition literature. Using the face
9 inversion paradigm, the current study investigated how holistic and featural
10 processing are differentially influenced by 2D and 3D faces. Twenty-five participants
11 completed a delayed face-matching task consisting of upright and inverted faces that
12 were presented as both 2D and 3D stereoscopic images. Recognition accuracy was
13 significantly higher for 3D upright faces compared to 2D upright faces, providing
14 support that the enriched visual information in 3D stereoscopic images facilitates
15 holistic processing that is essential for the recognition of upright faces. Typical face
16 inversion effects were also obtained, regardless of whether the faces were presented in
17 2D or 3D. Moreover, recognition performances for 2D inverted and 3D inverted faces
18 did not differ. Taken together, these results demonstrated that 3D stereoscopic effects
19 influence face recognition during holistic processing but not during featural
20 processing. Our findings therefore provide a novel perspective that furthers our
21 understanding of face recognition mechanisms, shedding light on how the integration
22 of stereoscopic information in 3D faces influences face recognition processes.

23

24 Keywords: face recognition, holistic processing, featural processing, 2D, 3D, face
25 inversion effect, stereoscopic images

26

27

1. Introduction

28

Face recognition is an innate ability that is essential to our daily social

29

interactions. Most of us are able to recognize and distinguish faces instantly,

30

suggesting that faces are a special category of our visual expertise (Heisz, Watter, &

31

Shedden, 2006; Richler, Mack, Gauthier, & Palmeri, 2009; Taubert, Apthorp, Aagten-

32

Murphy & Alais, 2011; Maurer, Grand, & Mondloch, 2002). Research largely

33

supports that a face is processed holistically such that individual features are

34

integrated and represented as a whole (Behrmann, Richler, Avidan, & Kimchi, 2014;

35

Diamond & Carey, 1986; Maurer et al., 2002; Tanaka & Farah, 1993). Such holistic

36

processing has been shown to be more important for recognition of faces than for

37

other objects, as the resulting Gestalt representations help us make sense of the visual

38

information and perceive different identities (Behrmann et al., 2014). This is in

39

contrast to featural processing, where the visual stimulus is recognized based on its

40

individual components (e.g., eyes, mouth, nose, face contour, colour, brightness, etc.)

41

rather than as a whole (Tanaka & Farah, 1993; Diamond & Carey, 1986). Featural

42

processing occurs when faces are inverted and subsequently processed more similarly

43

to objects based on their individual features instead, hence leading to upside-down

44

faces not being recognized as Gestalt representations with inherent identities (Tanaka

45

& Farah, 1993; Tanaka & Sengco, 1997; Farah, Wilson, Drain & Tanaka, 1995;

46

Rossion & Gauthier, 2002). This is referred to as the “face inversion effect”, whereby

47

inversion deteriorates our face recognition ability drastically compared to the

48

recognition of non-face stimuli (Yin, 1969).

49

The existing knowledge about face processing, however, is based on studies that

50

examined 2-dimensional (2D) faces presented on computer screens and lacking the

51 visual depth information inherent in real life faces. In contrast, 3-dimensional (3D)
52 images provide greater depth and visual details (Häkkinen et al., 2008; Lambooj,
53 Ijsselsteijn, Bouwhuis, & Heynderickx, 2011), thereby leading to richer information
54 of both individual features as well as the spatial interrelationship between them
55 (configural information) (Schwaninger, Ryf, & Hofer, 2003), and thus a more
56 “comprehensive” Gestalt representation. These enhanced featural and spatial details
57 in 3D faces also provide additional visual information that could help to make the
58 stimuli more closely resemble the real-life perceptions that our visual systems are
59 attuned to. Based on these premises, it is expected that 3D details would provide an
60 advantage over 2D images during face recognition for both holistic and featural
61 processing.

62 To date, no published studies have examined the manner in which 3D stereoscopic
63 faces influence the mechanisms of holistic and featural processing in face recognition.
64 This study therefore aims to address this research gap by comparing the recognition of
65 3D stereoscopic and 2D faces in a classic face inversion paradigm (Yin, 1969; Tanaka
66 & Farah, 1993). We examined to what extent the findings of the face inversion effect
67 on 2D faces could be generalized to 3D faces, and aimed to understand processing
68 that is involved more heavily for 3D faces. It is hoped that the findings from this
69 study not only extend our current understanding of face recognition but also provide
70 novel perspectives for research ideas that are enabled by the advancement of 3D
71 technology.

72 It has yet to be established how the proposed advantages of enhanced visual
73 details in 3D influence the mechanisms underpinning the holistic and featural
74 processing involved in face recognition. Existing literature suggests that faces are
75 processed over three different stages (Maurer et al., 2002). During the first stage, first-

76 order processing occurs based on the general organization of the face's features (i.e.
77 two eyes, above a nose, above a mouth), for the initial face detection. Subsequently,
78 holistic processing occurs in the second stage where facial features are integrated to
79 form a Gestalt representation. At the third stage, second-order processing takes place
80 in which the variance between faces is analyzed, such as the distance between the
81 eyes, to form accurate and distinct face representations (Diamond & Carey, 1986;
82 Freire, Lee & Symons., 2000; Taubert et al., 2011). Holistic face recognition is
83 therefore based on how basic attributes are spatially arranged to form the prototypical
84 representation of a face (Diamond & Carey, 1986; Taubert et al., 2011). When faces
85 are presented upside-down, however, it disrupts the spatial relationship among the
86 facial features (the first-order information), slowing down face detection and
87 impairing holistic face processing consequently (Yin, 1969; Farah, Wilson, Drain, &
88 Tanaka, 1998; Farah et al., 1995; Maurer et al., 2002; Sekuler, Gaspar, Gold, &
89 Bennett, 2004; Tanaka & Farah, 1993). As a result, inverted faces are recognized as
90 an amalgamation of facial parts rather than as a congruent face.

91 A substantial number of studies have manipulated the upright or inverted
92 orientation of faces as a reliable method of eliciting holistic or featural processing
93 (Itier & Taylor, 2002; Leder & Carbon, 2006; Rossion et al., 1999; Sekuler et al.,
94 2004; Tanaka & Farah, 1993; Tanaka & Sengco, 1997; Taubert et al., 2011). It is
95 typically shown that holistic processing contributes to greater face recognition
96 accuracy and faster response time, as it facilitates the formation of a coherent
97 representation of a face (Itier & Taylor, 2002; Jacques, D'Arripe, & Rossion, 2007;
98 Rossion et al., 1999; Tanaka & Farah, 1993; Taubert et al., 2011). On the other hand,
99 inversion leads to an increase in cognitive demand due to the disruption of first-order
100 information (Behrmann et al., 2014; Maurer et al., 2002; Rock, 1974), resulting in

101 lower accuracy and slower reaction times during featural processing (Itier & Taylor,
102 2002; Jacques et al., 2007; Jiang, Dricot, Blanz, Goebel, & Rossion, 2009; Maurer et
103 al., 2002; Rossion et al., 1999; Tanaka & Farah, 1993; Tanaka & Sengco, 1997;
104 Taubert et al., 2011).

105 Results from event-related potential (ERP) studies focusing on the face-sensitive
106 N170 modulation have corroborated the behavioural findings. In particular, the
107 amplitude and latency of N170 are thought to index the degree and onset of early
108 structural encoding of faces (Bentin, Deouell, & Soroker, 1999; Bentin, Allison, Puce,
109 Perez, & McCarthy, 1996; Eimer, Kiss, & Nicholas, 2010; Heisz et al., 2006; Itier &
110 Taylor, 2002; Jacques et al., 2007; Maurer et al., 2002). Studies have consistently
111 shown that N170 has a later onset for inverted faces compared to upright faces,
112 supporting the notion of delayed processing speed for inverted faces (Heisz et al.,
113 2006; Itier & Taylor, 2002, 2004; Jacques et al., 2007; Rossion et al., 1999; Rossion
114 et al., 2000; Rossion & Gauthier, 2002; Sadeh & Yovel, 2010). Moreover, these
115 studies have also shown a larger N170 amplitude (more negative) for inverted faces
116 compared to upright ones, suggesting more complex structural encoding for inverted
117 relative to upright faces (Bentin et al., 1996; Eimer et al., 2010; Heisz et al., 2006;
118 Itier & Taylor, 2002; Maurer et al., 2002; Rossion et al., 1999; Rossion et al., 2000;
119 Sekuler et al., 2004). Taken together, evidence from both behavioral and ERP studies
120 provide convincing evidence that upright faces are associated with higher accuracy
121 and shorter processing time due to less complex structural encoding compared to
122 inverted faces (Eimer et al., 2010; Itier & Taylor, 2002; Jacques et al., 2007; Rossion
123 et al., 1999; Rossion et al., 2000). Here, we employed the face inversion paradigm to
124 investigate whether such holistic and featural processes in 2D face recognition are
125 similarly engaged during processing of 3D faces.

126 As aforementioned, 3D stereoscopic images provide greater depth and visual
127 details compared to their 2D counterparts (Häkkinen et al., 2008; Lambooi, et al.,
128 2011). Therefore, it is expected that 3D faces would provide i) enriched configural
129 information between facial features which could be beneficial to holistic processing,
130 and ii) richer visual details of the individual facial parts which could be beneficial to
131 featural processing (Liu, Collin & Chaudhuri, 2000). In order to test for these
132 differences and their effects on holistic and featural processing during face
133 recognition, participants completed a delayed face-matching task for 2D and 3D faces
134 that were presented upright or inverted.

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

146

147 **2. Methodology**

148 2.1 Participants

149 Twenty-four undergraduates and two recent graduates, with normal or corrected-
150 to-normal vision, were recruited from four local universities in Singapore – Nanyang

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

159

160 2.2 Stimuli

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

171 The 3D faces were created following the setup illustrated in Figure 1. A face
172 model was first translated close to the “Screen Z plane” during rendering. Following
173 which, the face was rendered between two different planes - the “Near Z plane” and
174 “Far Z plane” - to the rendering plane. In the 2D mode, only one image is projected;
175 whereas in the 3D model, the monitor displays images from the left and right

176 alternately. Due to the high refresh rate (200Hz) of the monitor, viewers perceived the
177 left and right images appearing simultaneously as one stimulus. More importantly,
178 although the perception of faces in 2D is based on a symmetrical viewpoint, the
179 perception of 3D faces is facilitated by left and right eye views being non-parallel (see
180 Figure 1b). In the 3D display mode, the left eye looks towards the “centre-right”
181 direction while the right eye looks towards the “centre-left” direction instead. All
182 stimuli were presented on a black background displayed using an Alienware laptop
183 and viewed with a pair of 3D Vision wireless active shutter glasses (see Figure 2).

184

185 <Insert Figure 1 here>

186

187 Stimuli were set according to the following parameters, where image height and
188 width measured 10cm by 10cm. The “Near Z”, “Far Z” and “Screen Z” planes were
189 rendered within the Normalized Devices Coordinate (NDC) space, which is a
190 reference frame used within the application programme interface, OpenGL. The NDC
191 space refers to a coordinate system that depicts the positions of plotted virtual
192 points, wherein the visual coordinates of a stimulus seen from a real-world eye view
193 may be mapped into this space to render 2D and 3D vector graphics (Hughes et al.,
194 2013). Within the cube-shaped NDC space, the “Near Z-plane” was located away
195 from the participant at a distance of 30cm, while the “Far Z-plane” was located
196 further away at 300cm. The “Screen Z-plane” was placed at a distance of 100cm, and
197 the interpupillary distance was maintained at 5cm. As a result, participants were able
198 to experience the stereo effect that led to the perception of faces in 3D faces appearing
199 stereoscopically at a distance of approximately 2cm in front of the monitor screen. 2D

200 faces were also viewed through the same set up, albeit appearing flat on the monitor
201 screen (see Figure 1).

202

203 2.3 Procedure

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

217

218 <Insert Figure 2 here>

219

220 Each trial commenced with an audio signal “START” along with the presentation
221 of a fixation cross in the center of the screen for 1000ms. This was followed by a
222 1000ms blank screen prior to the onset of the study face for 1000ms. A blank screen
223 was then presented for 1000ms, followed by an audio signal “GO” along with a
224 1000ms fixation cross. The fixation cross was then offset for 1000ms (i.e. blank

225 screen) before the onset of the test face for 1000ms during which participants were
226 required to make the binary “match” or “non-match” responses by pressing the left
227 and right key respectively. A 1000ms inter-stimulus blank screen is then presented
228 before the onset of the next trial. Illustrations of the trial sequences are presented in
229 Figure 3.

230

231

<Insert Figure 3 here>

232

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

239

240 2.4 Data Analysis

241 Incorrect trials and correct trials that exceeded the 1000ms response window were
242 coded as outliers, as participants were able to execute delayed face recognition tasks
243 successfully within similar time frames in previous studies (Caharel, Jiang, Blanz &
244 Rossion, 2009; Rossion et al., 1999; Taubert et al., 2011). In the current study, the
245 highest mean RT for the current study was 705ms (SE = 12) obtained from the 2D
246 inverted face condition. Analyses of accuracy rates and RT were restricted to trials in
247 which participants provided correct responses. A three-step analysis was conducted:
248 1) A 2-way repeated measure ANOVA of modality (2D/3D) by orientation
249 (upright/inverted) was conducted for accuracy and reaction time separately. In the

250 analysis of accuracy outcomes, a Friedman two-way analysis of variance of ranks was
251 also conducted due to the non-normal distribution in at least one condition (see
252 below); 2) Follow-up tests were conducted through planned contrasts with Bonferroni
253 corrections to compare the effects of interest. In this analysis, paired-comparison *t*-
254 tests were conducted to examine i) the performance differences between 2D and 3D
255 stimuli of the same orientation (2D upright vs 3D upright, 2D inverted vs 3D
256 inverted) and, ii) inversion effects (2D upright vs 2D inverted, 3D upright vs 3D
257 inverted). Wilcoxon signed-rank tests were also run for the above comparisons
258 pertaining to accuracy results in view of their non-normal distribution; and 3) to
259 compare the extent of inversion effects between 2D and 3D conditions, subtraction
260 scores were obtained by measuring differences in accuracy and reaction time
261 separately (e.g. 2D upright accuracy score minus 2D inverted accuracy score). Paired-
262 comparison *t*-tests were then conducted between the subtraction scores of 2D and 3D
263 conditions for both accuracy and reaction time.

264

265 **3. Results**

266 The behavioral performances are displayed in Figure 4. Paired-samples *t*-tests
267 showed that participants performed significantly above chance level for all 4
268 conditions [$ts > 52.52$, $ps < .001$].

269

270 <Insert Figure 4 here>

271

272 Accuracy results were non-normally distributed in two of the four conditions (2D
273 upright and 3D upright conditions), therefore we conducted a non-parametric
274 Friedman test in addition to the initial planned ANOVA. The pattern of results for

275 both tests were similar - the Friedman test indicated a significant difference among
276 the accuracy scores across the four conditions [$\chi^2(3) = 50.72, p < .001$, Kendall's $W =$
277 $.68$]. This result was confirmed by ANOVA which revealed significant main effects
278 of orientation [$F(1, 24) = 58.56, p < .001, \eta_p^2 = .71$] and modality [$F(1, 24) = 4.71, p =$
279 $.040, \eta_p^2 = .16$] on accuracy. There was no significant interaction between the effects
280 of modality and orientation on accuracy [$F(1, 24) = 1.28, p = .269, \eta_p^2 = .05$].

281 Post hoc multiple comparisons were conducted following ANOVA. To
282 account for the non-normal distributions of accuracy data, Wilcoxon signed-rank tests
283 with multiple comparisons were conducted (Bonferroni corrected p -values = $.0125$).
284 These results were consistent with the results from Bonferroni-corrected paired-
285 samples t -tests. Both parametric and non-parametric results are reported in Table 1.
286 As predicted, 3D upright faces were recognized with significantly greater accuracy
287 than 3D inverted faces. Consistent with previous literature, 2D upright faces elicited
288 significantly greater accuracy than 2D inverted faces. Interestingly, it was found that
289 3D upright faces were recognized with significantly greater accuracy than 2D upright
290 faces. However, no differences were found between 3D inverted and 2D inverted face
291 conditions.

292 Reaction time data for all conditions was found to be normally distributed.
293 The ANOVA revealed significant main effects of orientation [$F(1, 24) = 79.34, p <$
294 $.001, \eta_p^2 = .77$], and modality [$F(1, 24) = 4.40, p = .047, \eta_p^2 = .16$] on reaction time
295 performance. There was no significant interaction between the effects of modality and
296 orientation on reaction time [$F(1, 24) = .01, p = .933, \eta_p^2 = .00$].

297 Paired-samples t -tests (Bonferroni corrected p -values = $.0125$) were conducted
298 to compare reaction time performances between the different conditions (see Table 1).

299 It was found that the reaction time performance for 3D upright faces was significantly
300 faster than 3D inverted faces. The same pattern was observed for the comparison
301 between 2D upright and 2D inverted faces. It was also found that there were no
302 significant differences between the 3D upright and 2D upright face conditions, as well
303 as between the 3D inverted and 2D inverted face conditions.

304

305 <Insert Table 1 here>

306

307 Finally, to determine the extent of inversion effects in 2D and 3D stimuli, we
308 compared the differences in performances between upright and inverted conditions of
309 2D and 3D stimuli [Accuracy: (3D upright - 3D inverted) vs. (2D upright - 2D
310 inverted); RT: (3D upright - 3D inverted) vs. (2D upright - 2D inverted)]. No
311 significant differences were found, suggesting that the inversion effects did not differ
312 reliably between 2D and 3D stimuli for both accuracy [$t(24) = 1.13, p = .269, d = .23$]
313 and reaction time [$t(24) = 0.08, p = .933, d = .02$].

314

315 **4. Discussion**

316 The current study examined the differential processes between 2D and 3D face
317 recognition using the face inversion paradigm. Our results showed that the
318 mechanisms involved in holistic and featural processing likely underlie these
319 differences. We found 3D upright faces were recognized with greater accuracy than
320 2D upright faces with no difference in reaction time. This is consistent with the
321 hypothesis that enriching the contours and visual details via 3D increases the
322 prominence of a face's first-order information, hence leading to enhanced accuracy
323 during holistic processing. Simultaneously, it was demonstrated that regardless of

324 whether the stimuli were presented in 2D or 3D, inverted faces were associated with
325 slower reaction times and lower accuracy compared to upright faces. This provides
326 evidence for the first time that face inversion effects that are commonly obtained for
327 2D faces (Itier and Taylor, 2002; Jacques et al., 2007; Rossion et al., 1999; Tanaka &
328 Farah, 1993; Taubert et al., 2011) could also be observed for 3D faces. The
329 unexpected finding was that 3D inverted faces did not yield any advantages in
330 accuracy or reaction time over 2D inverted faces, and the size of the face inversion
331 effect did not differ between 2D and 3D faces. Collectively, the results of the current
332 study suggest that i) the influence of 3D is more likely to manifest in conditions that
333 require holistic processing where face stimuli retain intact first-order information, and
334 ii) less likely in conditions that require featural processing where first-order
335 information is disrupted. The significance of the current findings will be discussed
336 hereafter.

337 The key finding of this study was that 3D upright faces were recognized with
338 greater accuracy than 2D upright faces, albeit with no difference in reaction time. At
339 first glance it appears that the enriched depth information in 3D does not help to
340 enhance processing speed; however, it is crucial to note that within this similar time
341 frame, 3D faces provide a greater volume of visual information compared to those
342 presented in 2D. This provides evidence that the additional visual information
343 provided by 3D faces do not lead to an increase in information processing load and
344 instead result in an improvement of accuracy. This is counter-intuitive to the idea of
345 information computational speed, where larger amounts of information would lead to
346 longer processing time (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007;
347 Fink & Neubauer, 2001). In sum, the additional information provided by 3D faces
348 benefits accuracy at no cost of processing speed. Interestingly, this result indicates

349 that processing speed or cognitive load do not necessarily correlate with the volume
350 of information but are instead influenced by the *resolution* of information. The current
351 result thus suggests that enriching the depth information between facial features and
352 facial details via 3D enhances face recognition performance during holistic
353 processing. This might be attributed to 3D details helping to make the stimuli appear
354 more alike the real-life face perceptions that our visual systems are attuned to. In
355 addition, this is in line with the notion that faces are distinguished from one another
356 based on second-order information (Diamond and Carey, 1986), the variation in the
357 spatial relationships between local features (e.g. distance between the eyes) among
358 different faces. Accordingly, the more prominent second-order information provided
359 by 3D upright faces make individual faces appear more distinctive from each other,
360 leading to higher accuracy. Moreover, this is consistent with the hierarchical order of
361 visual information processing, where the extraction of second-order information is not
362 only facilitated by, but also necessitates a Gestalt representation where first-order
363 information remains intact (Behrmann et al., 2014; Maurer et al., 2002).

364 Second, there was no reliable difference in either accuracy or reaction time
365 between 2D inverted and 3D inverted faces. This suggests that inverted faces were
366 processed similarly regardless of the image modality, and that enhancing visual
367 information via 3D does not aid in featural processing in the current study. This is in
368 line with the evidence that similar inversion effects occur in both 2D and 3D stimuli
369 (i.e. 2D upright vs. 2D inverted, 3D upright vs. 3D inverted), indicating that the
370 inversion effect did not change as a function of face modality. Although 3D stimuli
371 are typically associated with enhanced visual information, the current results suggest
372 that such information does not offer any behavioral advantage for inverted face
373 recognition supported by featural processing (Maurer et al., 2002; Sadeh & Yovel,

374 2010; Sekuler et al., 2004; Tanaka & Farah, 1993). Indeed, although 3D enriches
375 visual details of facial parts, it does not enhance the visual information that facilitates
376 the construction of a cogent face representation (i.e. first-order information) during
377 inversion. This study hence suggests that the benefits of 3D are abated when first-
378 order information is disrupted and do not aid in featural processing, as reflected by 3D
379 consistently failing to improve accuracy or reaction time when face stimuli were
380 inverted.

381 Finally, while previous literature has established that inversion disrupts holistic
382 processing by distorting first-order information, these studies have only been based on
383 2D face stimuli. The current study, however, not only replicated these findings but
384 also demonstrated for the first time that there was a significant difference in accuracy
385 and reaction time between 3D upright and 3D inverted face conditions. Our results
386 thus reinforce the importance of first-order information in creating a meaningful
387 Gestalt representation by which we are able to perceive an inherent identity that
388 facilitates face recognition. Furthermore, these findings also suggest that during
389 inversion, the incorporation of 3D only provides a “superficial” benefit of refining
390 visual details and does not address the incoherence brought about by the disruption of
391 first-order information. These results appear to be consistent with existing literature.
392 That is, face recognition relies on global information - derived from the face’s shape
393 and physical structure - to a greater extent than local information of facial features
394 (Caharel et al., 2009; Jiang et al., 2009; Jiang et al., 2011), and such global diagnostic
395 cues constitute second-order information crucial in distinguishing faces (Russell,
396 Sinha, Biederman & Nederhouser, 2006; Russell, Biederman, Nederhouser & Sinha,
397 2007). However, the disruption of first-order information during inversion impairs the
398 extraction of second-order information (Behrmann et al., 2014; Maurer et al., 2002).

399 As such, inversion effects remain un-mitigated during 3D conditions as the
400 enhancement of individual features fail to ameliorate the disruption of first-order
401 information necessary for the initial construction of a holistic face representation and
402 ensuing face recognition processes. Hence, even though the current study did not
403 support the hypothesis that 3D would facilitate featural processing, the findings have
404 furnished important preliminary results in this previously unexplored area of research.
405 More specifically, the study has helped to further elucidate that first-order information
406 serves as a “catalyst” for holistic processing, while providing evidence suggesting that
407 the benefits of 3D during face processing also appear to be dependent on the validity
408 of first-order information.

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

416 The current study has provided a good starting point for the understanding of face
417 recognition processes in 3D, although the use of behavioural measures has limited the
418 interpretation of the underlying mechanisms behind the observed effects. This is
419 especially pertinent given that our study was indicative of enhanced processing speed
420 for 3D upright faces but could not pinpoint the causality behind such a change. Future
421 research could look into using neuroimaging techniques, such as
422 electroencephalography (EEG), to detect possible underlying differences in the
423 processing of 2D and 3D faces that are not captured by the behavioral approach.

424 Furthermore, a large number of ERP studies have previously demonstrated that face
425 inversion elicits delayed latency and increased amplitude in N170. Future studies
426 could thus explore how the amplitude and latency of N170 deviate when 3D is
427 incorporated during holistic processing, and also examine whether 3D inverted faces
428 elicit similar object-associated N170 responses as 2D inverted faces during featural
429 processing. Other potential studies could also explore different methods of observing
430 holistic and featural processing other than the face inversion effect. Indeed, a large
431 number of studies (e.g. Agguire, Singh & D' Esposito, 1999; Farah et al., 1998;
432 Tanaka and Farah, 1993; Taubert et al., 2011; Zhao & Hayward, 2010) have also
433 employed different methods, such as scrambled faces, as a marker of disrupted
434 holistic processing. While suggesting that the use of inverted faces might offer greater
435 external validity than scrambled faces due to the former being more “natural”, Tanaka
436 and Farah (1993) also underlined that the distinction between holistic and featural
437 processing occurs along a continuum rather than as a strict dichotomy. Thus, based on
438 the evidence suggesting that scrambled faces elicit a greater degree of featural
439 processing than inverted faces (Lê, Raufaste, & Démonet, 2003; Taubert et al., 2011),
440 future studies could consider using scrambled faces to compare against upright and/or
441 inverted faces, where differences in behavioral performance could indicate the extent
442 to which the disruption of first-order information impacts 3D face processing.
443 Findings from these studies would therefore establish the threshold at which 3D
444 ameliorates or hinders featural processing. Lastly, another avenue of future research
445 could involve exploring how the incorporation of 3D details might prove informative
446 for existing models of face recognition (e.g. Bruce and Young, 1986), including the
447 influence of 3D in the construction and subsequent identification of a face.
448

449

5. Conclusion

450 Our results provide a preliminary understanding of how 3D is (or is not)
451 assimilated into our facial recognition processes. In particular, the current study found
452 that 3D further increases accuracy during holistic processing while also providing
453 evidence indicating that our facial recognition systems are capable of operating at
454 greater efficiency - contrary to the view that a greater volume of information is
455 associated with longer processing time. At the same time, the findings suggest that the
456 advantages of 3D are not utilized in a manner that might ameliorate featural
457 processing as we initially predicted, hence further emphasizing the importance of
458 first-order information and adding to current literature on its role in facial recognition.
459 The findings of this study have not only shed light on our current understanding of
460 holistic face processing, but also provide practical implications. The enhanced
461 accuracy found for 3D upright faces would presumably prove useful in
462 telecommunication between parties, and also in forensic settings where success rates
463 in the identification of criminal suspects may be improved using facial composites
464 rendered in 3D. In all, there remains much room for further research into the effects of
465 3D on face recognition. The current study has initiated the bridging of this gap and
466 serves as a platform for future research.

467

468

6. Acknowledgements

469 This research is supported by BeingThere Centre (NTU), funded by the Singapore
470 National Research Foundation under its International Research Centre @ Singapore
471 Funding Initiative and administered by the IDM Programme Office.

472 **References**

473

474 Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the
475 responses of face and object-sensitive cortical areas. *Neuroreport*, *10*(1), 189–94.
476 <http://doi.org/10.1097/00001756-199901180-00036>

477

478 Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time
479 and cognitive load in working memory. *Journal of Experimental Psychology.*
480 *Learning, Memory, and Cognition*, *33*(3), 570–585. [http://doi.org/10.1037/0278-](http://doi.org/10.1037/0278-7393.33.3.570)
481 [7393.33.3.570](http://doi.org/10.1037/0278-7393.33.3.570)

482

483 Baudouin, J. Y., Gilibert, D., Sansone, S., & Tiberghien, G. (2000). When the smile is
484 a cue to familiarity. *Memory (Hove, England)*, *8*(5), 285–292.
485 <http://doi.org/10.1080/09658210050117717>

486

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

490

491 Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996).
492 Electrophysiological Studies of Face Perception in Humans. *Journal of*
493 *Cognitive Neuroscience*, *8*(6), 551–565. <http://doi.org/10.1162/jocn.1996.8.6.551>

494

495 Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face
496 recognition: evidence from developmental prosopagnosia. *Neuroreport*, *10*(4),
497 823–827.

498

499 Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of*
500 *Psychology*, *77*(Pt 3), 305–327. [http://doi.org/10.1111/j.2044-](http://doi.org/10.1111/j.2044-8295.1986.tb02199.x)
501 [8295.1986.tb02199.x](http://doi.org/10.1111/j.2044-8295.1986.tb02199.x)

502

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

507

508 Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of
509 expertise. *Journal of Experimental Psychology. General*, *115*(2), 107–17.
510 <http://doi.org/10.1037/0096-3445.115.2.107>

511

512 Eimer, M., Kiss, M., & Nicholas, S. (2010). Response profile of the face-sensitive
513 N170 component: A rapid adaptation study. *Cerebral Cortex*, *20*(10), 2442–
514 2452. <http://doi.org/10.1093/cercor/bhp312>

515

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

520

521 Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special”

522 about face perception? *Psychological Review*, *105*(3), 482–498.

523 <http://doi.org/10.1037/0033-295X.105.3.482>

524

525 Fink, A., & Neubauer, A. C. (2001). Speed of information processing, psychometric

526 intelligence: And time estimation as an index of cognitive load. *Personality and*

527 *Individual Differences*, *30*(6), 1009–1021. <http://doi.org/10.1016/S0191->

528 [8869\(00\)00089-1](http://doi.org/10.1016/S0191-8869(00)00089-1)

529

530 Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in

531 encoding of configural information: Direct evidence. *Perception*, *29*(2), 159 –

532 170. doi:10.1068/p3012

533

534 Häkkinen, J., Kawai, T., Takatalo, J., Leisti, T., Radun, J., Hirsaho, A., & Nyman, G.

535 (2008). Measuring stereoscopic image quality experience with interpretation

536 based quality methodology. *Proceedings of SPIE*, *6808*, 68081B–68081B–12.

537 <http://doi.org/10.1117/12.760935>

538

539 Heisz, J. J., Watter, S., & Shedden, J. M. (2006). Automatic face identity encoding at

540 the N170. *Vision Research*, *46*(28), 4604–4614.

541 <http://doi.org/10.1016/j.visres.2006.09.026>

542

543 Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face
544 recognition: a meta-analytic review. *Visual Cognition*, 21(9 – 10), 1306 – 1336.
545 <http://dx.doi.org/10.1080/13506285.2013.823140>.
546

547 Hughes, J. F., Van Dam, A., McGuire, M., Sklar, D., Foley, J., Feiner, S., & Akeley,
548 K. (2013). *Computer Graphics: Principles and Practice (3rd ed.)*. New Jersey:
549 Addison-Wesley Professional.
550

551 Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both
552 encoding and recognition processes of unfamiliar Faces: a repetition study using
553 ERPs. *NeuroImage*, 15(2), 353–372. <http://doi.org/10.1006/nimg.2001.0982>
554

555 Itier, R. J., & Taylor, M. J. (2004). Source analysis of the N170 to faces and objects.
556 *Neuroreport*, 15(8), 1261–1265.
557 <http://doi.org/10.1097/01.wnr.0000127827.73576.d8>
558

559 Jacques, C., D'Arripe, O., & Rossion, B. (2007). The time course of the inversion
560 effect during individual face discrimination. *Journal of Vision*, 7(8), 3.
561 <http://doi.org/10.1167/7.8.3>
562

563 Jiang, F., Dricot, L., Blanz, V., Goebel, R., & Rossion, B. (2009). Neural correlates of
564 shape and surface reflectance information in individual faces. *Neuroscience*,
565 163(4), 1078–1091. <http://doi.org/10.1016/j.neuroscience.2009.07.062>
566

- 567 Jiang, F., Blanz, V., & Rossion, B. (2011). Holistic processing of shape cues in face
568 identification: Evidence from face inversion, composite faces, and acquired
569 prosopagnosia. *Visual Cognition*, 19(8), 1003–1034.
570 <http://doi.org/10.1080/13506285.2011.604360>
571
572
- 573 Lambooj, M., IJsselsteijn, W., Bouwhuis, D., & Heynderickx, I. (2011). Evaluation
574 of stereoscopic images: Beyond 2D quality. *IEEE Transactions on Broadcasting*,
575 57(2), 432–444. <http://doi.org/10.1109/TBC.2011.2134590>
576
- 577 Lê, S., Raufaste, E., & Démonet, J. F. (2003). Processing of normal, inverted, and
578 scrambled faces in a patient with prosopagnosia: Behavioural and eye tracking
579 data. *Cognitive Brain Research*, 17(1), 26–35. [http://doi.org/10.1016/S0926-
580 6410\(03\)00077-6](http://doi.org/10.1016/S0926-6410(03)00077-6)
581
- 582 Leder, H., & Carbon, C.-C. (2006). Face-specific configural processing of relational
583 information. *British Journal of Psychology*, 97(Pt 1), 19–29.
584 <http://doi.org/10.1348/000712605X54794>
585
- 586 Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition – Women’s faces
587 make the difference. *Brain and Cognition*, 50(1), 121 – 128.
588
- 589 Liu, C. H., Collin, C. a, & Chaudhuri, A. (2000). Does face recognition rely on
590 encoding of 3-D surface? Examining the role of shape-from-shading and shape-
591 from-stereo. *Perception*, 29(6), 729–743. <http://doi.org/10.1068/p3065>

592

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

596

597 Palmer, M. A., Brewer, N., & Horry, R. (2013). Understanding gender bias in face
598 recognition: Effects of divided attention at encoding. *Acta psychologica*, 142(3),
599 362-369.

600

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

604

605 Rock, I. (1974). The perception of disoriented figures. *Scientific American*, 230(1),
606 78–85. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6221066>

607

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

612

613 Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., &
614 Crommelinck, M. (2000). The N170 occipito-temporal component is delayed
615 and enhanced to inverted faces but not to inverted objects: an

616 electrophysiological account of face-specific processes in the human brain.
617 *Neuroreport*, 11(1), 69–74. <http://doi.org/10.1097/00001756-200001170-00014>
618
619 Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted
620 faces? *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 63–75.
621 <http://doi.org/10.1177/1534582302001001004>
622
623
624 Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (2006). Is pigmentation
625 important for face recognition? Evidence from Contrast Negation. *Perception*, 35
626 (6), 749 – 759. doi: 10.1068/p5490
627
628 Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2006). The utility of surface
629 reflectance for the recognition of upright and inverted faces. *Vision Research*, 47
630 (2), 157 – 165. <http://dx.doi.org/10.1016/j.visres.2006.11.002>
631
632 Sadeh, B., & Yovel, G. (2010). Why is the N170 enhanced for inverted faces? An
633 ERP competition experiment. *NeuroImage*, 53(2), 782–789.
634 <http://doi.org/10.1016/j.neuroimage.2010.06.029>
635
636 Schwaninger, A., Ryf, S., & Hofer, F. (2003). Configural information is processed
637 differently in perception and recognition of faces. *Vision Research*, 43 (14), 1501
638 – 1505. [http://dx.doi.org/10.1016/S0042-6989\(03\)00171-8](http://dx.doi.org/10.1016/S0042-6989(03)00171-8)
639

640 Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to
641 quantitative, not qualitative, changes in face processing. *Current Biology*, *14*(5),
642 391–396. <http://doi.org/10.1016/j.cub.2004.02.028>
643
644 Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The*
645 *Quarterly Journal of Experimental Psychology. A, Human Experimental*
646 *Psychology*, *46*(2), 225–245. <http://doi.org/10.1080/14640749308401045>
647
648 Tanaka, J. W., & Sengco, J. a. (1997). Features and their configuration in face
649 recognition. *Memory & Cognition*, *25*(5), 583–592.
650 <http://doi.org/10.3758/BF03211301>
651
652 Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic
653 processing in face perception: Evidence from the face inversion effect. *Vision*
654 *Research*, *51*(11), 1273–1278. <http://doi.org/10.1016/j.visres.2011.04.002>
655
656 Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental*
657 *Psychology*, *81*(1), 141 – 145.
658
659 Yokoyama, T., Noguchi, Y., Tachibana, R., Mukaida, S., & Kita, S. (2014). A critical
660 role of holistic processing in face gender perception. *Frontiers in Human*
661 *Neuroscience*, *8*, 477. <http://doi.org/10.3389/fnhum.2014.00477>
662

663 Zhao, M., & Hayward, W. G. (2010). Holistic processing underlies gender judgments
664 of faces. *Attention, Perception, & Psychophysics*, 72(3), 591–596.

665 <http://doi.org/10.3758/APP.72.3.591>

666

667 **Figure Captions**

668

669 Figure 1. Illustrations of projection of visual stimuli. (A) Image of set-up depicting
670 coordinates of the Far Z, Screen Z and Near Z planes in the projection of the 3D face
671 stimuli. (B) Top view of projection of visual stimuli presented in 2D and 3D. The
672 black solid line denotes the rendering plane. Black dashed lines represent the
673 cyclopean view in 2D display mode, while red dashed lines and green dashed lines
674 represent the left and right eye views respectively in 3D display mode.

675

676 Figure 2. Image of experimental set-up depicting participant equipped with 3D Vision
677 wireless active shutter glasses.

678

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

686

687 Figure 4. Behavioral results. (A) Mean accuracy for correct responses (\pm standard
688 error) (B) Mean reaction time for correct responses (\pm standard error). $*p < .010$.

689

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

Table 1

Pairwise Comparisons	Paired-sample <i>t</i> -tests [#]			Wilcoxon signed-rank tests [#]		
	<i>t</i>	<i>p</i>	<i>d</i>	<i>Z</i>	<i>p</i>	<i>r</i>
Accuracy:						
3D upright vs 3D inverted	7.37	<.001	1.47	-4.21	<.001	-.60
2D upright vs 2D inverted	5.71	<.001	1.14	-3.89	<.001	-.55
2D upright vs 3D upright	-3.31	.003	0.66	-2.73	.006	-.39
2D inverted vs 3D inverted	-0.54	.597	0.11	-0.54	.592	-.08
Reaction Time:						
3D upright vs 3D inverted	-7.04	<.001	1.41	-	-	-
2D upright vs 2D inverted	-7.53	<.001	1.51	-	-	-
2D upright vs 3D upright	1.62	.118	0.32	-	-	-
2D inverted vs 3D inverted	1.65	.113	0.33	-	-	-

Paired-sample *t*-tests: *df* = 24

[#]Bonferroni corrected *p*-values = .0125

Figure 1
[Click here to download high resolution image](#)

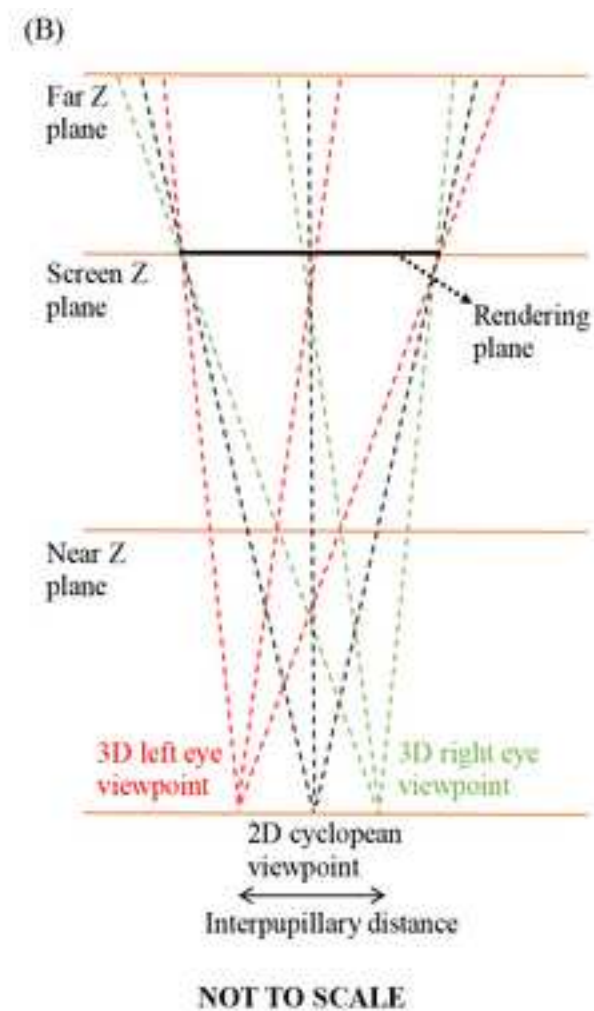
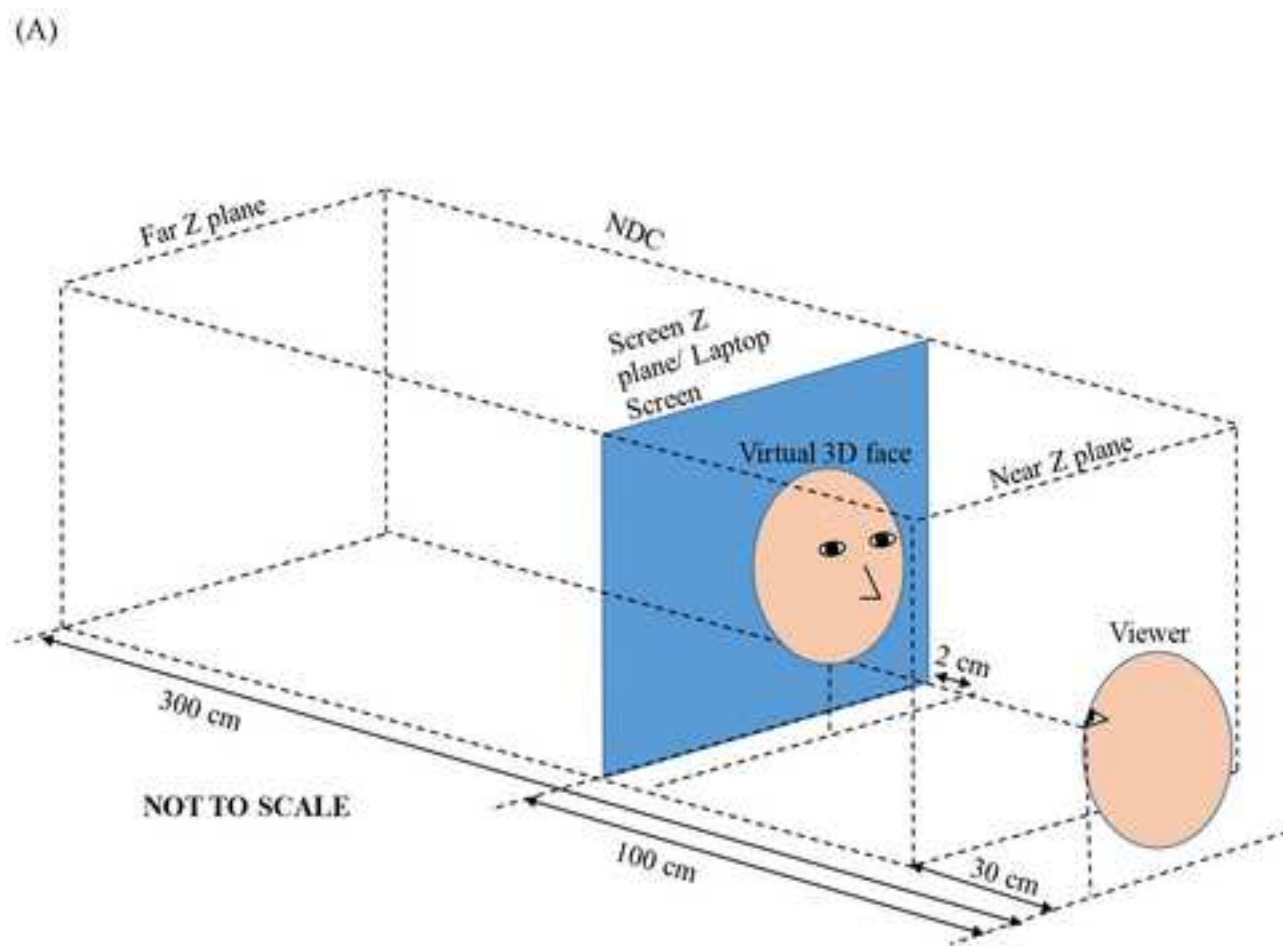
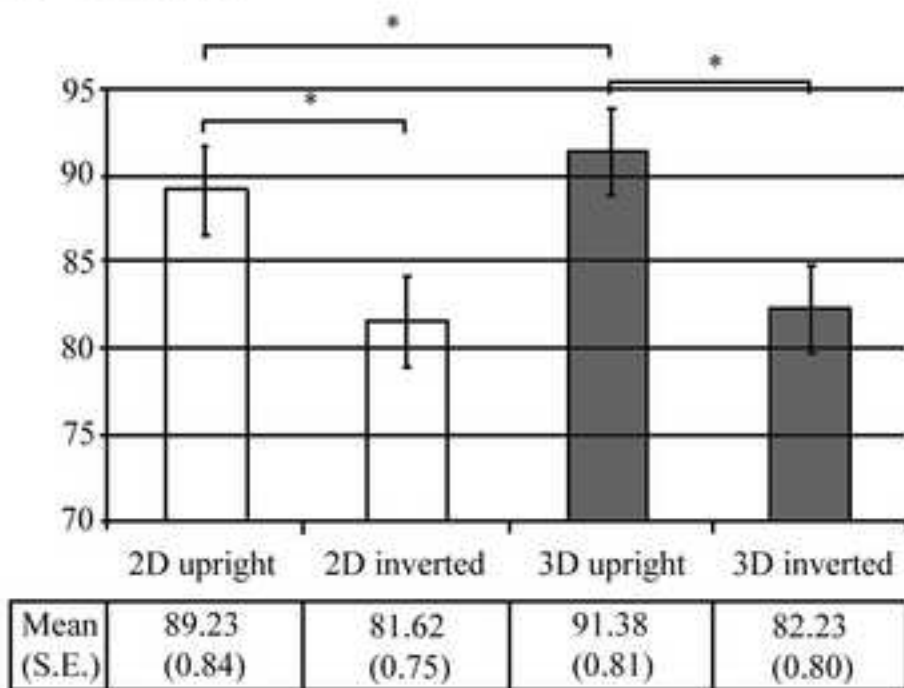


Figure 2
[Click here to download high resolution image](#)



Figure 4
[Click here to download high resolution image](#)

(A) Accuracy (%)



(B) Reaction time (ms)

