
Peer reviewed version

Link to published version (if available):
10.1068/p7940

Link to publication record in Explore Bristol Research

PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms
SHORT REPORT

The Headscarf Effect Re-Visited:
Further Evidence For A Culture-Based Internal Face Processing Advantage

Yin Wang¹, Justin Thomas², Sophia C. Weissgerber¹, Sahar Kazemini¹, Israr Ul-Haq¹, & Susanne Quadflieg¹,³

¹New York University Abu Dhabi, Abu Dhabi, UAE
²Zayed University Abu Dhabi, Abu Dhabi, UAE
³University of Bristol, Bristol, UK

Corresponding Author:
Susanne Quadflieg
School of Experimental Psychology, University of Bristol
12A Priory Road, BS8 1TU Bristol
United Kingdom
email: s.quadflieg@bristol.ac.uk
telephone: +44 (0) 117 928 8568

Declaration of Conflict of Interest: The authors declare that they have no conflict of interest with respect to their authorship or the publication of this article.

Author Contributions: SQ and JT designed the study. SCW, IUH, and SK prepared its materials. YW, SCW, and SQ contributed to data collection. SQ and SK performed the data analysis. SQ, JT, and YW drafted the paper. SK, SCW, and IUH provided critical revisions. All authors approved the final version of the paper for submission.

Acknowledgment: The research was supported by a capstone grant awarded to SK by New York University Abu Dhabi.
Abstract

Encoding the internal features of unfamiliar faces poses a perceptual challenge that occasionally results in face recognition errors. Extensive experience with faces framed by a headscarf may, however, enhance perceivers’ ability to process internal facial information. To examine this claim empirically, participants in the United Arab Emirates and the United States of America completed a standard part-whole face recognition task. Accuracy on the task was examined using a 2 (perceiver culture: Emirati vs. American) x 2 (face race: Arab vs. White) x 2 (probe type: part vs. whole) x 3 (probe feature: eyes vs. nose vs. mouth) mixed measures analysis of variance. As predicted, Emiratis outperformed Americans on the administered task. Although their recognition advantage occurred regardless of probe type, it was most pronounced for Arab faces and for trials that captured the processing of nose or mouth information. The findings demonstrate that culture-based experiences hone perceivers’ face processing skills.

Abstract Word Count: 150 words

Key Words: cross-cultural differences, face discrimination, face matching, person identification
The Headscarf Effect Re-Visited: Further Evidence For A Culture-Based Internal Face Processing Advantage

Face researchers habitually marvel at people’s ability to differentiate between hundreds of human faces in everyday life (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000). Their astonishment usually reflects the fact that faces of a similar age, race, and sex tend to look highly alike, requiring perceivers to note rather fine-grained perceptual variations between them (Bruce & Humphreys, 1994). Although this process frequently unfolds accurately, it can be compromised upon encountering unfamiliar others (Alenezi & Bindemann, 2013; Bindemann, Avetisyn, & Rakow, 2012). Such difficulties can become particularly worrisome in contexts that call for the reliable identification of a specific individual, such as border control checks or police line-ups (Burton & Jenkins, 2011).

Initial attempts to study the nature of errors in unfamiliar face recognition have revealed that accurate identification is partially hindered by perceivers’ tendency to encode a face’s external appearance (e.g., its hair or head pose) at the cost of its internal morphology (e.g., its eyes, nose, and mouth; Frowd et al., 2012; Nachson, Moscovitch, & Umlita, 1995). In other words, when encountering unfamiliar faces, people are easily distracted by cues that are poor markers of a person’s unique identity (Fletcher, Butavicius, & Lee, 2008; Toseeb, Keeble & Bryant, 2012). It must be noted, however, that this observation rests largely upon research with participants from Western, Educated, Industrialized, Rich and Democratic societies (i.e., WEIRD participants, see Henrich, Heine, & Norenzayan, 2010). Yet, conclusions based on
these studies may not generalize to large parts of the world’s population, given that face recognition skills are profoundly shaped by experience (Tarr & Cheng, 2003).

The ability to recognize faces of a certain race, for instance, is largely a product of a person’s exposure to such faces (Rossion & Michel, 2011). Similarly, the culture-specific exposure to headscarf-framed faces has been argued to hone a person’s ability to process internal facial information (Megreya, Memon, & Havard, 2012; Toseeb, Bryant, & Keeble, 2014). In support of this claim, Egyptians have recently been found to display a reliable internal face recognition advantage over British perceivers (Megreya & Bindemann, 2009). Given that both groups of participants differ in other sociocultural practices aside from their headscarf use, however, the ultimate cause of this processing advantage remains speculative (Norenzayan & Heine, 2005). Equally uncertain is its underlying mechanism, as increases in internal face processing can result from changes in both the holistic as well as the featural encoding of faces. To address these questions further, the current study re-visited the effects of regular exposure to headscarf-framed faces on holistic and featural internal face processing in a novel cross-cultural context.

**Method**

**Participants**

Arab citizens of the United Arab Emirates (i.e., Emiratis) were expected to have more exposure to faces framed by headscarves than White citizens of the United States of America (i.e., Americans) due to prevalent differences in clothing practices across the two cultures (Thomas, 2013). To ensure 80% power of detecting a medium size effect
(cf. study 4, Megreya & Bindemann, 2009) in the current study, we recruited a total of 86 Emirati (68 females, 18 males) and 84 American (66 females, 18 males) students for participation. Data of one Emirati woman were excluded due to a lack of task compliance (i.e., she kept checking her phone during task completion). Participants included in the final sample were between 18 and 37 years of age (Emirati $M = 21.01$, $SD = 2.08$; American $M = 20.38$, $SD = 2.95$; $t(167) = 1.61, p = .11$) and participated in exchange for course credit or money ($10.00$/AED 50.00). All participants provided informed consent prior to study participation. The study followed a protocol that was jointly approved by the Institutional Review Boards of Zayed University, New York University Abu Dhabi, and New York University New York.

**Procedure**

Emirati participants were recruited at Zayed University (in Abu Dhabi, UAE) and American participants were recruited at New York University (in New York, USA). Since English is the main language of instruction at both institutions, identical procedures and materials were used across sites. All participants were tested individually in a quiet room. Upon arrival, an experimenter asked participants to complete a standard part-whole face recognition task (Tanaka & Farah, 1993) and a brief questionnaire (modified from Islam & Hewstone, 1993). For the face recognition task, participants were seated at a desk equipped with a MacBook Pro laptop (15 inch anti-glare display, screen resolution 1680 x 1050 pixels). Instructions and experimental stimuli were presented and responses recorded using Psyscope Software (Build 57; Cohen, MacWhinney, Flatt, & Provost, 1993).
The part-whole face recognition task was set-up as a two-alternative, forced-choice procedure. Each trial began with a central fixation cross for 500 ms. A whole study face was then presented centrally for 1500 ms. After a 300 ms blank black screen, two probe stimuli were presented until the participant responded. Participants chose the probe that correctly matched the initial study face (or its parts) by pressing the appropriate button (A = left probe, L = right probe) on the keyboard. An intertrial interval of 700 ms elapsed before the next trial began (see Figure 1). Throughout the entire task, each face was shown six times as a study stimulus. Each set of probe stimuli consisted of the study face (or its parts) and one distractor face (or its parts). The position of the study face (or its parts) was equally often on the right and the left side of the computer screen. Half of the probe stimuli consisted of whole faces varying in only one part (i.e., the mouth, nose, or the eyes; each comprising 1/3 of trials), whereas the other half comprised the isolated face parts. Arab whole trials, Arab part trials, White whole trials, and White part trials were randomly intermixed and spread across two blocks of 120 trials each, giving participants a short, self-paced break during task completion. Block order was counterbalanced across participants and six practice trials displaying training faces were administered before the actual experiment.

Following the computerized face recognition task, participants completed a paper-and-pencil questionnaire to record their demographic information and their exposure to headgear-framed faces. A five item scale was constructed (cf. Islam & Hewstone, 1993) with a stem statement, “In everyday life, how much contact do you
have with faces that are covered by a hijab/ghutra” followed by: (1) “… in public settings (e.g., malls, restaurants)?”, (2) “… in your neighborhood?”, (3) “… at work or university?”, (4) “… in your circle of friends?”, (5) … through the media (e.g., on TV, in magazines etc. Response options ranged from 1 (= none at all) to 7 (= very much). In the American version of the questionnaire a picture of a woman wearing the hijab and a man wearing the ghutra illustrated the referred to headgear. Perceivers’ average scores on the reliable 5-item scale (α = .76) confirmed that Emiratis experienced more exposure to headgear-framed faces (M = 4.89, SD = .94) than Americans [M = 2.90, SD = 1.02; t(167) = 13.10, p < .001]. Upon questionnaire completion, participants were debriefed and thanked for their participation.

**Stimulus Material**

To control for effects of race during task completion, facial photographs of 20 Arab (10 female) and 20 White (10 female) individuals were used. All faces were unfamiliar to the perceivers and displayed people of approximately 18 to 25 years of age. Photographs were selected from relevant face databases of Arab (e.g., Megreya & Burton, 2008; van der Schalk, Hawk, Fischer, & Doosje, 2011) and White faces (e.g., Langner et al., 2010; Minear & Park, 2004). In addition, eight Arab females were photographed at the final author’s lab to overcome a shortage in the available databases. All faces were free of facial hair, accessories, or visible make-up and displayed a neutral frontal pose with direct gaze. In order to specifically probe internal face processing, all faces were

---

1To further disambiguate the stem statement, participants were also orally informed that we meant coverage of a face’s external features (e.g., hair, ears) rather than its internal appearance.
cropped to remove external facial features such as ears, hair, or garments that could aid facial recognition. Cropped faces were grey-scaled, standardized to a common height (340 pixels), and centrally inserted on a uniform black background (350 x 350 pixels). For the whole condition of the recognition task, additional distractor faces were created for each study face by replacing the mouth, nose, or eyes of the original individual with a corresponding feature of another gender- and race-matched individual using Adobe Photoshop (Version 12.0.4). Distractor features as well as original features of the study face were additionally inserted on a uniform black background (350 x 350 pixels) to act as part images (see Figure 1). Finally, study faces were presented in their standardized size (350 x 350 pixels), whereas probe stimuli were presented slightly reduced in size (320 x 320 pixels) to prevent participants from using low-level visual matching strategies to solve the task. Corresponding study and probe stimuli were otherwise identical.

Results

An analysis of variance was performed on recognition accuracy (see Table 1, Figure 2) with the between-subjects factor perceiver culture (American vs. Emirati) as well as the within-subject factors face race (Arab vs. White), probe type (part vs. whole), and probe feature (eyes vs. nose vs. mouth). A significant culture by probe feature interaction \( F(2,334) = 3.37, p = .035, \eta^2_p = .020 \) qualified a marginally significant main effect of culture \( F(1,167) = 2.90, p = .091, \eta^2_p = .017 \). Thus, Emiratis outperformed Americans on nose trials [Emirati: \( M = 68.99, SD = 9.37 \); American: \( M = 65.37, SD = 8.02 \); \( t(167) = 2.69, p = .004 \) (one-sided), \( d = 0.42 \)] and marginally so on mouth trials [Emirati: \( M = 70.12, SD = 8.36 \); American: \( M = 68.02, SD = 8.90 \); \( t(167) = 1.58, p = .058 \).
For eye trials, however, both groups performed equally well
[Emirati: $M = 83.66$, $SD = 9.41$; American: $M = 84.05$, $SD = 9.48$; $t(167) = 0.27$, $p = .396$
(one-sided), $d = .04$].

Additionally, a marginally significant culture by face race interaction emerged
$[F(1,167) = 3.19$, $p = .076$, $\eta^2_p = .019]$, signaling that the processing advantage was more pronounced for Arab faces [Emirati: $M = 76.86$, $SD = 7.43$; American: $M = 74.39$, $SD = 7.28$; $t(167) = 2.18$, $p = .016$ (one-sided), $d = .34$] than for White faces [Emirati: $M = 71.65$, $SD = 7.70$; American: $M = 70.57$, $SD = 6.44$; $t(167) = .99$, $p = .162$ (one-sided), $d = .15$]. Importantly, neither the two-way interaction of culture with probe type
$[F(1,167) = 0.99$, $p = .322$, $\eta^2_p = .006]$, nor any culture-related three-way interaction
[culture x face race x probe type: $F(1,167) = 2.16$, $p = .144$, $\eta^2_p = .013$; culture x face race x probe feature: $F(2,334) = 0.33$, $p = .721$, $\eta^2_p = .002$; culture x probe type x probe feature: $F(2,334) = 0.56$, $p = .570$, $\eta^2_p = .003$] or four-way interaction [culture x face race x probe type x probe feature: $F(2,334) = 0.14$, $p = .866$, $\eta^2_p = .001$] reached statistical significance.

Less relevant in the context of the current study, but reported for reasons of completeness (cf. Tanaka & Farah, 1993), are additional main effects of face race
$[F(1,167) = 135.63$, $p < .001$, $\eta^2_p = .448]$, probe type
$[F(1,167) = 95.45$, $p < .001$, $\eta^2_p = .364]$, and probe feature
$[F(2,334) = 275.90$, $p < .001$, $\eta^2_p = .623]$. These main effects were qualified by several two-way interactions
[probe type x probe feature: $F(2,334) = 14.25$, $p < .001$, $\eta^2_p = .079$; face race x probe feature: $F(2,334) = 16.20$, $p < .001$, $\eta^2_p = .088$; face race x probe type: $F(1,167) = 2.96$, $p = .087$, $\eta^2_p = .017$] as well as a marginally significant three way interaction
[face race x probe type x probe feature: $F(2,334) = 1
2.34, \( p = .098 \), \( \eta^2_p = .014 \). A final analysis examined correlations between participants’ self-reported everyday exposure to headgear-framed faces and their face recognition skills, but revealed no significant results (see Table 2).

**Discussion**

The successful recognition of unfamiliar faces based on their internal appearance can be surprisingly poor (Burton & Jenkins, 2011; Fletcher et al., 2008). A life-long experience with faces framed by a headscarf may, however, train perceivers’ ability to accurately process internal face information. To re-examine this claim, the current study compared the performance of Emirati and American perceivers on a standard part-whole recognition task. As expected, Emiratis outperformed Americans in the recognition of unfamiliar faces based on their internal appearance. This finding replicates an earlier report of increased internal face processing skills in individuals with culturally enhanced exposure to headscarf-framed faces (Megreya & Bindemann, 2009). The current findings go beyond existing work by demonstrating that the exact size of the effect is modulated by probe feature. In other words, Emiratis showed a medium-sized recognition advantage for nose trials, a small advantage for mouth trials, but no advantage for eye trials. In addition, these recognition advantages were more pronounced for same-race than other-race target faces, but occurred for both facial features and facial wholes.

Further inquiry is necessary to learn whether these recognition differences may reflect culture-dependent variations in face scanning patterns (cf. Blais, Jack, Scheepers, Fiset, & Caldara, 2008). It must be noted, however, that the observed
processing advantages for lower portions of the face (such as the mouth or the nose) did not simply result from a processing disadvantage for the eye region. In other words, people’s cultural background did not merely re-direct their attention to different parts of the face. Instead, despite equivalent recognition of the eye region, Emiratis showed additional improved processing of nose and, marginally so, of mouth information. Although the null finding with regard to eye trials requires further investigation (as previous work did not examine the effect of probe feature), it may reflect the fact that even WEIRD perceivers do particularly well on these trials, limiting the room for improvement in other cultures (see Table 1; cf. Tanaka & Farah, 1993).

It should also be noted that the current study did not find any significant correlations between perceivers’ exposure to headgear-framed faces and their internal face processing skills as measured via the part-whole recognition task. This lack of association may not be surprising, given that exposure was measured in reference to perceivers’ momentary experience. In other words, the questionnaire scores confirmed that in everyday life Emirati participants were generally more exposed to headscarf-framed faces than American participants. However, the questionnaire did not capture participants’ accumulated lifetime experience with headgear-framed faces within their respective cultural contexts. Considering that face perception skills are particularly honed during people’s childhood and puberty (Carey, 1981; Lawrence et al., 2008; Itier & Taylor, 2004), future work should aim to quantify accumulated exposure in order to specifically predict internal face recognition skills in adulthood.

Furthermore, speculations about the actual cause of the observed cultural differences remain limited by the fact that the two cultures investigated differ across
important sociocultural factors beyond their difference in headscarf use. Anecdotal evidence suggests, for instance, that Emiratis assign particular importance to a person's nose when it comes to judgments of beauty. The phrase 'sallat al-saif' is commonly used in Emirati poetry to describe a nose that looks like a drawn sword (i.e., a nose characterized by a straight bridge, narrow nostrils, and a pointed tip). Emiratis also frequently practice the custom of greeting close friends or family members with a nose rub (yukhashim). In consequence, Emiratis may be particularly good at encoding information about another person's nose because of its particular cultural significance for rapid judgments of beauty and kinship.

Alternatively, the high prevalence of consanguineous marriages in the UAE (Thomas, 2013, Table 4.2) may have created social environments in which the faces of different individuals are physically more alike than in countries with a low prevalence of consanguineous marriages such as the USA or the UK. Given that consanguinity is also common in Egypt (Temptamy & Aglan, 2012), perceivers in both the UAE and Egypt (as previously tested) may display enhanced internal face processing skills ensuing from a life-long experience with faces of particular similarity due to enhanced family resemblance. In order to rule out these alternative explanations for cross-cultural differences in internal face processing skills, future research should examine the effects of prolonged and accumulating exposure to headscarf-framed faces in an experimental manner (e.g., by conducting training studies with headscarf-framed faces, cf. Bate & Bennetts, 2014).
References


Table 1.

Mean accuracies in percentage (including their standard deviations) on the part-whole face recognition task based on perceiver culture and trial type. Cohen’s $d$ quantifies the size of the observed difference across perceiver culture for each trial type.

<table>
<thead>
<tr>
<th>Probe Type</th>
<th>Probe Feature</th>
<th>Americans</th>
<th>Emiratis</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arab Faces</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parts</td>
<td>Eyes</td>
<td>83.81 (12.89)</td>
<td>84.94 (10.22)</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>63.87 (12.21)</td>
<td>67.06 (10.86)</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>67.08 (11.70)</td>
<td>70.41 (11.78)</td>
<td>0.29</td>
</tr>
<tr>
<td>Wholes</td>
<td>Eyes</td>
<td>88.15 (10.32)</td>
<td>87.82 (11.32)</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>68.27 (12.65)</td>
<td>72.82 (13.08)</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>75.18 (12.76)</td>
<td>78.12 (13.47)</td>
<td>0.23</td>
</tr>
<tr>
<td><strong>White Faces</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parts</td>
<td>Eyes</td>
<td>81.01 (11.71)</td>
<td>78.82 (11.87)</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>64.76 (12.66)</td>
<td>66.59 (12.28)</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>61.25 (11.06)</td>
<td>61.65 (11.43)</td>
<td>0.04</td>
</tr>
<tr>
<td>Wholes</td>
<td>Eyes</td>
<td>83.21 (12.29)</td>
<td>83.06 (12.39)</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>64.58 (10.42)</td>
<td>69.47 (14.76)</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>68.57 (11.73)</td>
<td>70.29 (11.11)</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Table 2.

*Correlation coefficients (including their corresponding p-values) between perceivers’ everyday exposure to headgear-framed faces and their accuracy on the part-whole face recognition task.*

<table>
<thead>
<tr>
<th>Probe Type</th>
<th>Probe Feature</th>
<th>Americans</th>
<th>Emiratis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arab Faces</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parts</td>
<td>Eyes</td>
<td>-.04 (.738)</td>
<td>.10 (.377)</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>-.11 (.332)</td>
<td>-.18 (.091)</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>.19 (.085)</td>
<td>-.08 (.446)</td>
</tr>
<tr>
<td>Wholes</td>
<td>Eyes</td>
<td>-.07 (.551)</td>
<td>-.12 (.285)</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>-.01 (.941)</td>
<td>-.20 (.074)</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>.02 (.878)</td>
<td>.11 (.306)</td>
</tr>
<tr>
<td><strong>White Faces</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parts</td>
<td>Eyes</td>
<td>.00 (.993)</td>
<td>-.07 (.514)</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>.20 (.069)</td>
<td>-.12 (.282)</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>.03 (.765)</td>
<td>-.17 (.124)</td>
</tr>
<tr>
<td>Wholes</td>
<td>Eyes</td>
<td>-.08 (.454)</td>
<td>-.04 (.688)</td>
</tr>
<tr>
<td></td>
<td>Nose</td>
<td>.08 (.470)</td>
<td>-.11 (.335)</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
<td>.10 (.361)</td>
<td>.05 (.650)</td>
</tr>
<tr>
<td><strong>Overall Accuracy</strong></td>
<td></td>
<td>.05 (.667)</td>
<td>-.12 (.272)</td>
</tr>
</tbody>
</table>
Figure 1. Sample trials from the part-whole face recognition task. The study stimulus was always a whole face; probe stimuli required participants to either select from two whole faces (left) or two facial parts (right). Portrayed individuals gave written consent to appear as illustrations in this article.
Figure 2. Mean differences in recognition accuracy for American and Emirati perceivers based on trial type (in percentage). Positive scores signal a processing advantage for Emiratis. Error bars show the 95% confidence interval of the difference score.