Optimal face recognition performance involves a balance between global and local information processing: Evidence from cultural difference

Zhijie Cheng (chengzj@hku.hk)
Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong

William G. Hayward (whayward@hku.hk)
Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong
Department of Psychology and ARC Centre of Excellence in Cognition and Its Disorders, The University of Auckland, Private Bag 92019, Auckland, New Zealand

Antoni B. Chan (abchan@cityu.edu.hk)
Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Hong Kong

Janet H. Hsiao (jhsiao@hku.hk)
Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong

Abstract

In face recognition, eye gaze to the eye region is reported to be associated with better performance than to the center of a face. Nevertheless, Caucasians and Asians differ in how much they look at the eyes when they scan a face, but have comparable identification performance. To resolve this issue, here we test the hypothesis that optimal face recognition performance involves a balance between global and local face processing. Thus, Asians may benefit from enhancement of local processing and vice versa for Caucasians. We showed that local attention priming using hierarchical letter stimuli led to more eye-focused eye movement patterns compared to global attention priming in both Asians and Caucasians. However, Asians had better performance after local priming than global priming, whereas Caucasian showed the opposite effect. These results suggest that engagement of global/local attention leads to face-center/eye biased eye movements respectively, and optimal recognition performance involves both global and local processing/gaze transitions between the face center and eyes.

Keywords: eye movement, face recognition, cultural difference, hidden Markov model, EMHMM

Introduction

Humans have a remarkable ability to recognize individual faces. Nevertheless, it remains unclear what kind of information use can lead to optimal face recognition performance. Recent studies have reported substantial individual differences in eye movement patterns in face recognition (Peterson & Eckstein, 2013), which may reflect individual differences in information use and recognition performance. To account for these individual differences in eye movement data analysis, Chuk, Chan, and Hsiao (2014) proposed the Eye Movement analysis with Hidden Markov Models (EMHMM) approach, in which they modeled each participant’s eye movement pattern in face recognition with a hidden Markov model (HMM, a type of machine learning model for time series data), including personalized regions of interest (ROIs) and transition probabilities among the ROIs. Through clustering these individual models according to their similarities, they discovered two common patterns: “holistic” pattern, in which observers mainly looked at the face center; and “analytic” pattern, in which observers looked at the eye region in addition to the face center. Interestingly, analytic patterns were associated with better recognition performance (Chan, Chan, Lee, & Hsiao, 2018; Chuk, Chan, & Hsiao, 2017; Chuk, Crookes, Hayward, Chan, & Hsiao, 2017) and higher activations in brain regions important for top-down visual attention control such as the frontal eye field and the intraparietal sulcus (Chan, Wong, Chan, Lee, & Hsiao, 2016). In contrast, holistic patterns were associated with cognitive decline in older adults (Chan et al., 2018). Miellet, Caldara, and Schyns (2011) found that during face viewing, looking at the nose/face center was associated with global information processing, whereas looking at the eyes was associated with local information processing. Accordingly, since analytic patterns involve looking at both the face center and the eyes, their advantage in recognition performance may be due to the use of both global and local information. In other words, optimal face recognition performance may require both global and local information processing. Consistent with this speculation, while global/configural information is believed to be essential for face processing (e.g., Galton, 1883; Richler, Cheung, & Gauthier, 2011; Tanaka & Farah, 1993), recent studies have suggested the importance of local/featural information in addition to global information (e.g., Burton, Schweinberger, Jenkins, & Kaufmann, 2015; Cabeza & Kato, 2000).

Recent research has suggested that East Asians and West Caucasians differ in cognitive style: Asians are more likely to attribute the cause of an event to the context (holistic cognition), whereas Caucasians are more likely to attribute the cause of an event to isolated objects (analytic cognition; e.g., Nisbett & Miyamoto, 2005). This cultural difference is reflected in their eye movements in scene viewing: Asians looked at the background more often and are more attuned to contextual information, whereas Caucasians pay more
attention to salient foreground objects and are less sensitive to contexts (e.g., Masuda, Ishii, & Kimura, 2016; Miyamoto, Nisbett, & Masuda, 2006; Nisbett, Choi, Peng, & Norenzayan, 2001). Some studies have reported that this cultural difference could also be observed in eye movements in face recognition: at the group level, Asians are shown to predominantly fixate on the face center, whereas Caucasians fixated mainly on the eyes and the mouth (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Caldara, Zhou, & Miellet, 2010; Kelly, Miellet, & Caldara, 2010; Miellet, He, Zhou, Lao, & Caldara, 2012). This phenomenon suggests that during face recognition, Asians may rely more on global information processing, whereas Caucasians engage more local information processing. Consistent with this speculation, when information outside central vision was restricted, Asians fixated at the eyes and mouth much like Caucasians, whereas when central vision was masked and peripheral vision was preserved, Caucasians started to look at the face center (Caldara, Zhou, & Miellet, 2010; Miellet, He, Zhou, Lao, & Caldara, 2012). Note however that regardless of this cultural difference in eye movement pattern and information use, the two cultural groups did not differ in face recognition performance (Blais et al., 2008; Caldara et al., 2010; Miellet et al., 2012). Since previous studies have suggested that optimal face recognition performance may involve both global and local information processing, Asians may benefit from enhancement of local face processing, whereas Caucasians may benefit from enhancement of global face processing.

To examine this possibility, here we used the Navon stimuli (Navon, 1977) to induce global and local face processing in Asian and Caucasian participants during face recognition. Navon stimuli are hierarchical stimuli with a global figure composed of local components and have been widely used to prime global or local attention biases (e.g., Hübner, 2000; Large & McMullen, 2006; Shedden, Marsman, Paul, & Nelson, 2003; Ward, 1982). We predicted that local priming may lead to better face recognition performance than global priming in Asians, and this effect may be associated with increased eye fixations to the eye region. In contrast, Caucasians may have better recognition performance after global priming than local priming, and this effect may be associated with increased eye fixations towards the face center.

Methods

Participants

35 Chinese participants (12 male, mean age 22.1, SD = 4.23) from the University of Hong Kong and 24 Caucasian participants (13 male, mean age 24.04, SD = 3.98) from the University of Auckland were recruited. All had normal or corrected-to-normal vision and reported right-handed except for 2 left-handers (1 Asian and 1 Caucasian).

Materials

Navon stimuli Sixteen hierarchical letters were created using letters D, E, F, and H in bold Helvetica font. They were white in color and presented on a black background. Each local letter subtended 0.8° × 1.2° visual angle under a 60 cm viewing distance. Each global letter consisted of 13 to 17 local letters and subtended 5.5° × 6.7° (Fig. 1).

Figure 1: Samples of Navon letters.

Face stimuli Images of 120 Chinese faces and 120 Caucasian faces with neutral expressions were used (half male and half female in each race). External features such as hair and ears were removed. All faces were grey-scaled with equal luminance and scaled and aligned with standard eye-to-eye and eye-to-mouth distances. Each face subtended 6° × 8° of visual angle. Chinese faces were used for Asian participants and Caucasian faces were used for Caucasian Participants.

Design and Apparatus

The design consisted of a between-subject variable group (Asian vs. Caucasian) and a within-subject variable priming level (baseline vs. global vs. local). Stimuli were shown on a 22” monitor with 1024 x 768 resolution. Participants sat in front of the screen with a chinrest to limit their head movement. Eye movements were recorded with an Eyelink 1000 eye tracker (sampling rate 1000 Hz). Participants viewed the stimuli with binocular vision, but only the dominant eye was tracked. A nine-point calibration procedure was used before the task. Drift correction was performed in the beginning of each trial. The calibration procedure was repeated when drift correction error was larger than 1° of visual angle.

Procedure

Asian and Caucasian participants performed the same face recognition task with face stimuli of their own race. Each participant performed three blocks of old-new judgment task: baseline, global priming, and local priming blocks. In the baseline block, during the study phase, 20 faces were shown on the screen one at a time, for 5 s each. Participants were asked to view and remember the faces. After a 5-minute break, in the test phase, the 20 old faces together with 20 new faces were presented one at a time in a random order. The position of each face was randomly assigned to be either at the upper or lower center of the screen. Participants made old/new judgments using a keyboard. Each face was presented until response.

In the global and local priming blocks, participants performed the same study phase as the baseline block. Afterwards, instead of having a 5-minute break, participants
performed a 5-minute Navon task (162 trials) between the study and test phases. In each trial, 2 Navon stimuli were presented simultaneously on the left and right of the screen, each at 5° of visual angle away from the center. In the global priming block, participants judged whether the global form of the stimuli were the same, whereas in the local priming block, they judged whether the local letters of the stimuli were the same. Participants made responses through a keyboard. The test phase procedure was similar to that in the baseline block, except that in each trial, participants performed a trial of the Navon task before the presentation of the face in order to maintain the priming effect. The order of the 3 blocks was counterbalanced across participants.

Results

Performance of the Navon task in the priming blocks

We conducted a 2 x 2 mixed ANOVA with group (Asian vs. Caucasian) and level (global vs. local) as the independent variables. In the 5-minute Navon task prior to the test phase, participants responded faster in matching global forms ($M = 634.78$ ms, $SE = 16.20$) than local letters ($M = 786.99$ ms, $SE = 17.18$), $F(1, 57) = 80.24$, $p < .001$. In the Navon trials during the test phase, participants were more accurate, $F(1, 57) = 4.974$, $p = .030$, and faster, $F(1, 57) = 75.91$, $p < .001$, in matching global forms (response time: $M = 822.08$ ms, $SE = 17.09$; accuracy: $M = 98.81\%$, $SE = 0.23$) than local letters (response time: $M = 960.20$ ms, $SE = 19.40$; accuracy: $M = 98.01\%$, $SE = 0.42$). No other main effect or interaction was found ($p > .12$). These results reflected the global precedence effect in visual perception.

Face recognition performance

Face recognition performance was measured by $d'$. A mixed ANOVA showed a significant interaction between group and priming level, $F(2, 114) = 3.714$, $p = .027$ (Fig. 2). Post-hoc comparisons showed that in Asians, local priming led to better performance than global priming, $t(34) = 2.21$, $p = .025$, whereas in Caucasians, a marginal effect indicated better performance after global priming than local priming, $t(24) = -1.95$, $p = .055$. This result was consistent with our hypothesis: Asians benefited more from enhancement of local face processing whereas Caucasians benefited more from enhancement of global face processing. There was no main effect of group or priming level, and no significant effect was found in response time ($p > .17$).

Eye movement pattern analysis

We used the EMHMM approach (Chuk et al., 2014. See http://visal.cs.cityu.edu.hk/research/emhmm/ for details) to quantitatively assess eye movement pattern changes due to priming. Following previous studies (Chuk et al., 2014, 2017; Chan et al., 2018), we used the first three fixations in each trial in the analysis since these fixations were shown to be particularly relevant to recognition performance (Chuk et al., 2017). Chan et al., (2018) identified representative holistic and analytic eye movement patterns for face recognition from a large sample across a large age span (34 young and 34 older adults) through clustering using the EMHMM approach (Fig. 3). The holistic pattern focused at the face center, whereas the analytic pattern focused at the eye region in addition to the face center. They assessed the similarity of an individual’s eye movement pattern to the representative holistic/analytic pattern as the log-likelihood of the individual pattern being generated by the HMM of the representative pattern. They then developed the Holistic-Analytic scale (H-A scale) to quantitatively assess one’s eye movement pattern along the holistic-analytic dimension: H-A scale = (holistic log-likelihood – analytic log-likelihood) / (holistic log-likelihood + analytic log-likelihood). Higher score indicated higher similarity to the holistic pattern. They found that participants’ H-A scale was correlated with their cognitive performance. Interestingly, the two representative models could be used to assess new participants’ H-A scale of eye movement patterns and show similar correlations. This result demonstrated the robustness of the representative models in quantifying one’s eye movement patterns. Since here we used the same face recognition task and image size as Chan et al. (2018), we used their two representative

Fig. 2. Face recognition performance

Fig. 4. H-A Scale of Asians’ and Caucasians’ eye movement patterns in face recognition.

---

1 Chuk, Chan, and Hsiao (2017) found that face recognition performance did not correlate with the similarity between the eye movement patterns during face learning and recognition, and it was only correlated with eye movement patterns during recognition but not learning, suggested that eye movement patterns during encoding does not play an important role, but the retrieval of diagnostic information during recognition is essential for recognition. Giving that test phase play an important role and to exclude the confounding of encoding, we only manipulated the eye movement during test phase.
Here we tested the hypothesis that optimal face recognition performance involves a balance between global and local information processing through comparing Asians’ and Caucasians’ recognition performance and eye movement pattern changes in response to global and local attention priming using Navon stimuli. We first showed that without priming, Asians showed more face-center-focused, holistic eye movement patterns than Caucasians; nevertheless, the two groups did not differ in recognition performance. This result was consistent with previous findings in the literature (e.g., Blais et al., 2008). We then showed that global priming elicited more face-center-focused, holistic patterns, whereas local priming elicited more eye-centered, analytic patterns. Although this effect was consistent among Asians and Caucasians, local priming led to better face recognition performance than global priming in Asians, whereas Caucasians had better recognition performance after global priming than local priming. This result was consistent with our hypothesis, suggesting that optimal face recognition performance involves a balance between global and local face processing.

Chuk et al., (2017) examined cultural differences in face recognition with similar paradigm and eye movement data analysis methods (EMHMM) to the current study. In contrast to our results, they did not find strong evidence suggesting cultural difference in eye movement patterns.
More specifically, they recruited 24 Asian and 24 Caucasian young adult participants and discovered three representative eye movement patterns through clustering: holistic, left-eye-biased analytic, and right-eye-biased analytic. They found that the two race groups did not differ either in the log-likelihood or in the frequency of adopting the three patterns. Note however that their representative patterns were directly discovered from the 48 young adults, whose eye movement patterns in face recognition are shown to be more eye-focused than older adults (Chan et al., 2018). In contrast, the representative patterns used in the current study were developed from a larger sample with both young and older adults and captured better the difference between eye-focused and face-center-focused eye movement patterns (Chan et al., 2018). Also, their face images subtended 8° of visual angle horizontally, larger than the ones used here (6°). Thus, their representative patterns tend to be more eye-focused in general as compared with the ones used here (Fig. 6). Indeed, previous studies have shown that image size/viewing distance is an important factor influencing holistic face processing, as the effect declined sharply at viewing distances shorter than 2 meters (McKone, 2009; Ross & Gauthier, 2015). Since the image size used in Chuk et al. (2017) resembled the size of a real face under a viewing distance of 1 meter, both Asian and Caucasian observers might engage less global face processing, and consequently the cultural difference in eye movement patterns diminished. Thus, image size/viewing distance may be an important factor to consider in the examination of cultural difference in eye movements.

In the current study, we found that global/local attention priming using Navon stimuli had similar effects on eye movement patterns in Asians and Caucasians: Participants’ eye movement patterns were more holistic (face-center-focused) after global priming and more analytic (eye-focused) after local priming. This result is consistent with Miellet et al. (2011) and Lemieux, Collin and Nelson (2014), suggesting a strong link between engagement of local/global attention and eye movements in face recognition. Nevertheless, regardless of the consistent direction of change in eye movement pattern, Asians and Caucasians showed contrasting priming effects on recognition performance: Asians performed better after local priming than global priming, whereas Caucasians performed better after global priming than local priming. Since Asians’ eye movement patterns were more holistic whereas Caucasians’ were more analytic in the baseline condition, local priming may have helped Asians to direct attention to the eyes and global priming helped Caucasians to better process global information to facilitate recognition. This result suggests that optimal face recognition performance involves a balance between global and local information processing, consistent with recent studies suggesting the importance of both featural and configural information in face recognition (Burton et al., 2015; Cabeza & Kato, 2000). This result also suggests an inverted-U shape relationship between face recognition performance and eye movement patterns, where the optimal performance may be observed somewhere between the two extremes along the holistic-analytic dimension. Nevertheless, with the current sample, we did not observe a significant quadratic relationship between recognition performance and H-A scale of eye movements. We speculate that the variance in our current sample may be inadequate to reveal this potential relationship, since all participants were young adults, whose eye movements tend to be more analytic than older adults (Chan et al., 2018). Indeed, the range of H-A scale scores of the older adults in Chan et al. (2018) was from -.06 to .06, whereas in the current study with young adults, it was from -.08 to .02. With this H-A scale score range, we also failed to replicate the negative correlation between face recognition performance and H-A scale observed in Chan et al. (2018) (in the current Asian sample, r(103) = -.030, p = .764). Future work will examine this potential inverted-U shape relationship between face recognition performance and eye movement patterns with a larger, more representative participant sample.

Note that for both Asians or Caucasians, the priming procedure adopted here did not significantly improve recognition performance when compared with the baseline condition. This may be because in the priming blocks participants had to perform both the Navon and face recognition tasks, and thus their face recognition performance was interfered by the Navon task. Thus, it remains unclear whether it is possible to use attention priming to improve one’s face recognition performance. Indeed, recent studies have suggested that adults have limited plasticity for face recognition ability due to abundant experience with faces that may have led to the maximum level of capacity (Tree, Horry, Riley, & Wilmer, 2017). Future work will examine possible attention priming procedures that may improve face recognition performance.

In conclusion, here we showed that in face recognition, global and local attention priming could induce holistic (face-center-focused) and analytic (eye-focused) eye movement patterns respectively across cultures, suggesting a link between eye movement patterns and global/local information use. Nevertheless, Asians’ face recognition performance benefited more from local than global attention priming due to their tendency to adopt a holistic perceptual style, whereas Caucasians, who were more analytic, showed the opposite effect. These results suggest that optimal face
recognition performance involves a balance between global and local information processing through gaze transitions between the face center and the eye region.

Acknowledgments
We thank Tim Chuk for help in the EMHMM methods. We are grateful to the RGC (#17402814 to Hsiao; CityU110513 to Chan).

References