Culture shapes eye movements for visually homogeneous objects

David J. Kelly*, Sébastien Miellet and Roberto Caldara*

Department of Psychology and Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow, UK

Culture affects the way people move their eyes to extract information in their visual world. Adults from Eastern societies (e.g., China) display a disposition to process information *holistically*, whereas individuals from Western societies (e.g., Britain) process information *analytically*. In terms of face processing, adults from Western cultures typically fixate the eyes and mouth, while adults from Eastern cultures fixate centrally on the nose region, yet face recognition accuracy is comparable across populations. A potential explanation for the observed differences relates to social norms concerning eye gaze avoidance/engagement when interacting with conspecifics. Furthermore, it has been argued that faces represent a 'special' stimulus category and are processed *holistically*, with the *whole* face processed as a single unit. The extent to which the *holistic* eye movement strategy deployed by East Asian observers is related to *holistic* processing for faces is undetermined. To investigate these hypotheses, we recorded eye movements of adults from Western and Eastern cultural backgrounds while learning and recognizing visually homogeneous objects: human faces, sheep faces and *greebles*. Both group of observers recognized faces better than any other visual category, as predicted by the specificity of faces. However, East Asian participants deployed central fixations across *all* the visual categories. This cultural perceptual strategy was not specific to faces, discarding any parallel between the eye movements of Easterners with the *holistic* processing specific to faces. Cultural diversity in the eye movements used to extract information from visual homogenous objects is rooted in more general and fundamental mechanisms.

Keywords: culture, eye movements, face processing

INTRODUCTION

The term 'culture' is typically used to describe the particular behaviors and beliefs that characterize a social or ethnic group. Thus, by definition, culture represents a powerful deterministic force, which is responsible for shaping the way people think and behave. The potency of culture becomes evident when visiting foreign countries, especially if that country lies beyond our own continent. The cultural differences we observe can evoke feelings of surprise, intrigue and pleasure, but also we may experience confusion and anxiety. These intense feelings and emotions reflect the profound diversity of culture and the power it exerts over humans throughout ontogeny. Of course, such observations are not novel and accordingly it is uncontroversial to claim that culture affects thought and behavior. However, more recently a steadily growing body of literature has yielded evidence to suggest that culture also impacts upon visual perception.

Cultural differences of perception are best understood within the holistic/analytical framework (see Nisbett and Miyamoto, 2005 for a review). According to this view, adults from East Asian (EA; e.g., China, Korea and Japan) and Western Caucasian (WC; e.g., European and North American) cultures perceive the world and subsequently process information in a fundamentally different manner. In the past decade systematic differences have been found in a variety of perceptual tasks and paradigms including: scene perception (e.g., Miyamoto et al., 2006), scene description

(e.g., Masuda and Nisbett, 2001) and perceptual categorization (Norenzayan et al., 2002). All these studies point to a similar pattern of results, revealing that people from Western cultures process information *analytically* by focusing on salient objects and using categorical rules when organizing their environment. By contrast, people from Eastern cultures process information in a more *holistic* manner, showing interest in context and group objects according to relationships. Furthermore, pronounced cultural differences in perception have also been found in a comparison between European vs. African adults (Davidoff et al., 2008). Although the root causes of these different processing styles are currently unresolved, some authors (e.g., Masuda and Nisbett, 2001) have claimed that the organization of social systems is responsible. Western cultures are thought to be *individualistic* and encourage the pursuit of personal goals, which might lead to a bias for processing focal objects within a given context. By contrast, Eastern societies are more *collectivistic* and emphasize the importance of the group over individuals, which could lead to a tendency to think and process visual information in a more global manner. Intriguingly, Davidoff and colleagues report a striking *local* processing bias in the Namibian Himba tribe as measured by a Navon figure task (Navon, 1977), despite the fact that this population lives in a *collectivistic* society. This observation challenges the view that the organization of social systems is the critical factor which accounts for the perceptual cultural differences observed in the holistic/analytical framework. Yet, regardless

of potential theoretical explanations, the ever-growing catalog of cultural findings is challenging long-held notions of universality in perception and human cognition in general, which had previously been assumed.

In addition to presumed universality in general perception, the eye movement strategies employed by adults to encode and recognize facial identity has consistently been shown to be uniform. Typically, adults make fixations to the eye and mouth regions forming a triangular scanpath (e.g., Yarbus, 1965). However, this pattern of behavior has always been recorded in adults from Western cultures (e.g., Groner et al., 1984; Henderson et al., 2005). Blais et al. (2008) recently challenged this supposition by recording eye movements in an EA population. Contrary to expectations, EA participants fixated centrally on the nose region and generally avoided the eyes. This finding is not just a further example of cultural variations in the extraction of visual information, but instead represents the first demonstration of fundamental differences between peoples of different cultures when viewing biologically relevant stimuli. Furthermore and somewhat intriguingly, the distributed fixation patterns displayed by Westerners arguably resembles an *analytical* processing style whereas the central fixation pattern seen in Easterners appears consistent with a *holistic* processing strategy. Indeed, because retinal cell density and visual resolution decrease sharply toward the peripheral visual field, the center of the face is the most physiologically advantageous spatial position to capture facial feature information holistically.

Currently, the relationship between the *holistic* and *analytical* eye movement strategies respectively deployed by Eastern and Western observers and *holistic* vs. *featural* face processing is unclear. It is first necessary to distinguish between 'holistic processing' as defined within the cultural differences literature and 'holistic face processing' as described in the face recognition literature. Some researchers have argued that in contrast to objects, faces are processed in a holistic manner (e.g., Young et al., 1987; Tanaka and Farah, 1993; Hole, 1994; Le Grand et al., 2004). In other words, rather than processing facial features independently, the face is perceived and processed as a whole unit or Gestalt. However, it has also been argued that other-race faces may be processed more featurally/analytically (i.e., by attending to individual features; Tanaka and Farah, 1993). As reported by Blais et al. (2008), the eye movement strategies used by observers are not modulated by the race of face being viewed suggesting that EA and WC adults *extract* information using different strategies, but the information used and underlying cognitive processes are likely to be the same across populations. Regardless of their culture, observers showed a comparable performance in face recognition.

In support of this view is the fact that that fixation location does not unequivocally inform us about information use (Posner, 1980; Kuhn and Tatler, 2005). In the case of face processing, although EA adults may fixate centrally on the nose region under free viewing conditions, this does not mean that the information contained in this region is used to individuate faces. Indeed, available evidence suggests that there is insufficient variation contained in this region to allow accurate face recognition (Goldstein, 1979a,b; Caldara and Abdi, 2006; Caldara et al., 2010). Evidence from other studies suggests that the information used to accurately identify faces is contained in the eye region (e.g., bubbles technique: Gosselin and

Schyns, 2001; Caldara et al., 2005). Therefore, it seems probable that EA adults fixate the nose region when viewing faces, but actually exploit the eye region, processed parafoveally, to recognize faces. This hypothesis has recently been supported by Caldara and colleagues with a gaze-contingent moving aperture design.

Caldara et al. (2010) directly investigated the strategic differences displayed by EA and WC adults by restricting the visual information available to participants with Gaussian apertures (termed *spotlights*) sized 2°, 5° or 8°. In both the 2° and 5° conditions, the aperture was large enough for a single facial feature (e.g., eye or nose) to be viewed, but it was also small enough that the eyes and mouth were not visible when fixating the nose. By contrast, when participants fixated the nose in the 8° condition, the mouth and eyes could be simultaneously viewed. Analysis of fixations strategies showed that the differences reported by Blais et al. (2008) were abolished in the restrictive 2° and 5° conditions with both populations of participants predominantly directing their fixations to the eye region. However, in the 8° condition when the eyes were visible when a nose region fixation was made, the EA participants reverted to their preferred central landing position. The authors concluded that the facial information, and most likely the cognitive mechanisms required to accurately individuate conspecifics are invariant, but the strategies used to extract this information are likely to be modulated by social experience.

Whilst the purpose of the Caldara et al. (2010) study was to elucidate *how* EA participants were able to fixate the nose and still accurately recognize faces, the function of the current study is to better understand *why* these differential strategies exist between Eastern and Western populations. One possible explanation is that these divergent strategies are driven by simple social norms. In the East it is can be considered rude to look a person in the eyes during social interaction, while Westerners consider it impolite to *not* engage eye contact during communication (Argyle and Cook, 1976). If such an assertion is correct, it is likely that the contrasting eye movement strategies reported for conspecifics' faces will not be found when participants view unfamiliar visually homogeneous categories of stimuli. However, if the differences observed in previous studies represent a more fundamental biological disposition when exploring visual objects, we can expect to observe differences across all categories. Additionally, if the central fixation strategy shown by Easterners is related to *holistic* face processing, this eye movement strategy should not be deployed for sheep faces or *greebles*, as holistic face processing is – *by definition* – not recruited when processing non-human faces or objects. Thus, in the present study the eye movements of WC and EA adult participants were recorded whilst viewing human faces, non-human faces (sheep) and an artificial category of stimuli, *greebles* (Gauthier and Tarr, 1997).

MATERIALS AND METHODS
PARTICIPANTS
In total, 21 Western Caucasian adults (15 female; age range = 19–33 years) and 21 East Asian adults (10 female; age range = 20–27) participated in the experiment. The WC adults were psychology students from the Department of Psychology at the University of Glasgow. The EA adults (20 Chinese, 1 Japanese) were all newly arrived students at the University of Glasgow. They were recruited

through advertisements in the university library and had been in the UK for approximately 1–2 weeks at the time of testing. All participants had normal or corrected to normal vision and were reimbursed for their time at the rate of £6 per hour. Each participant gave written informed consent and the protocol was approved by the faculty ethics committee.

MATERIALS

Face stimuli were obtained from the KDEF (Lundqvist et al., 1998) and AFID (Bang et al., 2001) databases and consisted of 24 East Asian and 24 Western Caucasian identities holding neutral facial expressions and contained equal numbers of males and females. The images were 390 × 382 pixels in size, subtending 15.6° of visual angle horizontally and 15.3° of visual angle vertically, which represents the size of a real face (approximately 19 cm in height). All images were cropped around the face to remove clothing and were devoid of distinctive features (scarf, jewelry, facial hair etc.). Sheep stimuli were obtained from Prof. Mike Kendrick (University of Cambridge) and consisted of 48 unique identities. The images were 420 × 382 pixels in size subtending 16.8° of visual angle horizontally and 15.3° of visual angle vertically. *Greebles* were obtained courtesy of Prof. Michael J. Tarr (Brown University, http://www.tarrlab.org/) and consisted of 48 unique identities. The images were 420 × 382 pixels in size subtending 16.8° of visual angle horizontally and 15.3° of visual angle vertically. All images were mounted on a white background.

All images were viewed at a distance of 70 cm. This reflects a natural distance during human face-to-face interaction (Hall, 1966) and has been successfully used in previous studies (Blais et al., 2008; Caldara et al., 2010). Human and sheep faces were aligned on the eye and mouth positions using psychomorph software (Tiddeman et al., 2001). Luminance was normalized for all images and they were presented on a 800 × 600 pixel gray background displayed on a Dell P1130 19″ CRT monitor with a refresh rate of 170 Hz. Presentation of stimuli was controlled by code custom written in MatLab (The MathWorks, MA, USA). It is important to note that visually homogeneous categories were used as they permit both the normalization of visual information complexity *between* categories and the averaging of eye movement strategies for individual exemplars *within* categories.

EYE TRACKING

Eye movements were recorded at a sampling rate of 1000 Hz with the SR Research Desktop-Mount EyeLink 2K eyetracker (with a chin/forehead-rest), which has an average gaze position error of about 0.25°, a spatial resolution of 0.01° and a linear output over the range of the monitor used. Only the dominant eye of each participant was tracked although viewing was binocular. The experiment was implemented in Matlab (R2006a), using the Psychophysics (PTB-3) and EyeLink Toolbox extensions (Brainard, 1997; Cornelissen et al., 2002). Calibrations of eye fixations were conducted at the beginning of the experiment using a nine-point fixation procedure as implemented in the EyeLink API (see EyeLink Manual) and using Matlab software. Calibration was validated with the EyeLink software and repeated when necessary until the optimal calibration criterion was reached. At the beginning of each trial, participants were instructed to fixate a dot at the center of

the screen to perform a drift correction. If the drift correction was more than 1°, a new calibration was launched to insure optimal recording quality.

PROCEDURE

All 42 participants completed human face, sheep face and *greeble* stimulus conditions. The order in which conditions were completed was counterbalanced across participants. All participants began each stimulus condition with a training session, which comprised four examples of the images that would be displayed in that condition. Importantly, these images were sourced from the original databases from which the final stimulus sets were taken, but they did not form part of the final sets and were not displayed again subsequently. The purpose of the training session was simply to familiarize the participants with the stimuli.

Participants were informed that they would be presented with a series of images to learn and subsequently recognize. They were also told that they would be given two recognition blocks per stimulus condition. In each block, participants were instructed to learn 12 images. After a 30-s pause, a series of 24 images (12 targets from the learning phase plus 12 foils) were presented and participants were asked to indicate as quickly and as accurately as possible whether each stimulus was a target or foil by pressing designated keys (a, l) on the keyboard with the index fingers of their left and right hands. For the human face condition, faces of the two races were presented in separate blocks, with the order of presentation for same- and other-race blocks being counterbalanced across participants. Response buttons were counterbalanced across participants.

In contrast to previous studies (Blais et al., 2008; Caldara et al., 2010) where the faces displayed during learning and recognition phases contained different emotional expressions, participants viewed the same pictures during learning and recognition phases. This was simply because we had only picture per identity for sheep faces so could not change images across phases in this condition. We elected to not change images across phases in the other stimulus conditions to maintain consistency and to allow direct comparison of results across stimulus conditions.

Each test trial started with the presentation of a central fixation cross. Then four crosses were presented, one in the middle of each of the four quadrants of the computer screen. These crosses allowed the experimenter to check that the calibration was still accurate. In that way, calibration was validated between each test trial. Following this check, a final central fixation cross that served to monitor drift correction was displayed. Following these checks, a stimulus was then presented on the computer screen. All stimuli were presented for 5 s duration in the learning phase and until the participant made a key press response in the recognition phase. To prevent anticipatory strategies, images were randomly presented at different locations of the computer screen. Each stimulus was subsequently followed by the six fixation crosses, as described above, which preceded the next stimulus.

DATA ANALYSES

Only correct trials were analyzed. Fixation distribution maps were computed individually for WC and EA participants, for each stimulus condition and separately for the learning and recognition phases. The fixation maps were computed by summing, across

all (correct) trials, the fixation location coordinates (x, y) across time. Since more than one pixel is processed during a fixation, we smoothed the resulting fixation distributions with a Gaussian kernel with a sigma of 15 pixels. Then, the fixation maps of all the participants belonging to the same cultural group were summed together separately for each face condition, resulting in group fixation maps.

We Z-scored the resulting group fixation maps by assuming identical WC and EA eye movement distributions for a particular face race as the null hypothesis. Consequently, we pooled the fixation distributions of participants for both groups and used the mean and the standard deviation for each stimulus condition to separately normalize the data. Finally, to clearly reveal the difference of fixation patterns across participants of different cultures, we subtracted the group fixation maps of the EA participants from the WC participants and Z-scored the resulting distribution. To establish significance, we used a robust statistical approach correcting for multiple comparisons in the fixation map space, by applying a two-tailed *Pixel test* (Chauvin et al., 2005; $Z_{crit} > |4.38|$; $p < 0.05$) on the *differential* fixation maps. Finally, for each condition we individually extracted the average Z-score values for all participants within each region of interest showing significance in the differential fixation maps to estimate Cohen's d effect sizes (Cohen, 1988).

RESULTS

ACCURACY

A 2 (*Culture of Observer*: British or Chinese) × 3 (*Stimuli*: human faces, sheep faces, *greebles*) ANOVA was conducted on participant's recognition accuracy. The ANOVA yielded main effect of *Stimuli* [$F(2, 40) = 39.778$, $p < 0.001$, $\eta_p^2 = 0.499$] only. *Post hoc* analysis conducted on accuracy for stimulus categories showed that WC and EA participant's performance was comparable for all stimulus categories with both groups recognizing human faces most accurately, followed by *greebles* and sheep faces.

REACTION TIME

A 2 (*Culture of Observer*: British or Chinese) × 3 (*Stimuli*: human faces, sheep faces, *greebles*) ANOVA was conducted on participant's reaction time for correct responses. The ANOVA yielded main effects of *Culture of Observer* [$F(1, 40) = 37.809$, $p < 0.001$, $\eta_p^2 = 0.267$] and *Stimuli* [$F(2, 40) = 155.714$, $p < 0.001$, $\eta_p^2 = 0.395$]. *Post hoc* analysis revealed that regardless of culture, participants responded fastest for human faces followed by *greebles* and then sheep faces. EA participants also took longer to respond in all conditions (see **Figure 1**).

NUMBER OF FIXATIONS

A 2 (*Culture of Observer*: British or Chinese) × 3 (*Stimuli*: human faces, sheep faces, *greebles*) ANOVA was conducted on participant's number of fixations. The ANOVA yielded main effects of Stimuli [$F(2, 40) = 12.172$, $p < 0.001$, $\eta_p^2 = 0.092$] with participants from both groups making fewer fixations for human faces relative to sheep and *greebles* (see **Table 1**).

EYE MOVEMENTS

Similar to previous reports, when viewing human faces WC participants systematically fixated the eye and mouth regions during learning and the eye region during recognition. By contrast, the EA participants primarily fixated on the nose region during learning and recognition phases. The strategic group differences, as revealed by a two-tailed pixel test ($Z_{crit} > |4.64|$, $p < 0.05$), are clearly illustrated in the difference maps (see **Figure 2**). Areas fixated above chance are delimited by white borders and depict the relative fixation biases following map subtraction (WC–EA). Strategic differences for WC and EA faces were not explored as a previous study (Blais et al., 2008) showed that fixations patterns were not modulated by the race of the face stimuli.

Turning to the novel categories of stimuli, for sheep faces it is evident from **Figure 2** that the strategic differences found for human faces are also observed for sheep. WC participants cluster
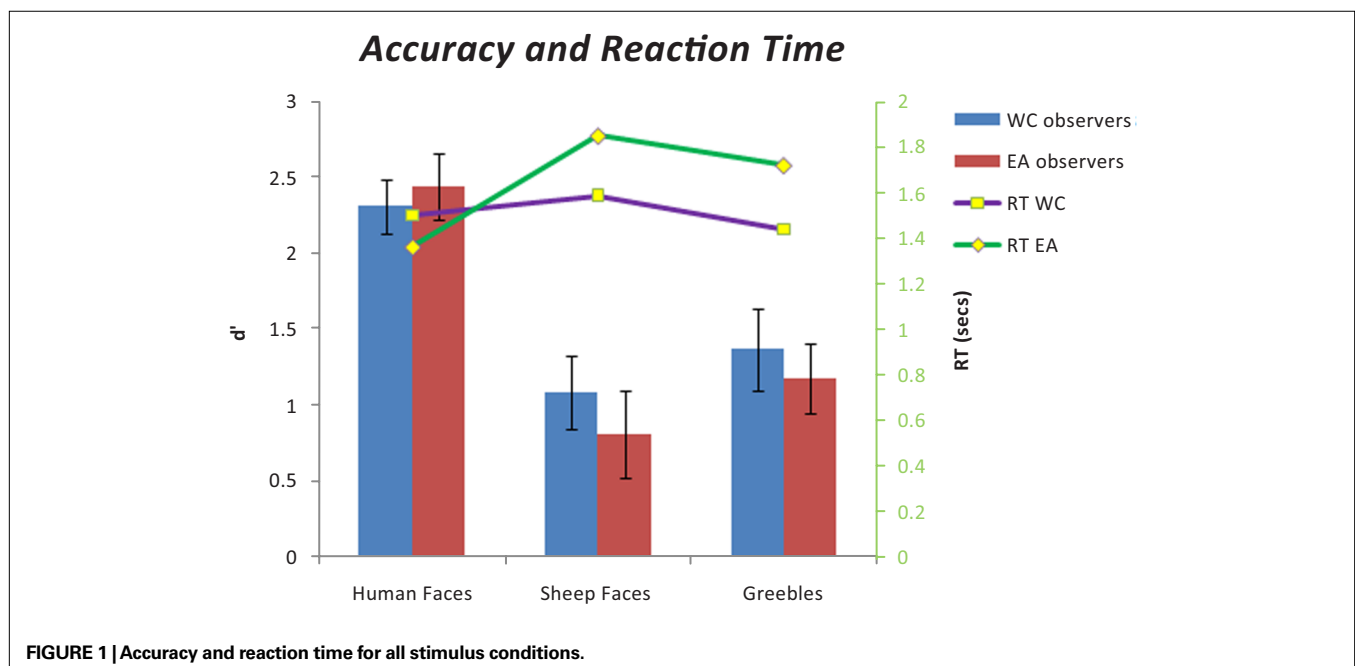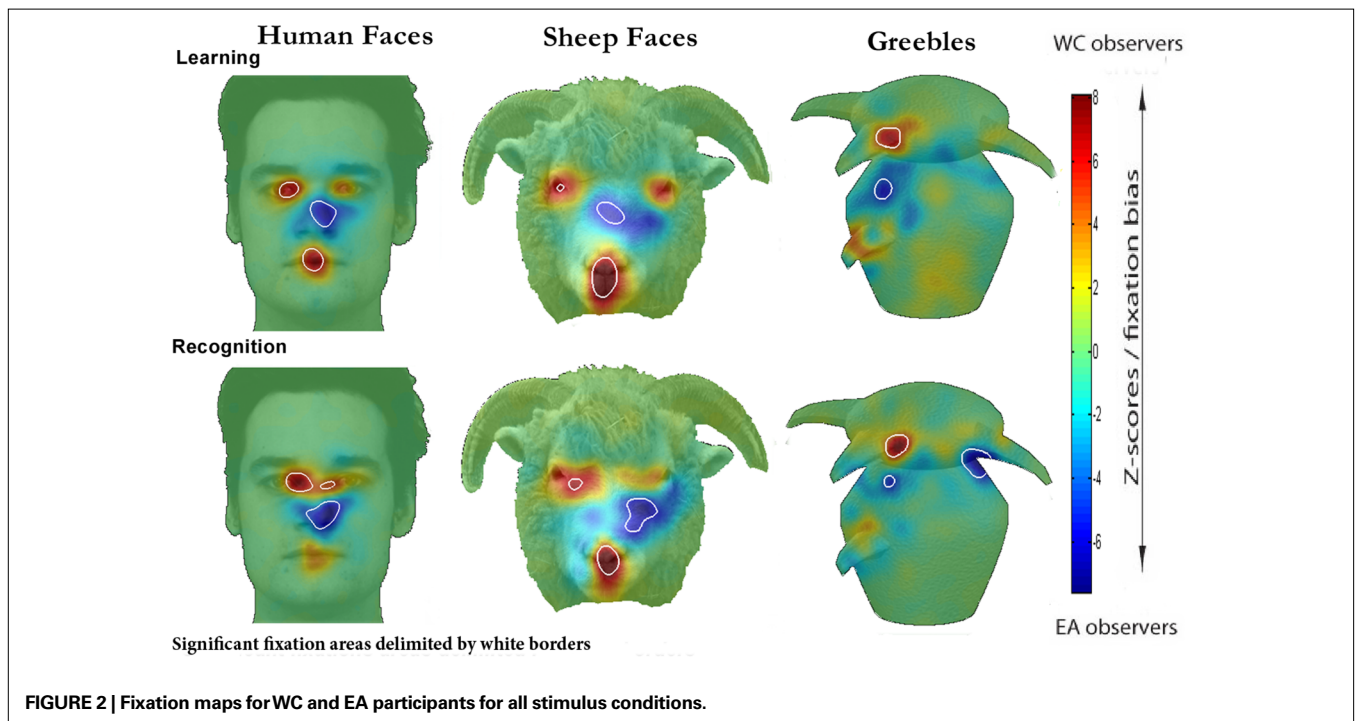

FIGURE 1 | Accuracy and reaction time for all stimulus conditions.

Table 1 | Number of fixations (standard deviations) for each all stimulus conditions.

Condition	Learning						Recognition					
Culture	Western			Eastern			Western			Eastern		
Stimulus	HF	SF	G	HF	SF	G	HF	SF	G	HF	SF	G
Fixations	13.80 (2.97)	14.37 (2.32)	12.68 (2.40)	12.04 (1.68)	13.25 (2.16)	12.58 (2.22)	2.92 (1.01)	4.48 (1.20)	3.90 (1.43)	2.76 (1.25)	5.00 (2.22)	4.21 (1.69)

FIGURE 2 | Fixation maps for WC and EA participants for all stimulus conditions.

their fixations around the eye and mouth regions, whereas EA participants fixate more centrally. Most surprisingly, strategic differences between groups are even found for *greebles*. EAs centered more fixations between the two prominent features (head and body limbs) whilst WCs landed their fixations directly on the central head limb.

In order to determine the magnitude of fixation biases, *Z*-scored fixation averages for each observer that landed within the significant regions were extracted. Effect sizes (*Cohen's d*) were then calculated from these values to reveal that fixation biases were large and robust across learning and recognition in all stimulus conditions (see **Table 2**).

DISCUSSION

It is clear from the fixation maps displayed in **Figure 2** that the strategic differences reported for face processing tasks in previous studies also extend to visually homogeneous categories with which participants had little or no prior familiarity. Unequivocally, any links between the eye movement strategies employed and *holistic* face processing are unsupported as both populations displayed the same pattern of fixations across all stimulus categories. The findings for

human faces replicates previous reports with WC participants primarily exploring the eye and mouth regions, whilst EA participants fixated centrally on the nose region (Blais et al., 2008; Caldara et al., 2010). Somewhat surprisingly these distinct strategies extended to sheep faces, with both groups of participants employing the same strategies used for human faces. Even more unexpectedly, clear and consistent differences in fixations were also found for *greebles*.

In the case of sheep faces, it could be argued that participants simply transferred their favored strategy for human faces and perhaps a similar result would be yielded with any category of non-human faces or face-like stimuli (e.g., cars in front-view, houses with window-eyes and door-mouth etc.). According to this interpretation, when participants are confronted with a stimulus that is novel but possesses a familiar configuration, individuals will deploy the strategy that they typically use for more familiar categories of stimuli (i.e., conspecifics' faces). For *greebles*, however, the result is more intriguing as the stimulus category was completely novel to all participants, yet group differences were still displayed. EA participant's fixations landed approximately in the center of the object's prominent features whereas WC participant's fixations actually landed on salient features. It has been argued by some authors

Table 2 | Cohen's *d* values for significantly fixated areas.

Area	Human faces				Sheep faces				Greebles			
	Eyes		Nose		Eyes		Nose		Head		Center	
Phase	Learn	Reco	Learn	Reco	Learn	Reco	Learn	Reco	Learn	Reco	Learn	Reco
d	1.05	1.96	1.34	1.87	0.88	1.04	0.80	0.83	0.78	1.38	0.92	1.22

(e.g., Kanwisher, 2000) that *greebles* are 'face-like' and therefore the same 'strategy transfer' argument proposed for sheep faces could also be applicable for *greebles*. However, this view is not one that we endorse as *greebles* do not possess eyes or any obvious mouth. It is our contention that in the absence of such attributes, which clearly define a face as such, it becomes obligatory to classify *greebles* as non-face objects. If this view is accepted, then it can be tentatively concluded that the differing strategies observed for human faces are not *solely* driven by eye gaze avoidance/engagement social norms as had been proposed. From the current data it is impossible to completely rule out the social norm explanation for human faces at the very least. However, Jack et al. (2009) found that when categorizing expressive faces by emotion, EA adults rely almost exclusively on the eye region when making judgments. This strongly supports the view that EA adults do not inherently avoid eye gaze and that the central fixation pattern reflects a cultural perceptual tuning in visual (homogenous) object identification.

If factors other than simple gaze avoidance are involved, it is of course necessary to delineate what these might be. The most obvious candidate is general differences in perceptual processing style as described earlier. It was noted by Blais et al. (2008) that the triangular scanpath displayed by WCs was indicative of an *analytical* processing style, while the central fixation strategy used by EAs resembles a *holistic* processing strategy. Along with the social norms hypothesis, Blais et al. (2008) advocated that the general differences observed in perception might extend to human faces and therefore account for the pattern of results reported. The eye movement data from the current study lend weight to this argument as participants from both cultural groups display strategies for all stimuli that are consistent with these processing styles when viewing visual objects (but see Rayner et al., 2007; Evans et al., 2009 for conflicting results from scene perception studies). However, it is important to concede that it remains unclear whether the pattern of results observed in adulthood are indeed shaped by cultural forces or whether genetic factors govern the differential eye movement behavior. It is essential to test different populations, such as adopted Chinese adults living in Western societies and also across developmental age groups to clarify when differences emerge and how they are shaped by the cultural environment.

Linking eye movements to cognitive processes must be done with great caution. The eye movement data reported in the current study cannot inform us about the cognitive processes underlying face or object recognition. Indeed, as in our previous study (Blais et al., 2008) the recognition accuracy and reaction times across populations were essentially identical, suggesting that the underlying processes used by both groups are most probably the same. Our contention that the *holistic* and *analytical* eye movement strategies do not relate to holistic and featural face processing is strongly supported by the fact that both populations' strategies extended to objects with which they had little or no prior familiarity. Holistic face processing is typically attributed to experience with own-race faces processed less holistically than other-race faces (Michel et al., 2006), yet we observed the same eye movement strategies for sheep and greebles (and other-race faces), indicating that natural eye movements cannot straightforwardly isolate the underlying cognitive processes. However, eye movements still remain an invaluable technique to understand how the visual system gathers information and to isolate diagnostic features used in information processing.

In conclusion, adults from different cultural backgrounds extract visual information in a fundamentally different manner. Whereas previous studies have shown differences for biologically relevant stimuli (conspecifics), the current study has extended these findings to biologically irrelevant (sheep) and non-biological stimuli (*greebles*). Interestingly, neither strategy appears to be more effective than the other, as both groups demonstrate comparable behavioral performance across recognition tasks with a range of stimuli. The differences reported are not trivial, but instead represent robust and consistent visual information extraction styles that further challenge notions of universality in perception.

ACKNOWLEDGMENTS

REFERENCES

Argyle, M., and Cook, M. (1976). Gaze and Mutual Gaze. Cambridge, England, Cambridge University Press.

Bang, S., Kim, D., and Choi, S. (2001). Asian Face Image Database PF01. Pohang, Intelligent Multimedia Lab, Pohang University of Science and Technology. http://nova.postech.ac.kr/.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., and Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE* 3, e3022. doi:10.1371/journal.pone.0003022.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436.

Caldara, R., and Abdi, H. (2006). Simulating the 'other-race' effect with autoassociative neural networks: further evidence in favor of the face-space model. *Perception* 35, 659–670.

Caldara, R., Schyns. P. G., Mayer, E., Smith, M., Gosselin, F., and Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *J. Cogn. Neurosci.* 17, 1652–1666.

Caldara, R., Zhou, X., and Miellet, S. (2010). Putting culture under the *spotlight* reveals universal information use for face recognition. *PLoS One.* 5, e9708. doi: 10.1371/journal.pone.0009708.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., and Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *J. Vis.* 5, 659–667.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Hillsdale, NJ, Lawrence Erlbaum Associates.

Cornelissen, F. W., Peters, E. M., and Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behav. Res. Methods Instrum. Comput.* 34, 613–617.

Davidoff, J., Fonteneau, E., and Fagot, J. (2008). Local and global processing: observations from a remote culture. *Cognition* 108, 702–709.

Evans, K., Rotello, C. M., Li, X., and Rayner, K. (2009). Scene perception and memory revealed by eye movements and receiver-operating characteristic analysis: does a cultural difference truly exist? *Q. J. Exp. Psychol.* 26, 279–285.

Gauthier, I., and Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Res.* 37, 1673–1682.

Goldstein, A. G. (1979a). Race-related variation of facial features: anthropometric data I. *Bull. Psychon. Soc.* 13, 187–190.

Goldstein, A. G. (1979b). Facial feature variation: anthropometric data II. *Bull. Psychon. Soc.* 13, 191–193.

Gosselin, F., and Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Cognition* 41, 2261–2271.

Groner, R., Walder, F., and Groner, M. (1984). Looking at faces: local and global aspects of scanpaths. In Theoretical and Applied Aspects of Eye Movements Research, A. G. Gale and F. Johnson, eds (Amsterdam, Elsevier), pp. 523–533.

Hall, E. T. (1966). *The Hidden Dimension.* New York: Doubleday.

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

Hole, G. J. (1994). Configural factors in the perception of unfamiliar faces. *Perception* 23, 65–74.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., and Caldara, R. (2009). Cultural confusions show facial expressions are not universal. *Curr. Biol.* 19, 1543–1548.

Kanwisher, N. (2000). Domain specificity in face perception. *Nat. Neurosci.* 3, 759–763.

Kuhn, G., and Tatler, B. W. (2005). Magic and fixation: now you don't see it, now you do. *Perception* 34, 1155–1161.

Le Grand, R., Mondloch, C. J., Maurer, D., and Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychol. Sci.* 15, 762–768.

Lundqvist, D., Flykt, A., and Öhman, A. (1998). The Karolinska Directed Emotional Faces – KDEF, CD ROM. Stockholm, Sweden, Department of Clinical Neuroscience, Psychology section, Karolinska Institute.

Masuda, T., and Nisbett, R. E. (2001). Attending holistically vs. analytically: comparing the context sensitivity of Japanese and Americans. *J. Pers. Soc. Psychol.* 81, 922–934.

Michel, C., Rossion, B., Han, J., Chung, C. S., and Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychol. Sci.* 17, 608–615.

Miyamoto, Y., Nisbett, R. E., and Masuda, T. (2006). Culture and the physical environment. Holistic versus analytic perceptual affordances. *Psychol. Sci.* 17, 113–119.

Navon, D. (1977). Forest before the trees. The precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383.

Nisbett, R. E., and Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *Trends Cogn. Sci.* 9, 467–473.

Norenzayan, A., Smith, E. E., Kim, B. J., and Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cogn. Sci.* 26, 653–684.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.

Rayner, K., Li, X., Williams, C. C., Cave, K. R., and Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Res.* 47, 2714–2726.

Tanaka, J. W., and Farah, M. J. (1993). Parts and wholes in face recognition. *Q. J. Exp. Psychol.* 46, 225–245.

Tiddeman, B., Burt, M., and Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Comput. Graph. Appl.* 21, 42–50.

Yarbus, A. L. (1965). Role of Eye Movements in the Visual Process. Moscow: Nauka.

Young, A. W., Hellawell, D., and Hay, D. C. (1987). Configurational information in face perception. *Perception* 16, 747–759.