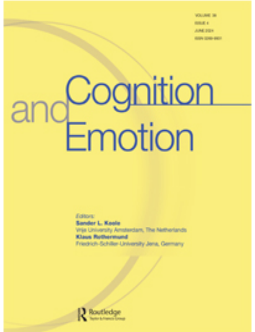

Cognition and Emotion

Routledge
Taylor & Francis Group

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pcem20

Registered report “Categorical perception of facial expressions of anger and disgust across cultures”

Xia Fang, Gerben A. van Kleef, Kerry Kawakami & Disa A. Sauter

To cite this article: Xia Fang, Gerben A. van Kleef, Kerry Kawakami & Disa A. Sauter (07 Jul 2024): Registered report “Categorical perception of facial expressions of anger and disgust across cultures”, Cognition and Emotion, DOI: 10.1080/02699931.2024.2370667

To link to this article: https://doi.org/10.1080/02699931.2024.2370667

COGNITION AND EMOTION
https://doi.org/10.1080/02699931.2024.2370667

REGISTERED REPORT

Registered report "Categorical perception of facial expressions of anger and disgust across cultures"

Xia Fang[a], Gerben A. van Kleef[b], Kerry Kawakami[c] and Disa A. Sauter[b]

[a]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, People's Republic of China; [b]Department of Social Psychology, University of Amsterdam, Amsterdam, the Netherlands; [c]Department of Social Psychology, York University, Toronto, Canada

ABSTRACT
Previous research has demonstrated that individuals from Western cultures exhibit categorical perception (CP) in their judgments of emotional faces. However, the extent to which this phenomenon characterises the judgments of facial expressions among East Asians remains relatively unexplored. Building upon recent findings showing that East Asians are more likely than Westerners to see a mixture of emotions in facial expressions of anger and disgust, the present research aimed to investigate whether East Asians also display CP for angry and disgusted faces. To address this question, participants from Canada and China were recruited to discriminate pairs of faces along the anger-disgust continuum. The results revealed the presence of CP in both cultural groups, as participants consistently exhibited higher accuracy and faster response latencies when discriminating between-category pairs of expressions compared to within-category pairs. Moreover, the magnitude of CP did not vary significantly across cultures. These findings provide novel evidence supporting the existence of CP for facial expressions in both East Asian and Western cultures, suggesting that CP is a perceptual phenomenon that transcends cultural boundaries. This research contributes to the growing literature on cross-cultural perceptions of facial expressions by deepening our understanding of how facial expressions are perceived categorically across cultures.

ARTICLE HISTORY
Received 12 October 2021
Revised 17 March 2024
Accepted 13 June 2024

KEYWORDS
Emotion; facial expressions; categorical perception; culture

It is well established that people across all cultures can identify emotions, including anger, disgust, fear, sadness, and happiness, from facial expressions. Recognising expressions entails mapping combinations of facial features onto specific emotion categories. To explain how this works, two accounts have been proposed (Etcoff & Magee, 1992; Fugate, 2013). One account holds that people perceive facial expressions as varying continuously along certain underlying dimensions, reflecting the incremental manifestation of different emotion-relevant features such as the activation of specific facial muscles. According to this view, membership in a particular emotion category (e.g. anger) is assigned by higher conceptual and linguistic systems, and perception is continuous rather than categorical. An alternative view is that people perceive facial expressions as belonging to discrete categories, because the perceptual mechanisms are specifically tuned to combinations of facial features that represent prototypical exemplars of emotion categories. Although evidence has accumulated showing that facial expressions of emotion are perceived categorically (Calder et al., 1996; Cong et al., 2019; De Gelder et al., 1997; Etcoff & Magee, 1992; Sauter et al., 2011; Young et al., 1997), this evidence is based primarily on studies of Western participants. To our knowledge, research has yet to examine whether Categorical Perception (CP) also characterises judgments of facial

CONTACT Xia Fang x.fang@zju.edu.cn Department of Psychology and Behavioral Sciences, Zhejiang University, 866 Yuhangtang Rd., Hangzhou, Zhejiang Province, People's Republic of China
Supplemental data for this article can be accessed online at https://doi.org/10.1080/02699931.2024.2370667.

expressions by East Asian participants. In the present study, our goal was to investigate whether East Asians perceive facial expressions categorically and if/how this process differs from Western participants.

Categorical perception

Categorisation is considered one of the most fundamental of human cognitive abilities. The ability to sort things into groups can be innate or learned via prototypes or exemplars that belong to that category or learned via boundaries that divide response regions (Ashby & Maddox, 2005; Medin & Schaffer, 1978; Nosofsky, 1986; Reed, 1972).

Here, we focus on CP, a phenomenon that accompanies the process of categorisation in some (but not all) domains of stimuli (Harnad, 1987). CP refers to the perception of strict boundaries when there is a gradual change in a variable along a continuum (Fugate, 2013). For example, continuous changes along the visible light spectrum are parsed into discrete colours. Continuous stimuli are thus perceived as belonging to distinct categories marked by a clear delineation, at which perception shifts from one category to the other (Harnad, 1987). A defining feature of CP is that when stimuli are separated by an equal physical distance, it is easier to discriminate between two stimuli that straddle a category boundary than when stimuli do not cross this boundary. For example, given the same wavelength difference, it is easier to distinguish between two colours from different categories (e.g. green and blue) than two colours from the same category (e.g. two different shades of green). CP has been found for colours (Bornstein & Korda, 1984), speech sounds (Liberman et al., 1957), and face identities (Angeli et al., 2008; Beale & Keil, 1995; Levin & Angelone, 2002). The present research focuses on CP of facial expressions of emotions (e.g. Fugate, 2013; Sauter, 2018).

Categorical perception of emotional facial expressions

CP of emotional facial expressions means that when viewing a full continuum between two prototypical emotional facial expressions (e.g. from happiness to sadness), individuals perceive discrete emotion categories rather than a gradual continuum of emotional change. Individuals thus perceive discrete shifts in emotion category membership.

The first investigation of CP of emotional expressions was conducted by Etcoff and Magee (1992). They surmised that if expressions are perceived categorically, then within a series of facial expression stimuli differing by equal physical increments between two different expressions, the probability of participants identifying a given expression as a particular emotion should not vary linearly across the series. Instead, perceptual judgments should change relatively abruptly at a boundary point. Specifically, pairs of faces should be discriminated more accurately when two stimuli straddle the category boundary (e.g. one predominantly angry face with some cues related to disgust and one predominantly disgusted face with some cues related to anger), compared to two stimuli that are identified as expressing the same emotion (e.g. two predominantly angry faces that both contain some cues to disgust but to different degrees).

To test their prediction, Etcoff and Magee (1992) converted prototypical photos of facial expressions into line-drawings and created a series of drawings representing equally interpolated steps between two different facial expressions posed by the same individual. These morphed faces were used in an ABX discrimination task, where participants were sequentially presented with faces A, B, and X and were asked to decide whether X was the same as A or B. Participants were more accurate in their judgments when A and B crossed a category boundary than when they were from the same category, despite A and B in both conditions being separated by an equal physical distance. These findings demonstrated that continua between emotional facial expressions were perceived categorically. Since then, the ABX discrimination task and its variants (e.g. the X-AB discrimination task, in which instead of presenting A, B, and X sequentially, A and B are presented simultaneously after X) have been used in the perception of facial expressions to study whether emotional expressions are perceived categorically or as varying continuously. Using this paradigm, Calder and others replicated and extended Efcoff and Magee's finding with photographic stimuli and a wider range of emotions (e.g. Calder et al., 1996; De Gelder et al., 1997; Young et al., 1997). Notably, researchers have found that CP of emotional facial expressions is also present in pre-verbal infants (Cong et al., 2019; Kotsoni et al., 2001; Lee et al., 2015).

Culture and emotion perception

The research on CP of facial expressions has largely been conducted in a Western context and only a limited number of studies have explored the impact of culture on this process. Wang et al. (2006) investigated how Chinese perceivers identified morphed emotional faces that were developed along the continua between expressions of happiness and surprise, expressions of surprise and fear, expressions of fear and sadness, expressions of sadness and disgust, expressions of disgust and anger, and expressions of anger and happiness by choosing one of the six corresponding emotion labels. They found that the frequency curves for correctly labelling each emotion across the morphed photographs showed six discrete peaks, indicating a relatively abrupt shift from one emotion to another around the boundary point of each emotion pair. This study, however, did not examine the defining feature of CP – that it is easier to discriminate between stimuli that straddle the category boundary than to discriminate between stimuli that do not cross the boundary. The results, however, do suggest that Chinese perceivers may also perceive emotional expressions categorically.

To date, only Sauter et al. (2011) have directly investigated CP of emotional facial expressions in a non-Western sample. In their study, CP related to anger and disgust expressions was examined among Germans and native speakers of Yucatec Maya, a language with no lexical labels that distinguish between these two emotions. The results demonstrated that both German and Yucatec participants were better in discriminating between-category pairs of faces (with one predominantly angry and one predominantly disgusted face) than within-category pairs of faces (with both showing predominantly anger or predominantly disgust). Notably, the magnitude of this effect did not differ across the language groups. These findings suggest that the perception of emotional expressions may be categorical across cultural groups, and that CP of emotional expressions is not driven by lexical labels.

A growing number of studies, however, suggest that people from different cultures vary in how well they recognise certain facial expressions (Beaupré & Hess, 2005; Jack et al., 2009; Matsumoto, 1992; Yik & Russell, 1999). Specifically, East Asians are known to make more misclassifications than Westerners in recognition tasks of facial expressions of anger, disgust, and fear. That is, East Asian perceivers more commonly misrecognise disgust as anger (and vice versa) and fear as surprise. It is noteworthy that these confusions also occur in Western perceivers, but to a lesser degree (Matsumoto, 1992). Anger and disgust (and also fear and surprise) are widely held to be distinct emotions, which are characterised by different patterns of facial, vocal, and autonomic physiological components and appraisals (Ekman & Cordaro, 2011; Jehna et al., 2011; Russell & Giner-Sorolla, 2013). However, some researchers have highlighted that these emotions are conceptually related and correspond to similar patterns of appraisals (Chapman & Anderson, 2013; Frijda et al., 1989; Smith & Ellsworth, 1985). For example, both anger and disgust are characterised by high arousal and negative valence, are other-focused, and both are considered to be morally-relevant emotions (Haidt, 2003). Disgust is also semantically close to anger (Russell & Fehr, 1994). It has even been shown that the term "disgust" is sometimes used by lay people as a synonym of "anger" (Nabi, 2002). In addition to conceptual similarities, facial expressions of anger and disgust also have perceptual (or morphological) similarities (Gagnon et al., 2010; Roy-Charland et al., 2014). For example, facial expressions of anger and disgust have in common the lowering of the inner part of the eyebrow. For anger, this change in appearance is produced by lowering the eyebrows, whereas for disgust it is produced by wrinkling the nose (Gagnon et al., 2010). By using electromyography (EMG) to measure facial muscle movements, Whitton et al. (2014) found that corrugator and levator activity seemed to rise and fall in tandem in response to manipulations of incidental anger and disgust. Because of these conceptual and perceptual similarities, researchers have found not only in adults (Jack et al., 2009) but also in children (Widen & Russell, 2003, 2008) that perceivers may confuse facial expressions of disgust with anger.

As to why East Asians recognise facial expressions of anger and disgust less accurately than Western Caucasians, recent research has found that it may be because East Asians are more likely than Westerners to see a mixture of emotions when viewing these facial expressions (Fang et al., 2018, 2019). For example, when observing a facial expression of anger, Chinese participants perceived primarily anger, but also – to a greater degree than Dutch participants – disgust. Likewise, when observing a disgusted face, Chinese participants were more likely than Dutch participants to perceive both anger and disgust (Fang et al., 2018). Moreover, this pattern of results was not limited to angry and disgusted faces based on Western prototypes (prototypes defined in the FACS manual), but also included angry and disgusted faces based on Chinese prototypes (facial expressions produced by Chinese actors posing facial expressions that would be best understood by friends from their own culture; Fang et al., 2019).

Together these findings raise a number of possibilities regarding East Asians' perceptions of facial expressions of anger and disgust. In the present research, we sought to explore three competing hypotheses: East Asians do not perceive angry and disgusted faces categorically; rather, they perceive angry and disgusted faces as varying continuously along certain dimensions (H1); East Asians exhibit CP for angry and disgusted faces, but the magnitude of the CP effect is smaller than that for Westerners (H2); East Asians exhibit CP for angry and disgusted faces, and the magnitude of the CP effect does not differ from that for Westerners (H3). We believe that a focus on anger and disgust is interesting and fitting in light of our goal to understand how culture shapes categorical emotion perception. That said, we acknowledge that other emotion pairs (e.g. fear-surprise) could also be investigated. Given the current project's focus on anger and disgust, any conclusions emerging from our data might be limited to the specific case of anger versus disgust. More research will be needed to empirically establish generalizability to other emotion pairs, although we submit that such generalizability would seem plausible on theoretical grounds.

The present research

The present research aimed to examine whether East Asians perceive facial expressions of anger and disgust categorically, and whether the size of the CP effect for East Asians was different from that for Westerners. Judgments of facial expressions on both Asian and White actors were made by participants from China and Canada. Samples from these two countries are commonly used to investigate East–West differences (e.g. Beaupré & Hess, 2005; Grossmann et al., 2016).

Following the standard approach in investigations of CP of emotional expressions, participants were instructed to complete two separate tasks. In the first task, participants discriminated between pairs of facial expressions that are morphs containing differing proportions of anger and disgust. The purpose of this discrimination task was to test the defining feature of CP by comparing the ability of participants to discriminate between faces belonging to the same category (Within-Category) and faces belonging to different categories (Between-Category) that vary by the same physical distance. Afterwards, participants identified emotions on these faces with a two-alternative forced identification task. The purpose of this identification task was to establish the category boundary on the morph continuum, where expressions to one side of the boundary were more likely to be perceived as anger and expressions on the other side of the boundary were more likely to be perceived as disgust.

The criterion for testing CP was whether better discrimination occurs for Between-Category pairs than Within-Category pairs. Better discrimination can be operationalised as higher accuracy or faster response latencies. In the present study, we examined both criteria. We would thus perform a 2 (Expresser Culture: Asian vs. White) × 2 (Trial Type: Between-Category vs. Within-Category) × 2 (Perceiver Culture: Chinese vs. Canadian) mixed-design ANOVA on accuracy and response latencies separately, with Expresser Culture and Trial Type as within-subject factors and Perceiver Culture as a between-subjects factor. If H1 – that Chinese participants perceive angry and disgusted faces as varying continuously along certain dimensions – was correct, then Chinese participants would show no difference in accuracy and response latencies between Between-Category and Within-Category pairs. Specifically, a significant interaction between Perceiver Culture and Trial Type would be expected, with the effect of Trial Type significant for Canadian but not Chinese participants. If H2 – that Chinese participants exhibit weaker CP effects for angry and disgusted faces than Canadian participants – was correct, then Chinese participants would identify Between-Category pairs more accurately and faster than Within-Category pairs, and Chinese participants would differ less than Canadian participants between the two trial types. Specifically, a significant interaction between Perceiver Culture and Trial Type would be expected, with a significant effect of Trial Type for both groups. If H3 – that Chinese and Canadian participants show a similar CP effect for angry and disgusted faces – was correct, then both groups would identify Between-Category pairs more accurately and faster than Within-Category pairs, and the two groups would differ similarly between the two trial types. Specifically, only a main effect of Trial Type would be expected, and the interaction between Perceiver Culture and Trial Type would be nonsignificant.

Method

Participants

To provide an estimate of the sample size, we performed a power analysis for Bayesian hypothesis testing because it is usually more conservative than null hypothesis significance testing (NHST). Using the BayesFactor package (Morey & Rouder, 2018) in R, we set the true effect at Cohen's $d = 0.43$ (the smallest effect size of the CP effect found in Sauter et al., 2011) with 100,000 simulations and Bayes factor = 6 in estimating an appropriate sample size. This analysis revealed that 150 participants from each cultural group were needed to achieve 80% power.[1] We thus intended to recruit 150 Chinese participants from a University in China and 150 European Canadian participants from a University in Canada for the experiment. The final sample consisted of 150 Chinese participants ($M_{age} = 20.40$, $SD = 2.11$; 50 men and 100 women) and 152 European Canadian participants ($M_{age} = 19.46$, $SD = 2.14$; 51 men and 101 women). Because participants in both cultures were university students, they were of similar education level and socioeconomic status. Participants received either partial course credit or money for their participation. The study was approved by the ethics committees of York University, Canada (ethics number: e2018-028) and Zhejiang University, China (ethics number: [2021]043).

Facial expression stimuli

One Asian female and one Asian male depicting facial expressions of anger and disgust were selected from the Taiwan Facial Expression Image Database (TFEID; Chen & Yen, 2007), and one White female and one White male depicting the same emotions were selected from the Amsterdam Dynamic Facial Expression Set (ADFES; Van der Schalk et al., 2011). Facial expressions in both databases were produced using instructions from the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002). Although each emotion can be expressed through a variety of combinations of muscle movements, the two expression databases we purposefully selected are generally consistent in terms of the muscle movements used for the angry and disgusted expressions. In particular, both databases contain brow lowerer (Action Unit 4 [AU4]), chin raiser (AU17) and lip presser (AU24) for the angry expressions, and nose wrinkler (AU9) and upper lip raiser (AU10) for the disgusted expressions. We FACS coded all of the stimuli, and established empirically that the crucial AUs of each emotional expression are similar across cultures and actors. See Table S1 for the FACS-coded AUs of all facial stimuli.

The ADFES consists of dynamic expressions changing from a neutral expression to a specific emotional state. To ensure the intensity of expression was similar across the White and Asian stimuli, we conducted a pilot study with 20 White Dutch ($M_{age} = 23.71$ years; 6 men, 14 women) and 20 Asian Chinese ($M_{age} = 27.60$ years; 12 men, 8 women) participants to select the frame of each White stimulus that best matched the intensity of the corresponding Asian stimulus. We extracted 150 frames from each original clip of the White actors (ranging from 6 to 6.5 s) to form new stimulus sequences. The number of frames was computed by multiplying 24 fps (the common frame rates used in films) by 6.25 (the mean of 6 and 6.5). A photograph of one of the Asian facial expressions was presented on the left side of the screen, while the corresponding sequential White facial expressions were presented on the right side of the screen. Participants were asked to drag the slider bar underneath the clip to choose the frame that was most similar in terms of intensity to the Asian stimulus. Each comparison between Asian and White stimuli included two trials with different initial positions of the slider bar, one starting from the first frame of the clip and the other starting from the last frame of the clip. Each participant completed a total of 16 trials of stimulus evaluation (4 actor pairs $\times$ 2 emotions [anger, disgust] $\times$ 2 initial positions of the slider bar).[2] An independent t-test was used to compare Chinese and Dutch participants' selected frames for each pair of stimuli. No significant differences were found between the two groups (Table S2). Therefore, the average frame across all participants was used to select the final stimuli for each White actor – resulting in sets of Asian and White facial expressions matched in terms of both AUs and perceived intensity. See Table S1 for the coded AU intensities of all facial stimuli.

We used a digital programme (Fantamorph; http://www.fantamorph.com) to morph the expressions of each actor to create continua of expressions with a mix of anger and disgust in different proportions. Following the most common approach used in research on CP of facial expressions (e.g. Calder et al., 1996; Cong et al., 2019; De Gelder et al., 1997; Etcoff &

Figure 1. Range of morphs on the anger-disgust continuum and the emotion pairs used in the discrimination task.

Magee, 1992), expressions were selected at 10% intervals in this successive variation of each actor's angry expression morphing into a disgusted expression. This generated 11 expressions for each actor, including the prototypes of the angry and disgusted expressions (0 and 10), and 9 morphs in between (morphs 1–9). See Figure 1 for an example.

Procedure

The experiments were programmed using a custom-written Psychopy programme (Psychophysics software in Python; Peirce, 2007). Each participant first completed the discrimination task and then proceeded to the identification task. After completing both the discrimination and identification tasks, participants were asked to provide basic demographic information (e.g. sex, age, race). The instructions for Chinese participants in both tasks were translated from English into Chinese by means of the standard translation/back-translation procedure.

Discrimination task

In accordance with earlier experiments (Calder et al., 1996), the stimulus set in the discrimination task was comprised of nine morphed facial expressions (morphs 1–9). Exclusion of the prototypical faces avoided participants using the small variation in picture quality between the prototypical faces and the morphed faces to discriminate between pairs of faces that include either of the prototypical faces. Because in sequential ABX discrimination tasks participants were required to hold representations of A and B in memory in order to decide whether face X matches face A or B, the results might reflect short-term memory performance rather than a perceptual phenomenon (see Calder et al., 1996). To rule out this possibility in the present study, we utilised an X-AB discrimination task (Cong et al., 2019; Sauter

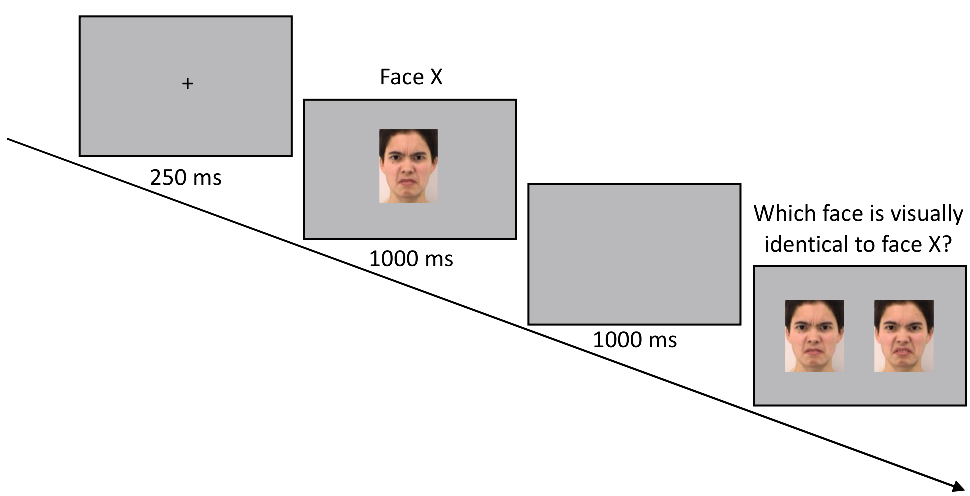

Figure 2. The procedure for the discrimination task.

et al., 2011) (see Figure 2). Specifically, in this task, each trial began with a 250 ms fixation followed by the first stimulus, face X, presented in the centre of the screen for 1000 ms. After a blank screen for 1000 ms (in order to allow the sensory memory of face X to fade), the target and distractor stimuli, faces A and B, appeared *side-by-side* until participants indicated which of the two stimuli, faces A or B, was visually identical to face X. Faces A and B differed by one step (10%) along the continuum, resulting in a total of 32 face pairs (8 pairs per actor × 4 actors).[3] All face pairs were presented eight times (twice in each of the four orders: A-AB, A-BA, B-AB, B-BA), resulting in a total of 256 trials. All pictures were subtended 8° × 12° of visual angle of participants seated at an average distance of 60 cm from the screen. Participants were instructed to respond as quickly and accurately as possible. No feedback was provided. The experiment started with six practice trials, using morphs of two actors from the Radboud Faces Database and the Taiwan Corpora of Chinese Emotions that were not be used in the experimental trials. The order of all experimental trials was randomised.

Identification task

After completing the X-AB discrimination task, participants were asked to identify the expression on each face (see Figure 3). In this task, each trial began with a central fixation for 250 ms, followed by one picture and two response alternatives below the image, "anger" and "disgust." The left or right position of the response alternatives were counterbalanced between participants. For Chinese participants, we provided the corresponding translations, "气愤(qifen/anger)" and "厌恶(yanwu/disgust)," which have been used in other cross-cultural studies of these emotions (Fang et al., 2018, 2019). Participants were asked to indicate which of the two emotions was expressed. Participants were shown the nine morphed faces from the X-AB discrimination task plus the two prototypical faces. Each face appeared once in each block and participants completed 4 blocks of 44 trials each for a total of 176 trials. The order of stimuli was random for each participant.

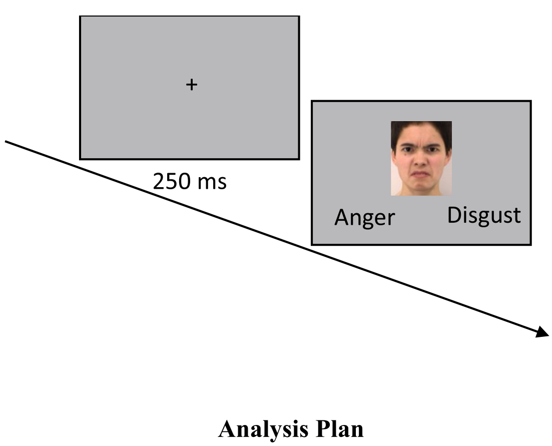

Figure 3. The procedure for the identification task.

Analysis plan

Analysis plan: identification task

The aim of the analysis of the identification data was to establish the point on the morph continuum where expressions on one side of the boundary point were more likely to be perceived as anger and expressions on the other side of the boundary were more likely to be perceived as disgust. In order to establish the boundary between anger and disgust, we first calculated the percentage of trials that were judged as "anger" for each picture and for each participant. The boundary was set at the point at which the proportion of trials that was judged as expressing anger was higher than 50% for morph N, while for morph $N + 1$, the proportion of trials that was judged as expressing anger was lower than 50%.

Analysis plan: discrimination task

We performed a 2 (Expresser Culture: White vs. Asian) × 2 (Trial Type: Between-Category vs. Within-Category) × 2 (Perceiver Culture: Canadian vs. Chinese) mixed-design ANOVA on accuracy and response latencies separately, with Expresser Culture and Trial Type as within-subject factors and Perceiver Culture as a between-subjects factor. Following the analytic strategy of Sauter et al. (2011), we conducted analysis on one Between-Category (faces A and B in a pair with different predominant emotions as established in the identification task) and other Within-Category pairs (faces A and B in a pair with the same predominant emotion as established in the identification task). For example, if the category boundary in the identification task was between morphs 4 and 5, then the Between-Category pair would be morphs 4 and 5 (pair 4) and the Within-Category pairs would be all other pairs including morphs 1 and 2 (pair 1), morphs 2 and 3 (pair 2), morphs 3 and 4 (pair 3), morphs 5 and 6 (pair 5), morphs 6 and 7 (pair 6), morphs 7 and 8 (pair 7), morphs 8 and 9 (pair 8). Because the category boundaries might be different for each participant, we calculated the category boundaries for each participant individually, rather than averaging the category boundaries of all participants. For example, the Between-Category pair for one participant might be morphs 4 and 5, while the Between-Category pair for another participant might be morphs 3 and 4. If the current X-AB task was valid, the main effect of Trial Type would be significant for Canadian participants.

The first competing hypothesis that Easterners do not perceive angry and disgusted expressions categorically (H1) would be supported by a significant two-way interaction between Perceiver Culture and Trial Type. In this scenario, simple effects analyses were expected to show that Canadian, but not Chinese, participants showed better accuracy and faster responding for Between-Category compared to Within-Category pairs. To provide evidence of an absence of CP in Chinese participants, a Bayesian *t*-test was conducted for the effect of Trial Type on Chinese participants' performance on accuracy and response latencies separately. The Bayes Factor (BF) for a null effect of Trial Type on accuracy and response latencies by Chinese participants should be larger than 3 to allow for a conclusion that no CP was present in Chinese participants (Wagenmakers et al., 2011).

The second competing hypothesis was that the size of the CP for East Asian perceivers of angry and disgusted expressions was *smaller* than for Western perceivers (H2). This hypothesis would be supported by a significant two-way interaction between Perceiver Culture and Trial Type. Although both Chinese and Canadian participants were expected to demonstrate significantly better accuracy and faster response latencies for Between-Category compared to Within-Category pairs, an independent *t*-test would indicate that the difference scores related to accuracy and response latencies of Between-Category and Within-Category pairs for Chinese participants were smaller than for Canadian participants.

The third competing hypothesis was that the magnitude of CP for East Asian perceivers of angry and disgusted expressions does not differ from Western perceivers (H3). This hypothesis would be supported by a main effect of Trial Type, not qualified by Perceiver Culture. Results would be expected to show that both groups of participants were faster and more accurate in identifying Between-Category pairs than Within-Category pairs. To provide further evidence that the magnitude of CP did not vary across cultures, a Bayesian mixed-design ANOVA was conducted to compare the main effects model (including main effects of Trial Type and Perceiver Culture) against the Null model and to compare the interaction model (including both main effects as well as their interactions) against the Null model, respectively. The BF for the main effects model (against the Null model) divided by the BF for the interaction model (against the Null model) should be larger than 3 to allow for a conclusion that no interaction between Perceiver Culture and Trial Type was present (Wagenmakers et al., 2011).

Results

All data and analysis code are available at https://osf.io/ftnb5/?view_only=da66dda5dd9140fdbfba1d31d6def5e9. Before proceeding with the primary analyses, we provide an overview of the identification performance among Canadian and Chinese participants for all actors. For each cultural group of perceivers, we computed the proportion of trials in which the stimuli were categorised as "anger" (see panel A in Figure 4). Similarly, to provide a summary of the discrimination performance across both cultures of perceivers for all actors, we calculated the proportion of trials in which the stimuli were correctly categorised for each pair, along with the corresponding response latencies (see panels B and C in Figure 4).

Pre-registered analysis

Data prescreening

According to the analysis plan, we computed the category boundaries for each participant with respect to each continuum. In the identification task, if any participant's recognition rate for the prototypical angry expression (morph 0) or the prototypical disgusted expression (morph 10) for a specific actor was below 50%, their discrimination data for that particular actor was excluded. This indicates that the participant did not recognise the prototypical angry expression as anger or the prototypical disgusted expression as disgust. Additionally, if any participant's judgment of a specific morph for a particular actor fell exactly at 50%, we were unable to define the "Between-Category" and "Within-Category" pairs for that actor.

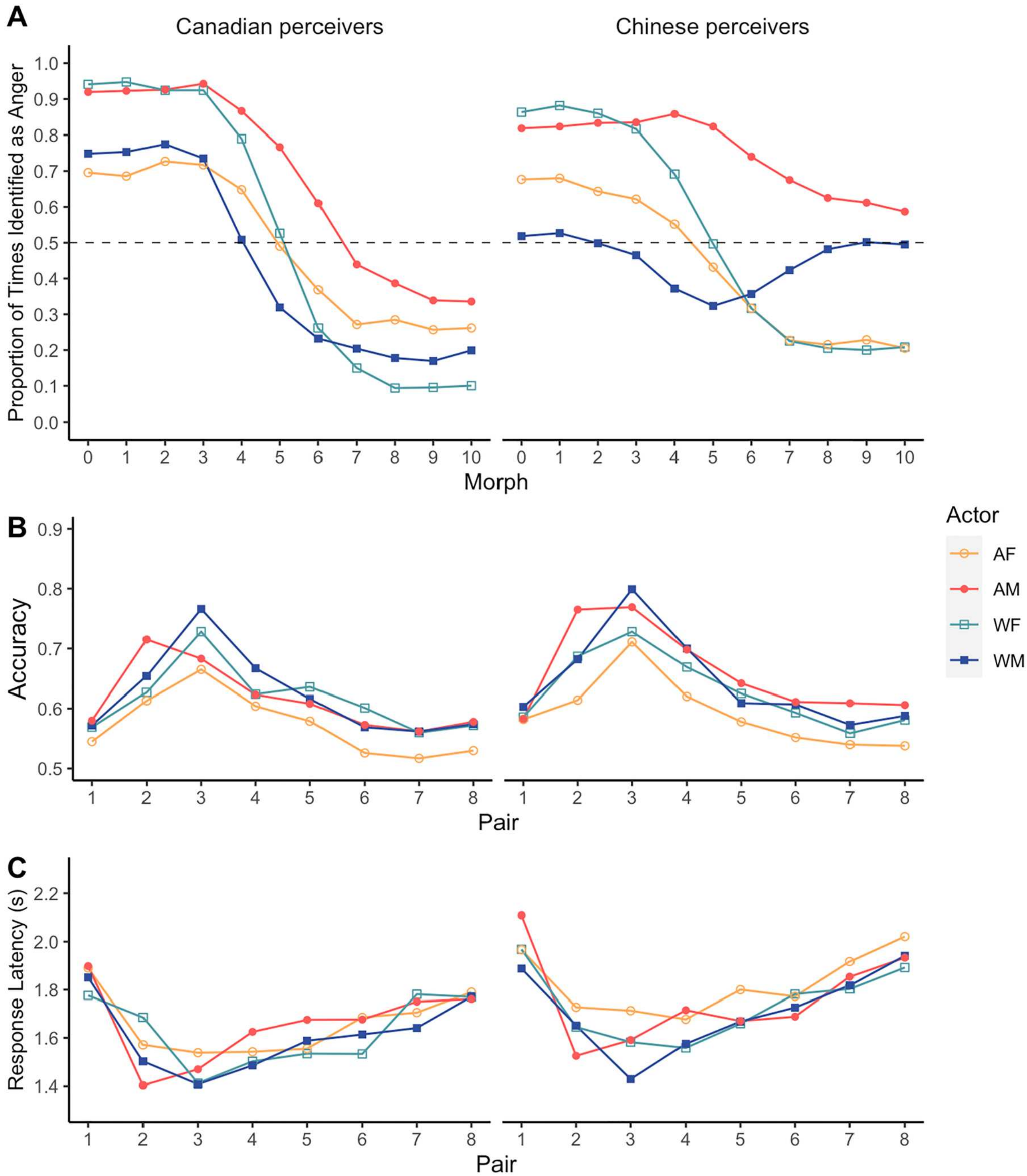

Figure 4. Proportion of face identification as anger (Panel A), accuracy (Panel B), and response latency (Panel C) in discriminating between face pairs.

Note: AF indicates Asian female, AM indicates Asian male, WF indicates White female, and WM indicates White male.

Therefore, the participant's discrimination data for that actor was excluded. In the discrimination task, if a participant's overall accuracy on the emotion discrimination task was more than two standard deviations below the mean or his/her overall response latency was more than two standard deviations above the mean, the participant's data was excluded (Cong et al., 2019).

Table 1. Expresser culture × trial type × perceiver culture mixed-design analysis of variance for discrimination accuracy and response latency based on individual category boundaries.

	Accuracy				Response latency			
Effect	*F*	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2
Expresser Culture (E)	10.44	(1, 126)	.002	0.077	10.72	(1, 126)	.001	0.008
Trial Type (T)	12.70	(1, 126)	<.001	0.092	20.85	(1, 126)	<.001	0.142
Perceiver Culture (*P*)	2.32	(1, 126)	.130	0.018	<0.01	(1, 126)	.997	<0.001
E × T	0.27	(1, 126)	.607	0.002	0.78	(1, 126)	.378	0.006
E × *P*	0.24	(1, 126)	.622	0.002	<0.01	(1, 126)	.984	<0.001
T × *P*	0.09	(1, 126)	.770	0.001	2.39	(1, 126)	.125	0.019
E × T × *P*	0.18	(1, 126)	.672	0.001	0.01	(1, 126)	.906	<0.001

Because the preregistered analysis was a 2 (Expresser Culture: White vs. Asian) × 2 (Trial Type: Between-Category vs. Within-Category) × 2 (Perceiver Culture: Canadian vs. Chinese) mixed-design ANOVA, it was essential for each participant to have data for both conditions of the Expresser Culture and Trial Type factors in order to proceed with the analysis. Based on these criteria, we obtained data from a total of 75 Canadian participants and 53 Chinese participants, which were utilised for the present analysis.[4] A complete overview of effects can be found in Table 1.

Accuracy

Central to our hypotheses, the analysis revealed a significant main effect of Trial Type, $F(1, 126) = 12.70$, $p < .001$, $\eta_p^2 = .092$, 90% CI [0.026, 0.177], which was not qualified by the two-way interaction between Trial Type and Perceiver Culture, $F(1, 126) = 0.09$, $p = .770$, $\eta_p^2 = .001$, 90% CI [0.000, 0.024]. Specifically, participants were more accurate when judging between-category pairs of expressions ($M = 0.65$; $SD = 0.12$) than within-category pairs of expressions ($M = 0.61$; $SD = 0.07$). These findings support H3, indicating that both cultural groups of perceivers exhibited CP for angry and disgusted expressions, with no significant difference in the magnitude of CP between East Asian and Western perceivers (see Figure 5(A)). Furthermore, we observed a significant main effect of Expresser Culture, $F(1, 126) = 10.44$, $p = .002$, $\eta_p^2 = .077$, 90% CI [0.018, 0.158], with expressions by White actors being discriminated with higher accuracy ($M = 0.65$; $SD = 0.10$) compared to those by Asian actors ($M = 0.61$; $SD = 0.09$). Despite efforts made during stimulus development to match facial muscle movements and perceived intensity between

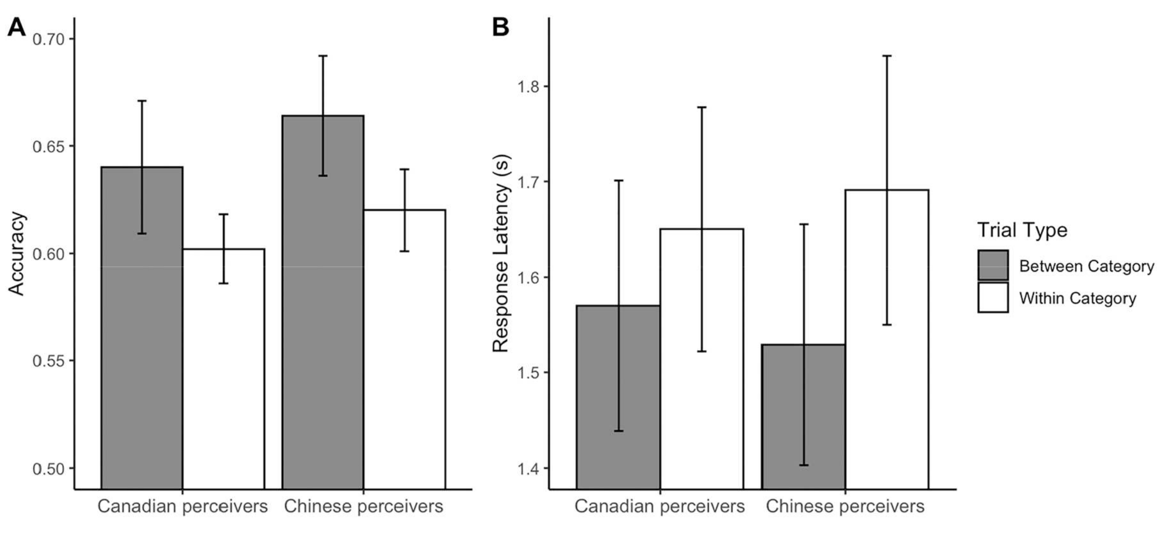

Figure 5. Accuracy (A) and response latency (B) of discrimination task as a function of trial type and perceiver culture based on individual category boundaries.

Note: Error bars indicate 95% confidence intervals.

White and Asian actors, these results suggest a possible higher similarity in the expressions of anger and disgust portrayed by Asian actors compared to White actors.[5]

To further investigate whether the magnitude of CP varied across cultures of perceivers, we conducted a Bayesian mixed-design ANOVA. The main effects model, which included the main effects of Trial Type and Perceiver Culture, was compared against the Null model. Additionally, the interaction model, including the main effects and their interactions, was compared against the Null model. The BF for the main effects model (against the Null model) divided by the BF for the interaction model (against the Null model) yielded a value of 5.417, providing moderate evidence that there was no significant interaction between Trial Type and Perceiver Culture.

Response latency

Consistent with the accuracy results, the main effect of Trial Type was significant, $F(1, 126) = 20.85$, $p < .001$, $\eta_p^2 = .142$, 90% CI [0.060, 0.235], and it was not qualified by the two-way interaction between Trial Type and Perceiver Culture, $F(1, 126) = 2.39$, $p = .125$, $\eta_p^2 = .019$, 90% CI [0.000, 0.074]. Specifically, participants were faster in judging between-category pairs of expressions ($M = 1.55$; $SD = 0.52$) than within-category pairs of expressions ($M = 1.67$; $SD = 0.54$; see Figure 5(B)). This finding indicates that the observed accuracy advantage for between-category trials cannot be attributed to a speed-accuracy trade-off. Furthermore, we observed a significant main effect of Expresser Culture, $F(1, 126) = 10.72$, $p = .001$, $\eta_p^2 = .008$, 90% CI [0.019, 0.161], with expressions by White actors being discriminated with faster response latencies ($M = 1.57$; $SD = 0.48$) compared to those by Asian actors ($M = 1.65$; $SD = 0.57$). Again, these results suggest a possible higher similarity in the expressions of anger and disgust portrayed by Asian actors compared to White actors.

To further investigate whether the magnitude of CP varied across cultures of perceivers, we conducted a Bayesian mixed-design ANOVA. The BF for the main effects model (against the Null model) divided by the BF for the interaction model (against the Null model) yielded a value of 1.781, providing anecdotal evidence that there was no significant interaction between Trial Type and Perceiver Culture.

Overall, the results revealed a similar pattern of CP for facial expressions across Canadian and Chinese participants. Specifically, participants from both cultural groups consistently demonstrated higher accuracy and faster response latencies when categorising between-category pairs of expressions compared to within-category pairs, highlighting a clear perceptual differentiation between these two expression types.

Exploratory analysis

Data prescreening

Consistent with the commonly employed analytical approach in previous research (e.g. Calder et al., 1996; Cong et al., 2019; Etcoff & Magee, 1992; Young et al., 1997), we calculated the category boundaries for each continuum separately for all Chinese and Canadian participants. As depicted in Figure 4(A), we were unable to identify category boundaries for the Asian male and White male models among the Chinese participants. Consequently, for the formal data analysis, we excluded the data pertaining to the Asian male and White male models. Likewise, in the discrimination task, participants were excluded if their overall accuracy on the emotion discrimination task deviated more than two standard deviations below the mean or if their overall response latency exceeded two standard deviations above the mean (Cong et al., 2019). Applying these criteria, we obtained data from a total of 143 Canadian participants and 139 Chinese participants, which were utilised for the 2 (Expresser Culture: White vs. Asian) × 2 (Trial Type: Between-Category vs. Within-Category) × 2 (Perceiver Culture: Canadian vs. Chinese) mixed-design ANOVA. A complete overview of effects can be found in Table 2.

Accuracy

Central to the hypotheses, the analysis revealed a significant main effect of Trial Type, $F(1, 280) = 24.40$, $p < .001$, $\eta_p^2 = .080$, 90% CI [0.036, 0.134], which was not qualified by the two-way interaction between Trial Type and Perceiver Culture, $F(1, 280) = 0.22$, $p = .638$, $\eta_p^2 = .001$, 90% CI [0.000, 0.015]. Specifically, participants were more accurate when judging between-category pairs of expressions ($M = 0.64$; $SD = 0.13$) than within-category pairs of expressions ($M = 0.60$; $SD = 0.07$). These findings support H3, indicating that both cultural groups of perceivers exhibited CP for angry and disgusted expressions, with no significant difference in the magnitude of CP between East Asian and Western perceivers (see Figure 6(A)). Furthermore, we observed a significant main effect

Table 2. Expresser culture × Trial type × Perceiver culture mixed-design analysis of variance for discrimination accuracy and response latency based on average category boundaries.

	Accuracy				Response latency			
Effect	*F*	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2
Expresser Culture (E)	31.17	(1, 280)	<.001	0.100	17.94	(1, 280)	<.001	0.060
Trial Type (T)	24.40	(1, 280)	<.001	0.080	41.36	(1, 280)	<.001	0.129
Perceiver Culture (*P*)	4.52	(1, 280)	.034	0.016	1.55	(1, 280)	.214	0.006
E × T	0.14	(1, 280)	.714	<0.001	0.15	(1, 280)	.704	0.001
E × *P*	0.38	(1, 280)	.536	0.001	2.29	(1, 280)	.131	0.008
T × *P*	0.22	(1, 280)	.638	0.001	1.38	(1, 280)	.242	0.005
E × T × *P*	1.25	(1, 280)	.264	0.004	1.30	(1, 280)	.256	0.005

of Expresser Culture, $F(1, 280) = 31.17$, $p < .001$, $\eta_p^2 = .100$, 90% CI [0.051, 0.158]. Specifically, expressions by White actors being discriminated with higher accuracy ($M = 0.64$; $SD = 0.10$) compared to those by Asian actors ($M = 0.60$; $SD = 0.10$), suggesting a possible higher similarity in the expressions of anger and disgust portrayed by Asian actors compared to White actors.

To further investigate whether the magnitude of CP varied across cultures of perceivers, we conducted a Bayesian mixed-design ANOVA. The BF for the main effects model (including the main effects of Trial Type and Perceiver Culture; against the Null model) divided by the BF for the interaction model (including the main effects and their interactions; against the Null model) yielded a value of 6.809, providing moderate evidence that there was no significant interaction between Trial Type and Perceiver Culture.

Response latency

Consistent with the accuracy results, the main effect of Trial Type was significant, $F(1, 280) = 41.36$, $p < .001$, $\eta_p^2 = .129$, 90% CI [0.073, 0.190], and it was not qualified by the two-way interaction between Trial Type and Perceiver Culture, $F(1, 280) = 1.38$, $p = .242$, $\eta_p^2 = .005$, 90% CI [0.000, 0.027]. Specifically, participants were faster in judging between-category pairs of expressions ($M = 1.53$; $SD = 0.54$) than within-category pairs of expressions ($M = 1.66$; $SD = 0.53$; see Figure 6 (B)), indicating that the observed accuracy advantage for between-category trials cannot be attributed to a speed-accuracy trade-off. Furthermore, we observed a significant main effect of Expresser Culture, $F(1, 280) = 17.94$, $p < .001$, $\eta_p^2 = .060$, 90% CI [0.023, 0.110]. Specifically, expressions by White actors being discriminated with faster response latencies ($M = 1.55$; $SD = 0.51$) compared to those by Asian actors ($M = 1.63$; $SD = 0.56$), suggesting a possible higher similarity in

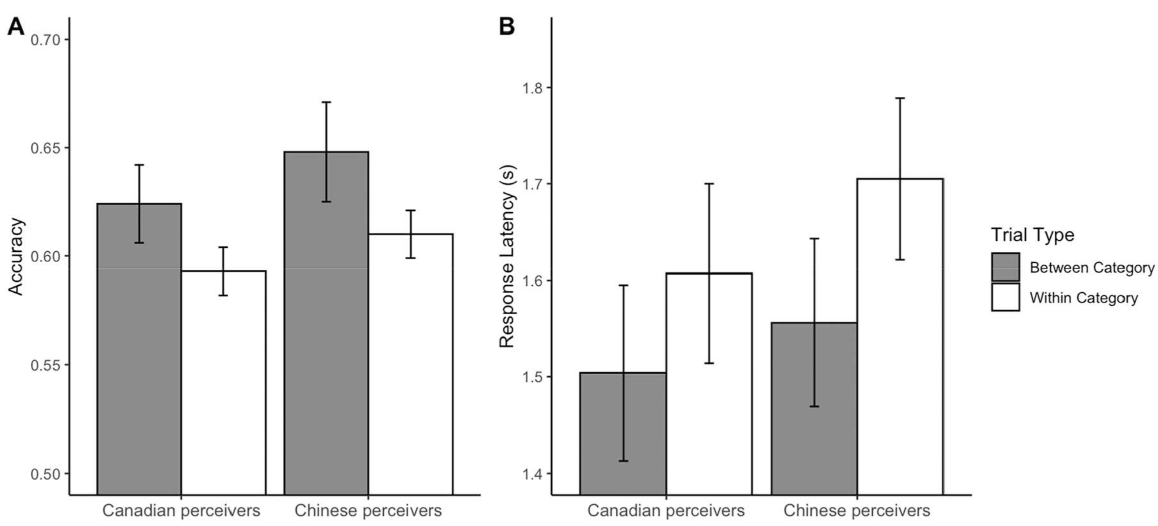

Figure 6. Accuracy (A) and response latency (B) of discrimination task as a function of trial type and perceiver culture based on average category boundaries.

Note: Error bars indicate 95% confidence intervals.

the expressions of anger and disgust portrayed by Asian actors compared to White actors.

To further investigate whether the magnitude of CP varied across cultures of perceivers, we conducted a Bayesian mixed-design ANOVA. The BF for the main effects model (against the Null model) divided by the BF for the interaction model (against the Null model) yielded a value of 3.845, providing moderate evidence that there was no significant interaction between Trial Type and Perceiver Culture.

Overall, consistent with the preregistered analysis, the results revealed a similar pattern of CP for facial expressions across Canadian and Chinese participants, with participants from both cultural groups demonstrating higher accuracy and faster response latencies when categorising between-category pairs of expressions compared to within-category pairs.

Discussion

Supporting H3, which posits that the magnitude of categorical perception (CP) for East Asian perceivers of angry and disgusted expressions does not differ from Western perceivers, the present research demonstrated that participants from both Canada and China exhibited CP of facial expressions of anger and disgust, and the magnitude of CP did not vary across cultures. We employed two approaches to examine this phenomenon: one based on individual category boundaries (as specified in the preregistered analysis) and the other based on average category boundaries (as specified in the exploratory analysis). In both cases, participants from Canada and China showed higher accuracy and faster response latencies when discriminating between pairs of expressions that crossed a category boundary compared to pairs within the same category. Importantly, these differences did not significantly vary between the two cultures. The Bayes factors (BFs) provided "moderate evidence" for the absence of interactions between Trail Type and Perceiver Culture in most cases, with the exception of the BF related to response latency based on individual category boundaries, which was characterised as providing "anecdotal evidence."

Consistent with the findings of Wang et al. (2006), we also found evidence of CP among Chinese participants. However, it is important to note that Wang et al. (2006) primarily focused on the identification of emotions from morphed photographs, which corresponds to the identification task in the present research. They observed an abrupt shift from perceptions of one emotion to another around the midpoint of each emotion pair, indicating a greater sensitivity to changes across emotion boundaries compared to changes within emotions. Consequently, they concluded that Chinese individuals exhibit CP of emotions.

Nevertheless, there are two primary concerns regarding their study. Firstly, the definition of an "abrupt shift from one emotion to another around the midpoint" lacks specificity and clear statistical methods. Their investigation relied solely on facial material from a single actor. Taking the example of the midpoint of the disgust-anger continuum for this actor, it was recognised as anger and disgust with probabilities of approximately 40% and 60%, respectively. Without further statistical analysis, Wang et al. (2006) based their assertion of CP on these observational recognition data. In contrast, our study utilised facial material from four different actors. When considering the identification of morphed photographs at the midpoint of a continuum, there were substantial differences among the actors. In particular, for Chinese perceivers, the midpoint of Asian female and White female (i.e. morph 5) was recognised as anger and disgust with probabilities around 50%, the midpoint of Asian male (i.e. morph 5) was recognised as anger and disgust with probabilities of 82% and 18%, respectively, and the midpoint of White male (i.e. morph 5) was recognised as anger and disgust with probabilities of 32% and 68%, respectively (see Figure 4(A)). If we apply the criteria of Wang et al. (2006), it remains unclear whether these data can provide evidence for the existence of CP among Chinese participants, or if Chinese participants exhibit CP only for certain actors (e.g. Asian and White males) but not for others (e.g. Asian and White females).

Secondly, Wang et al. (2006) did not investigate a crucial feature of CP – the superior discrimination of stimuli across category boundaries compared to stimuli within the same category. In our study, the identification task served as a basis for defining between-category pairs and within-category pairs in the discrimination task. Our findings revealed that Chinese participants exhibited higher accuracy and faster response latencies when discriminating between-category pairs of expressions compared to within-category pairs, providing further support for the existence of CP for expressions of anger and disgust among Chinese participants.

It is worth noting that in this research, category boundaries were defined using two different approaches: calculating individual category boundaries for each participant with respect to each continuum, and calculating category boundaries for each continuum separately for all Chinese and Canadian participants (as commonly employed in previous research, e.g. Calder et al., 1996; Cong et al., 2019; Young et al., 1997). The advantage of the first approach is that it effectively controls for individual differences. For instance, if participant A's category boundary for a particular continuum falls between morphs 4 and 5, while participant B's boundary for the same continuum falls between morphs 5 and 6, then the between-category pair for participant A would be morphs 4 and 5, while the between-category pair for participant B would be morphs 5 and 6. However, if we adopt the second approach, selecting one of these pairs as the between-category pair would render it inaccurate for the other participant. Nevertheless, the first approach has its limitations. A considerable portion of participants' identifications of specific morphed photographs as anger or disgust fell exactly at 50%. This made it impossible to define between-category pairs and within-category pairs for the discrimination task, resulting in the exclusion of a substantial amount of data. In contrast, the second approach allowed us to retain most of the data, and the average category boundaries obtained were applicable to the majority of participants. Importantly, regardless of the approach used, we obtained consistent results, indicating that both Canadian and Chinese participants exhibited similar magnitudes of CP for expressions of anger and disgust.

Previous studies have shown that East Asians are more likely to confuse anger and disgust compared to Westerners (Jack et al., 2009; Matsumoto, 1992), and they are more likely to perceive the coexistence of both emotions in anger and disgust expressions (Fang et al., 2018, 2019). It is important to note that these studies commonly employed tasks such as forced-choice or multidimensional emotion ratings, which share similarities with the identification task used in our study, as they involve emotion labels (words). As shown in Figure 5, the differences between the cultural groups in identification performance were much larger than their differences in discrimination performance. This suggests the possibility that language may amplify cultural differences in the perception of emotional facial expressions. Consistent with the findings of Sauter et al. (2011), when participants performed discrimination tasks that did not involve emotion labels, they exhibited better discrimination for pairs of expressions that crossed category boundaries compared to pairs within the same category. Furthermore, this difference was not modulated by the cultural background of the perceivers. These results indicate that both Canadian and Chinese participants exhibited CP for anger and disgust expressions, with no significant differences observed between the two groups. However, caution must be exercised when interpreting these findings. Although the BF related to accuracy provides moderate evidence to support the absence of an interaction between Trial Type and Perceiver Culture, the BF related to response latency only provides anecdotal evidence to support the absence of an interaction.

Limitations and future directions

The present research is not without limitations. Firstly, this study focused on examining and demonstrating CP for facial expressions of anger and disgust among individuals from Canada and China. However, it remains unclear whether these findings can be generalised to other emotion pairs. For example, fear and surprise share perceptual similarities, which might lead perceivers to easily misidentify fear as surprise, especially among East Asians who have shown a higher susceptibility to confusing these two expressions compared to Western individuals (Beaupré & Hess, 2005; Matsumoto, 1992; Yik & Russell, 1999). Therefore, future research could explore whether individuals from Canada and China exhibit similar levels of category perception for fear and surprise.

Secondly, to ensure consistent facial muscle movements for the same expressions across actors of different ethnicities, we relied on Western prototypes of anger and disgust in this study (Ekman et al., 2002). Specifically, all anger stimuli included brow lowerer (AU4), chin raiser (AU17), and lip presser (AU24), while all disgust stimuli included brow lowerer (AU4), eyelids tightener (AU7), nose wrinkler (AU9), and upper lip raiser (AU10) (see Table S1). It is noteworthy that previous research has shown that different prototypes of disgust, such as those involving AU9 without AU10 or AU10 without AU9, convey varying degrees of disgust and anger (Fang & Kawakami, 2024; Rozin et al., 1994; Widen & Russell, 2008). Future studies could explore alternative prototypes of disgust to determine if the findings of the present research remain consistent. Moreover, recent studies have highlighted subtle yet systematic differences in facial expressions of emotion when expressed by individuals from different cultures (Elfenbein et al., 2007; Fang et al., 2022). Hence, it would be of great value for forthcoming research to examine the generalizability of the current findings to prototypes of anger and disgust based on Eastern prototypes.

Lastly, the study compared participants from Canada and China, but it would be beneficial to include participants from additional cultural backgrounds. This would allow for more cross-cultural comparisons and better assessment of the generalizability of the findings. Moreover, investigating whether cultural differences in CP extend beyond East Asian and Western cultures would provide a richer understanding of the influence of culture on emotion perception.

Conclusions

In conclusion, the present study provides evidence for CP of facial expressions of anger and disgust among Canadian and Chinese participants. The findings suggest that both cultural groups exhibited similar patterns of CP, emphasising this perceptual phenomenon transcending cultural boundaries. However, further research is needed to explore other emotions and expressions based on non-Western prototypes, as well as to include participants from more diverse cultural backgrounds. These future studies will further deepen our understanding of how facial expressions are perceived categorically and shed light on the influence of culture on emotion perception.

Notes

1. It is noteworthy that conducting a power analysis for NHST with the same criteria (Cohen's $d = 0.43$ and $\alpha = .05$) reveals that a sample size of only 45 participants from each cultural group would be sufficient to achieve 80% power. This sample size exceeds the majority of previous studies on categorical perception of facial expressions (e.g. Etcoff & Magee, 1992; Sauter et al., 2011; Young et al., 1997), or is comparable to them (Cong et al., 2019).
2. This pilot study was conducted together with another pilot study (see Fang et al., 2019), the purpose of which was to match the perceived intensity of various emotional expressions on Asian and White faces for use in future studies.
3. The intervals between the presented face pairs in the discrimination task of this study were carefully selected. When utilizing two-step intervals, a substantial difference in the degree of mouth opening between face A and face B was observed. Consequently, participants were able to accurately discern which face corresponded to the previously presented target face solely based on the degree of mouth opening. Conversely, when employing one-step intervals, the disparity in the degree of mouth opening between face A and face B was relatively small. As a result, participants were unable to determine which face matched the previously presented target face based solely on the degree of mouth opening. Therefore, we opted to utilize a one-step interval (10%) for our experiment to enable a more nuanced examination of the participants' perceptual abilities compared to intervals involving two or more steps.
4. It is important to highlight that the exclusion of participant data in this research was not due to their failure to diligently complete the tasks as required. Rather, it was a consequence of the rigorous data selection criteria employed for the preregistered analysis. Nonetheless, even with this stringent approach, the valid sample size of participants in the present research still exceeds that of previous research on categorical perception of facial expressions. Furthermore, to mitigate data loss, we also conducted (generalized/) linear mixed-effects model analyses that were more tolerant of missing values. This analysis utilized data from 126 Canadian participants and 97 Chinese participants, yielding results consistent with those reported in the main text. For detailed information, please refer to the Supplementary Materials.
5. To better quantify the similarity between angry and disgusted expressions of each actor, we utilized computer vision techniques. By employing opencv-python for image grayscaling and dlib with a pre-trained model (Sagonas et al., 2013) for detecting facial landmarks, we generated 128-dimensional facial embeddings. These embeddings captured the essential features of each face. To measure the similarity between expressions, we calculated the Euclidean distance between their corresponding embeddings. The resulting distances were as follows: 0.37 (AF), 0.44 (AM), 0.45 (WF), and 0.53 (WM). These values indicate that Asian actors exhibited lower Euclidean distances compared to White actors, suggesting a higher degree of similarity within their angry and disgusted expressions.

Acknowledgments

The authors would like to express their gratitude to Zhihe Pan and Jinpeng Fu for her valuable contributions in data collection and analysis from China, as well as to Hannah Tran for her valuable contributions in data collection from Canada.

Author contributions

1. X. Fang and D. Sauter developed the study concept and design. X. Fang collected data. X. Fang performed

the data analysis and interpretation. X. Fang drafted the manuscript, and G. Kleef, K. Kawakami, D. Sauter provided critical revisions. All authors approved the final version of the manuscript for submission.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Social Sciences and Humanities Research Council of Canada [grant number 430-2020-00908].

ORCID

Disa A. Sauter http://orcid.org/0000-0003-4872-0536

References

Angeli, A., Davidoff, J., & Valentine, T. (2008). Face familiarity, distinctiveness, and categorical perception. *Quarterly Journal of Experimental Psychology*, *61*(5), 690–707. https://doi.org/10.1080/17470210701399305

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*(1), 149–178. https://doi.org/10.1146/annurev.psych.56.091103.070217

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*(3), 217–239. https://doi.org/10.1016/0010-0277(95)00669-X

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, *36*(3), 355–370. https://doi.org/10.1177/0022022104273656

Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research*, *46*(3), 207–222. https://doi.org/10.1007/BF00308884

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition*, *3*(2), 81–118. https://doi.org/10.1080/713756735

Chapman, H. A., & Anderson, A. K. (2013). Things rank and gross in nature: a review and synthesis of moral disgust. *Psychological Bulletin*, *139*(2), 300–327. https://doi.org/10.1037/a0030964

Chen, L. F., & Yen, Y. S. (2007). *Taiwanese facial expression image database*. Brain Mapping Laboratory. Institute of Brain Science, National Yang-Ming University.

Cong, Y. Q., Junge, C., Aktar, E., Raijmakers, M., Franklin, A., & Sauter, D. (2019). Pre-verbal infants perceive emotional facial expressions categorically. *Cognition and Emotion*, *33*(3), 391–403. https://doi.org/10.1080/02699931.2018.1455640

De Gelder, B., Teunisse, J. P., & Benson, P. J. (1997). Categorical perception of facial expressions: Categories and their internal structure. *Cognition & Emotion*, *11*(1), 1–23. https://doi.org/10.1080/026999397380005

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Ekman, P., Friesen, W. V., & Hager, J. V. (2002). *Facial action coding system* (2nd ed.). Research Nexus Ebook.

Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion*, *7*(1), 131–146. https://doi.org/10.1037/1528-3542.7.1.131

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227–240. https://doi.org/10.1016/0010-0277(92)90002-Y

Fang, X., & Kawakami, K. (2024). Decoding angry and disgusted faces across cultures: Facial prototypes and software matter. *Journal of Nonverbal Behavior*, 1–26. https://doi.org/10.1007/s10919-024-00453-0

Fang, X., Sauter, D. A., Heerdink, M. W., & van Kleef, G. A. (2022). Culture shapes the distinctiveness of posed and spontaneous facial expressions of anger and disgust. *Journal of Cross-Cultural Psychology*, *53*(5), 471–487. https://doi.org/10.1177/00220221221095208

Fang, X., Sauter, D. A., & Van Kleef, G. A. (2018). Seeing mixed emotions: The specificity of emotion perception from static and dynamic facial expressions across cultures. *Journal of Cross-Cultural Psychology*, *49*(1), 130–148. https://doi.org/10.1177/0022022117736270

Fang, X., Van Kleef, G. A., & Sauter, D. A. (2019). Revisiting cultural differences in emotion perception between easterners and westerners: Chinese perceivers are accurate, but see additional non-intended emotions in negative facial expressions. *Journal of Experimental Social Psychology*, *82*, 152–159. https://doi.org/10.1016/j.jesp.2019.02.003

Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*(2), 212–228. https://doi.org/10.1037/0022-3514.57.2.212

Fugate, J. M. (2013). Categorical perception for emotional faces. *Emotion Review*, *5*(1), 84–89. https://doi.org/10.1177/1754073912451350

Gagnon, M., Gosselin, P., Hudon-ven Der Buhs, I., Larocque, K., & Milliard, K. (2010). Children's recognition and discrimination of fear and disgust facial expressions. *Journal of Nonverbal Behavior*, *34*(1), 27–42. https://doi.org/10.1007/s10919-009-0076-z

Grossmann, I., Huynh, A. C., & Ellsworth, P. C. (2016). Emotional complexity: Clarifying definitions and cultural correlates. *Journal of Personality and Social Psychology*, *111*(6), 895–916. https://doi.org/10.1037/pspp0000084

Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852–870). Oxford University Press.

Harnad, S. (1987). *Categorical perception: The groundwork of cognition*. Cambridge University Press.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, *19*(18), 1543–1548. https://doi.org/10.1016/j.cub.2009.07.051

Jehna, M., Neuper, C., Ischebeck, A., Loitfelder, M., Ropele, S., Langkammer, C., Ebner, F., Fuchs, S., Schmidt, R., Fazekas, F., & Enzinger, C. (2011). The functional correlates of face perception and recognition of emotional facial expressions as evidenced by fMRI. *Brain Research*, *1393*, 73–83. https://doi.org/10.1016/j.brainres.2011.04.007

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115–1125. https://doi.org/10.1068/p3155

Lee, V., Cheal, J. L., & Rutherford, M. D. (2015). Categorical perception along the happy–angry and happy–sad continua in the first year of life. *Infant Behavior and Development*, *40*, 95–102. https://doi.org/10.1016/j.infbeh.2015.04.006

Levin, D. T., & Angelone, B. L. (2002). Categorical perception of race. *Perception*, *31*(5), 567–578. https://doi.org/10.1068/p3315

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358–368. https://doi.org/10.1037/h0044417

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-cultural Psychology*, *23*(1), 72–84. https://doi.org/10.1177/0022022192231005

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238. https://doi.org/10.1037/0033-295X.85.3.207

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes factors for common resigns. R package version 0.9.12-4.2*. https://CRAN.R-project.org/package=BayesFactor.

Nabi, R. L. (2002). The theoretical versus the lay meaning of disgust: Implications for emotion research. *Cognition & Emotion*, *16*(5), 695–703. https://doi.org/10.1080/02699930143000437

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*(3), 382–407. https://doi.org/10.1016/0010-0285(72)90014-X

Roy-Charland, A., Perron, M., Beaudry, O., & Eady, K. (2014). Confusion of fear and surprise: A test of the perceptual-attentional limitation hypothesis with eye movement monitoring. *Cognition and Emotion*, *28*(7), 1214–1222. https://doi.org/10.1080/02699931.2013.878687

Rozin, P., Lowery, L., & Ebert, R. (1994). Varieties of disgust faces and the structure of disgust. *Journal of Personality and Social Psychology*, *66*(5), 870–881. https://doi.org/10.1037/0022-3514.66.5.870

Russell, J. A., & Fehr, B. (1994). Fuzzy concepts in a fuzzy hierarchy: varieties of anger. *Journal of Personality and Social Psychology*, *67*(2), 186–205. https://doi.org/10.1037/0022-3514.67.2.186

Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: what it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological Bulletin*, *139*(2), 328–351. https://doi.org/10.1037/a0029319

Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., & Pantic, M. (2013). 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *Proceedings of the IEEE international conference on computer vision workshops* (pp. 397–403).

Sauter, D. A. (2018). Is there a role for language in emotion perception? *Emotion Review*, *10*(2), 111–115. https://doi.org/10.1177/1754073917693924

Sauter, D. A., Le Guen, O., & Haun, D. B. M. (2011). Categorical perception of emotional facial expressions does not require lexical categories. *Emotion*, *11*(6), 1479–1483. https://doi.org/10.1037/a0025336

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*(4), 813–838. https://doi.org/10.1037/0022-3514.48.4.813

Van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, *11*(4), 907–920. https://doi.org/10.1037/a0023853

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432. https://doi.org/10.1037/a0022790

Wang, K., Hoosain, R., Lee, T. M., Meng, Y., Fu, J., & Yang, R. (2006). Perception of six basic emotional facial expressions by the Chinese. *Journal of Cross-Cultural Psychology*, *37*(6), 623–629. https://doi.org/10.1177/0022022106290481

Whitton, A. E., Henry, J. D., Rendell, P. G., & Grisham, J. R. (2014). Disgust, but not anger provocation, enhances levator labii superioris activity during exposure to moral transgressions. *Biological Psychology*, *96*, 48–56. https://doi.org/10.1016/j.biopsycho.2013.11.012

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, *39*(1), 114–128. https://doi.org/10.1037/0012-1649.39.1.114

Widen, S. C., & Russell, J. A. (2008). Children's and adults' understanding of the "disgust face". *Cognition & Emotion*, *22*(8), 1513–1541. https://doi.org/10.1080/02699930801906744

Yik, M. S. M., & Russell, J. A. (1999). Interpretation of faces: A cross-cultural study of a prediction from Fridlund's theory. *Cognition & Emotion*, *13*(1), 93–104. https://doi.org/10.1080/026999399379384

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*(3), 271–313. https://doi.org/10.1016/S0010-0277(97)00003-6