Journal of Nonverbal Behavior (2024) 48:439–464
https://doi.org/10.1007/s10919-024-00453-0

ORIGINAL PAPER

Decoding Angry and Disgusted Faces Across Cultures: Facial Prototypes and Software Matter

Xia Fang[1] · **Kerry Kawakami**[2]

Accepted: 5 January 2024 / Published online: 14 February 2024

Abstract
Although the same emotion can be related to different facial prototypes, little is known about the impact of differences in these prototypes on emotion perception. The present research aimed to investigate the impact of different facial prototypes on emotion perception, focusing specifically on anger and disgust. Four major prototypes of anger and six major prototypes of disgust by Asian and White targets were generated by FaceGen and MetaHuman programs. These prototypes were presented to participants from Canada and China, who were asked to identify the corresponding emotions. The results revealed that emotional signals conveyed by different facial prototypes were not uniform. While some prototypes cued distinctive emotions, others cued mixed emotions. Moreover, the signal value of specific prototypes varied across perceiver culture, with prototypes generally perceived as more mixed in Chinese compared to Canadian participants. Notably, emotional signals were also influenced by the facial generation software. Together these findings highlight the importance of considering the impact of specific facial prototypes on emotion perception, especially in cross-cultural contexts, and caution researchers to be mindful of the facial generation software when examining potential variability in emotional decoding associated with different facial prototypes.

Keywords Emotion identification · Facial prototypes · Culture · Specific · Mixed emotions

Introduction

What does an angry face look like? Is the mouth open or closed? Are the eyebrows furrowed or arched? How about the eyes? Notably, as with other emotions, there are several variants of prototypes related to anger. These variants consist of different combinations of facial muscle movements. The Facial Action Coding System (FACS) developed by Ekman and Friesen (1978) is the most commonly used approach for decomposing facial expressions into action units (AUs: minimal units of discrete changes in facial expression that

✉ Xia Fang
x.fang@zju.edu.cn

[1] Department of Psychology and Behavioral Sciences, Zhejiang University, 866 Yuhangtang Rd., Hangzhou, Zhejiang Province, China

[2] York University, Toronto, Canada

are due to the innervations of one or several muscles). Based on theoretical considerations and systematic observation, Ekman et al. (2002) developed a list of combinations of AUs that are associated with basic emotions (anger, disgust, fear, sadness, surprise, and happiness), with each emotion having more than one facial prototype. For example, anger can be expressed either by lowering eyebrows (AU4), widening eyes (AU5), and pressing lips (AU24) or by lowering eyebrows (AU4) and opening the mouth (AU25) and dropping the jaw (AU26).

Surprisingly, researchers have devoted little attention to examining how these variations in facial prototypes associated with one emotion impact emotion perception. As demands for more rigorous control over stimuli to investigate fine-grained discriminations between emotions increase (Gosselin & Kirouac, 1995; Pochedly et al., 2012; Widen et al., 2013), understanding the role of variations in facial expressions is important. The primary goal of the present research was to investigate the extent to which different facial prototypes related to anger and disgust impact emotion perception. A further goal was to examine whether the impact of these prototypes on emotion perception differ across East Asian and Western cultures.

Facial Prototypes of Anger and Disgust

Anger and disgust are widely held to be distinct emotions, characterized by different patterns of facial, vocal, and autonomic physiological components and appraisals (Ekman & Cordaro, 2011). Yet these two emotions are conceptually related (Giner-Sorolla et al., 2018) and both emotions are associated with appraisals of negative events caused by others (Frijda et al, 1989; Scherer, 1988; Smith & Ellsworth, 1985). Moreover, the facial morphological features of these two emotions share a certain degree of similarity. For example, some studies have shown that lowered eyebrows (AU4) and tightened eye lids (AU7) are present in both anger and disgust expressions (Cordaro et al., 2018; Fang et al., 2022). People, therefore, often confuse facial expressions related to disgust with anger and vice versa (Aviezer et al., 2008; Ekman & Friesen, 1976; Jack et al., 2009; Widen & Russell, 2008).

Because the confusion between anger and disgust is partially caused by overlapping morphological features, different facial prototypes might be related to the ability to distinguish between these emotions. Specifically, when prototypes of anger and disgust share more morphological features, people may be less able to distinguish between these emotions. Alternatively, when prototypes share fewer overlapping features, people may be better at decoding a specific emotion. Despite these potential differences in responses to emotions depicted by different prototypes, few studies have investigated whether and how different prototypes impact emotion identification.

Previous studies on judgements of standard facial expressions of disgust provide some suggestive evidence that different prototypes may differ in their perceived meanings (Rozin et al., 1994; Yoder et al., 2016). Although Ekman and Friesen (1978) specified that standard expressions of disgust must consist of at least one or both the AU9 (nose wrinkler) and the AU10 (upper lip raiser), research has shown that disgust prototypes that consist of the AU9 or AU10 are associated with different disgust elicitors (Rozin et al., 1994). Russell and colleagues (Pochedly et al, 2012; Widen & Russell, 2008; Widen et al., 2013; Yoder et al., 2016) further suggest that standard facial expressions of anger and disgust may convey a mixture of both. For example, in an experiment, Widen et al. (2004) asked participants to indicate yes or no for the questions "Is this person angry/disgusted/happy?" for standard expressions of anger, disgust, fear, and happiness proposed by Ekman et al.

(2002). They found that many participants perceived both standard angry and disgusted faces as angry and disgusted. Furthermore, when asked to select angry or disgusted targets, participants included both standard angry and disgusted faces in both conditions. While these findings suggest that some prototypes of anger and disgust may convey mixed emotions instead of a singular emotion, researchers did not investigate which prototypes were more likely to convey one emotion or a mixture of both emotions. In the present research, we summarized major variants of facial prototypes of anger and disgust from previous work, with the aim of investigating how each prototype of anger and disgust is perceived to be predominantly angry, predominantly disgusted or a mix of both emotions.

Culture and Emotion Perception

A further goal was to examine whether these facial prototypes are decoded differently by members of Eastern and Western cultures. While early cross-cultural studies on emotion perception found evidence supporting the universality of facial expressions (Ekman & Friesen, 1971; Ekman et al., 1969, 1987; Izard, 1971), more recent research suggests that cultural differences in emotion recognition may exist. For example, East Asians compared to Westerners may exhibit lower recognition accuracy for standard expressions of certain negative emotions (Matsumoto & Ekman, 1989; Nelson & Russell, 2013; Russell, 1994; Yik, 1999). The findings regarding these cross-cultural differences, however, have not been consistently replicated (Ekman et al., 1987). For example, while Jack et al. (2009) found that East Asians were worse than Western Europeans at recognizing expressions of disgust and fear but not anger, Beaupré and Hess (2005) demonstrated that participants from China compared to Canada were less accurate at recognizing expressions of sadness but not anger, disgust, or fear.

Although these inconsistencies may be related to the particular emotions included in the study (Nelson & Russell, 2013; Russell et al., 1994; Yik et al., 2013), the choice of emotion alternatives (Russell et al., 1994; Widen et al., 2013) or language and translations (Kollareth & Russell, 2017; Matsumoto & Hwang, 2017; Matsumoto et al., 2008; Yoder et al., 2016), it is reasonable to speculate that the different results may also be related to variations in facial prototypes. For example, it is possible that some prototypes related to disgust are perceived similarly across cultures, while other disgust prototypes are perceived differently across cultures. When different prototypes related to the same emotion are used in different studies, cultural differences on emotion perception may also differ. It is therefore important to compare across cultures the signal value of different prototypes related to the same emotion.

Recent research has also suggested that Easterners' lower overall recognition accuracy for negative facial expressions may be related to their tendency to perceive mixed emotions in a facial expression (Fang et al., 2018, 2019). For example, when judging a face with a standard disgusted expression, both East Asians and Westerners perceive disgust as the predominant emotion. However, East Asians compared to Westerners also perceive other non-intended emotions, such as anger and fear to a greater degree (Fang et al., 2018). This greater inclination of East Asians to perceive mixed emotions may be associated with their holistic cognitive style and dialectical thinking. Because East Asians place greater emphasis on the entire field and rely less on categorical distinctions (Ji et al., 2000; Markus & Kitayama, 1991; Nisbett et al., 2001), and demonstrate a higher tolerance for inconsistencies in their emotions, thoughts, and behaviors (Peng & Nisbett, 1999; Peng et al., 2006), they may perceive facial expressions as containing multiple emotions. In the present

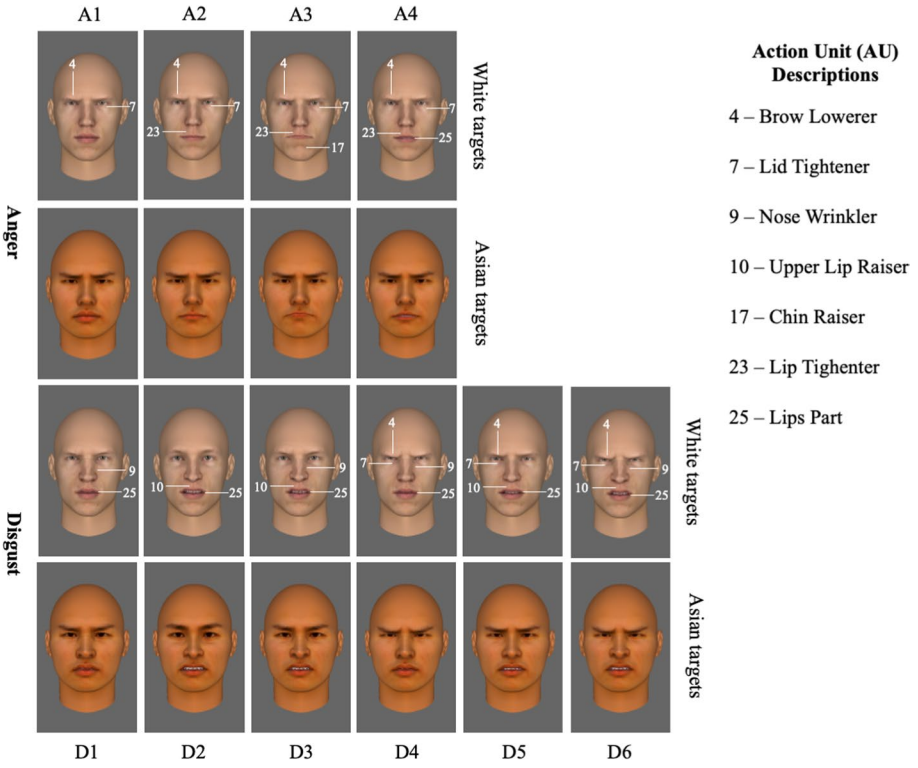

Fig. 1 Facial prototypes of anger and disgust of White and Asian targets Generated by FaceGen Modeller in Experiment 1. *Note.* All prototypes of anger and disgust meet the criteria set by Ekman et al. (2002) and are considered standard facial expressions

research we thus expected that East Asians would be more likely than Westerners to perceive mixed emotions across facial prototypes.

Emotional Facial Expression Software

Computer-generated faces are widely used as experimental stimuli in the field of psychology to investigate perception of faces and facial expressions (Friesen et al., 2019; Malek et al., 2019; Oosterhof & Todorov, 2008). Although compared to real human faces, computer-generated faces lack a certain level of realism, they allow for precise control over facial features associated with age, race, gender, head position, and even individual muscle movements (Friesen et al., 2019; Mende-Siedlecki et al., 2019). This capability ensures that faces can be generated and used as experimental stimuli that are consistent across all aspects, except the manipulated factors. Among these face synthesis software, one of the most widely used face synthesis software programs is FaceGen (https://facegen.com/) Modeller (Hehman et al., 2015, 2017; Thorstenson et al., 2019). For examples of images created with this software, see Fig. 1. An important reason why faces generated by FaceGen may be perceived as artificial is that this software does not fully simulate specific facial muscle movements. For example, Miller et al. (2022) found that the FaceGen AU tools did not adequately produce "crow's feet"

wrinkles around the eyes or the pouches around the eyes that are associated with the Duchenne marker; it only narrowed the eye aperture. Due to the poor simulation of the Duchenne marker, they did not find any differences between happy faces with and without the Duchenne marker with FaceGen stimuli but they found that happy faces with the Duchenne marker in real human faces were perceived as more authentic and had higher intensity compared to those without the Duchenne marker (but see Friesen et al., 2019 for alternative findings).

In 2021, new facial software was released called MetaHuman Creator (https://www.unrealengine.com), which produces high-quality digital faces with significant improvements in realism. For examples of images created with this software, see Fig. 3. Notably, when manipulating the same muscle movements in MetaHuman and FaceGen, slight differences in muscle changes can be observed. For example, when manipulating the muscle movement of nose wrinkler (AU9) in facial expressions, the resulting changes in FaceGen generated faces (FG faces) primarily involved a lowered brow, narrowed eyes, and a horizontal wrinkle across the bridge of the nose, see D1 in Fig. 1. Changes in MetaHuman generated faces (MH faces), alternatively, included a slight lowering of the brow, narrowed eyes, several vertical wrinkles on the bridge of the nose, and deepened nasolabial lines, see D1 in Fig. 3. Furthermore, when manipulating the movement of brow lowerer (AU4) in facial expressions, FG faces did not exhibit any furrows (e.g., A1 in Fig. 1), while MH faces had several vertical furrows between the eyebrows (e.g., A1 in Fig. 3).

Overall, although MetaHuman compared to FaceGen appears to be more realistic in simulating facial muscle changes associated with various emotional expressions, no studies have been conducted to compare the perception of facial expressions generated by these two software programs. If there are indeed differences in the perception of facial expressions from these programs, then researchers should be cautious in selecting software for generating facial expressions in future studies. A further aim of the present research was therefore to compare facial expressions produced by FaceGen and MetaHuman and examine whether the specific software affects perceptions of facial expressions.

The Present Research

In summary, the goals of the present research were to investigate (1) whether different prototypes associated with the same emotion (i.e., anger or disgust) were decoded differently, (2) whether Chinese compared to Canadian participants perceived more mixed emotions in these facial prototypes, and (3) whether prototypes generated by different software (FaceGen and MetaHuman) were decoded differently. To this end, three experiments were conducted in which Chinese and Canadian participants decoded facial prototypes related to anger and disgust, respectively. In particular, Chinese and Canadian participants viewed FG faces of White (Experiment 1a) and East Asian (Experiment 1b) targets, and MH faces of White and East Asian targets (Experiment 2). On each trial, participants made judgments about whether the target expression depicted anger (yes or no) or disgust (yes or no). To explore whether these findings could be extended to different evaluation paradigms, Experiment 3 asked Chinese and Canadian participants to rate the intensity of anger and disgust for each expression generated by MH faces of White and East Asian targets.

Table 1 Facial action units for the facial expressions of anger and disgust

References	Reference AUs for anger	Reference AUs for disgust	Physical description
Ekman et al., (2002)	4+5+7+10+22+23+25, 26	9	Upper face: 4: Brow lowerer 5: Upper lid raiser 6: Cheek raiser 7: Lid tightener 9: Nose wrinkler Lower face: 10: Upper lip raiser 12: Lip corner puller 16: Lower lip depressor 17: Chin raiser 22: Lip funneler 23: Lip tightener 24: Lip presser 25: Lips part 26: Jaw drop
	4+5+7+l0+23+25, 26	9+16+25, 26	
	4+5+7+23+25, 26	9+17	
	4+5+7+17+23	10	
	4+5+7+17+24	10+16+25, 26	
	4+5+7+23	10+17	
	4+5+7+24		
Elfenbein et al., (2007)	4, 5, 7, 17	4, 6, 9, 10, 12, 17, 25, 26	
Cordaro et al., (2018)	4, 7	4, 6, 7, 9, 10, 25, 26	
Fang et al., (2022)	4, 7, 23	4, 7, 9, 10, 17, 25	

In Ekman et al.'s work (2002), AUs for expressions are denoted by combinations of AUs (AUs combined with "+"; except AU26, which can be present or not for a given combination). Other emotion production work (Cordaro et al., 2018; Elfenbein et al., 2007; Fang et al., 2022) represents AUs for expressions using single AUs (AUs separated with commas, indicating they are not necessarily all present in a single expression)

Development of Major Facial Prototypes of Anger and Disgust

In order to secure a comprehensive yet concise set of facial prototypes of anger and disgust, we took a number of steps to develop major variants of prototypes related to anger and disgust. First, we summarized empirical work on facial prototypes of anger and disgust. In particular, we included the FACS manual (Ekman et al., 2002), as well as cross-cultural studies on emotion production (Cordaro et al., 2018; Elfenbein et al., 2007; Fang et al., 2022). Specific AUs (or combinations of AUs) related to expressions of anger and disgust were extracted from this work and summarized in Table 1.

Based on the frequency of occurrence of AUs in Table 1, we identified critical AUs for specific emotions. For instance, AU4 (brow lowerer) and AU7 (lid tightener) are present in all prototypes associated with anger, thus establishing them as critical AUs for anger. Similarly, because the presence of AU9 (nose wrinkler) and/or AU10 (upper lip raiser) is included in all prototypes related to disgust, they are also identified as critical AUs for disgust. With these critical AUs as the foundation, we constructed emotion prototypes ranging from minimal to more complex representations. The simplest prototypes encompass only the critical AUs (e.g., AUs 4+7 for Anger 1), while more comprehensive prototypes progressively incorporate additional AUs. This approach enables a systematic examination of emotional expressions through the combination of critical and supplementary AUs.

Specifically, the inclusion of Action Units (AUs) in the prototypes was guided by several criteria:

1. Critical AUs for anger and disgust were included in all prototypes related to anger and disgust, respectively. This ensured that the most salient AUs associated with these emotions were represented.

2. AUs were included sequentially based on their relative importance, as determined by their frequency of occurrence in the expression of a specific emotional prototype. This approach allowed for the inclusion of AUs that are most commonly associated with a particular emotion and thus are likely to be critical for the accurate representation of that emotion.
3. A new AU was only included if it produced perceptual differences in facial appearance from an existing AU. This ensured that the addition of a new AU contributed to the overall differentiation from the existing emotional prototypes. For example, AU5 (upper lip raiser) was not included because when AU5 was added to AU4, it resulted in a facial appearance similar to having AU4 (brow lowerer) alone in FaceGen. Similarly, AU24 (lip presser) was not included because it produced a similar facial appearance to AU23 (lip tightener).

Based on these criteria, we selected a set of facial AUs to develop prototypes of anger and disgust. Specifically, we included AU4 (brow lowerer), AU7 (lid tightener), AU23 (lip tightener), AU17 (chin raiser), and AU25 (lips part) to develop four major prototypes of anger (Anger 1: AUs 4+7; Anger 2: AUs 4+7+23; Anger 3: AUs 4+7+17+23; and Anger 4: AUs 4+7+23+25). For disgust, we included AU9 (nose wrinkle), AU10 (upper lip raiser), AU25 (lips part), AU4 (brow lowerer), and AU7 (lid tightener) to develop six major prototypes of disgust (Disgust 1: AUs 9+25; Disgust 2: AUs 10+25; Disgust 3: AUs 9+10+25; Disgust 4: AUs 4+7+9+25; Disgust 5: AUs 4+7+10+25; and Disgust 6: AUs 4+7+9+10+25).

Experiment 1

In Experiment 1, we investigated the perception of Canadian and Chinese participants of 10 facial prototypes of expressions. In Experiment 1a all targets were White men and in Experiment 1b all targets were Asian men. Because the only difference between these two experiments was target race, with all other aspects of the experiments identical, we reported the results of these two experiments together.

Method

Participants

To maximize power, we used a 2 (Emotion Judgment: Anger, Disgust) × 10 (Facial Prototype: A1, A2, A3, A4, D1, D2, D3, D4, D5, D6) × 2 (Perceiver Culture: Canadian, Chinese) mixed design, with the first two factors as within-subject and Perceiver Culture as between-subjects. Given this design, we aimed to recruit approximately 100 people in each cultural group for both Experiments 1a and 1b. The final sample in Experiment 1a included 116 (60 men and 56 women; $M_{age} = 19.60$, $SD = 2.16$) participants born and raised in Canada and recruited at a university in Ontario and 109 (53 men and 56 women; $M_{age} = 20.81$, $SD = 2.43$) participants born and raised in China and recruited at several universities in China. The final sample in Experiment 1b included 73 (16 men and 57 women; $M_{age} = 18.74$, $SD = 1.24$) participants born and raised in Canada and recruited from a university in Ontario and 74 (15 men and 59 women; $M_{age} = 19.80$, $SD = 1.23$) participants born and raised in China and recruited from several universities in China. Canadian and

Chinese participants received either course credit for their participation or volunteered, respectively. We merged Experiment 1a and 1b samples and treated Target Race as a between-subjects factor. A sensitivity analysis using G*Power (Faul et al., 2007) indicated that our final sample could detect a small effect size, Cohen's $d = 0.15$, for the differences between anger and disgust ratings (with $\alpha = .05$).

Stimuli and Procedure

Ten White male faces with neutral emotion from the York Face Database (Vinglis-Jaremko et al., in preparation) and ten Asian male faces with neutral emotion from the Taiwan Facial Expression Image Database (TFEID; Chen & Yen, 2007) were imported into FaceGen.[1] We manipulated specific AUs to create10 prototypes for each model, resulting in a total of 100 White faces and 100 Asian faces (10 prototypes per model × 10 models). Both Experiment 1a with White targets and Experiment 1b with Asian targets were conducted on a Qualtrics survey platform (www.qualtrics.com). In each study, each face was presented twice, with "Anger" versus "Not Anger" under the image in one trial and "Disgust" versus "Not Disgust" under the image in another trial. Participants were instructed to choose the label that best described the displayed emotion. Each participant completed a total of 200 trials (100 faces × 2 types of trials), with the order of faces and alternative choices randomized.

Transparency and Openness

How we determined our sample size, all manipulations, and all measures in the current experiments are reported. All data, analysis code, and research materials are available at https://osf.io/nytj9/vs.view_only=97f92f0f4d3a46aeabd81fcfb724fb09.

Results

Participants' anger scores were calculated as the probability of responding "yes" to the anger judgment trials and disgust score were calculated as the probability of responding "yes" to the disgust judgment trials for each prototype. For example, if a participant rated 6 out of 10 anger judgment trials as expressing anger for the targets of A1, their anger score would be 0.6. Alternatively, if a participant rated 5 out of 10 disgust judgment trials as expressing disgust for the same prototype, their disgust score would be 0.5. The dependent measure for prototypes of anger was calculated as the difference between the anger and disgust scores, while the dependent measure for prototypes of disgust was calculated as the difference between disgust and anger scores. We treated Target Race as a between-subjects factor, related to Experiment 1a and 1b, in a 2 (Perceiver Culture: Canadian, Chinese) × 2 (Target Race: White, Asian) × 4 (Facial Prototype: A1, A2, A3, A4) mixed-design ANOVA for prototypes of anger, and a 2 (Perceiver Culture: Canadian, Chinese) × 2 (Target Race:

[1] When importing the facial images of our models into the FaceGen software, the default is to generate bald-headed avatars. However, the appearance of bald-headed female avatars did not align with our portrayal of female faces and seemed rather unusual. Additionally, although the FaceGen software allows for the inclusion of additional hairstyles, the available options are limited in number and lack realism. Taking all these factors into consideration, we made the deliberate decision to exclusively use male models for Experiment 1.

White, Asian) × 6 (Facial Prototype: D1, D2, D3, D4, D5, D6) mixed-design ANOVA for prototypes of disgust. In both analyses, Facial Prototype was a within-subjects factor, while Perceiver Culture and Target Race were between-subjects factors, see Table 2.

Facial Prototypes of Anger

The primary aim of this study was to investigate whether different facial prototypes related specifically to anger and disgust were perceived differently and whether these effects were moderated by culture. In our first set of analyses, we examined prototypes of Anger (A1, A2, A3, and A4). Our results revealed a significant two-way interaction between Perceiver Culture and Facial Prototype, see Table 2. To initially decompose this effect, we analyzed the impact of Facial Prototype for Canadian and Chinese cultures separately. The left panels of Fig. 2 include the means and standard deviations of emotion scores for all prototypes in both cultures, while the right panels present the Standardized Mean Differences (SMD, Cohen's *d*) between the two scores and their confidence intervals. To account for pairwise comparisons across the four facial prototypes within each cultural group, we used a Bonferroni corrected $p = 0.008$ (0.05/6).

For Canadian participants, pairwise comparisons indicated that the difference score for A3 was significantly higher than for A1, A2, and A4, $ps \leq .005$, Cohen's $ds \geq 0.21$. No other significant differences were found, $ps > .072$, Cohen's $ds \leq 0.13$. Furthermore, we conducted *t*-tests to compare the difference score for each facial prototype with zero, with a significance threshold Bonferroni adjusted to $p = .013$ (0.05/4). Our results indicated that the difference scores for all facial prototypes were significantly greater than zero, $ps < .001$, see Fig. 2 for Cohen's *d*s. Together these findings suggest that all anger prototypes were primarily perceived as conveying anger by Canadian participants, with A3 being the most effective in conveying this emotion.

For Chinese participants, however, no pairwise comparisons were significant among the four anger prototypes, $ps > .031$, Cohen's $ds \leq 0.16$. Furthermore, *t*-tests comparing the difference score for each facial prototype with zero indicated that the difference score for A1 was significantly less than zero, $p < .001$, while the difference scores for A2, A3, and A4 did not significantly differ from zero, $ps > .107$, see Fig. 2 for Cohen's *d*s. Together these results suggest that while A1 was primarily perceived as disgust rather than anger, the other anger prototypes were perceived as a mixture of anger and disgust by Chinese participants.

To further examine the impact of Eastern versus Western cultures on perceptions of emotion prototypes, we decomposed the two-way interaction between Perceiver Culture and Facial Prototype by Facial Prototype. The significance threshold was Bonferroni adjusted to $p = .013$ (0.05/4) due to the comparisons between two cultures across four facial prototypes. For all anger prototypes, the difference scores for Canadian compared to Chinese participants were significantly higher, $ps < .001$, Cohen's $ds \geq 0.49$. As depicted in Fig. 2, Chinese compared to Canadian participants perceived more mixed emotions across all anger prototypes, with difference scores closer to 0.

Facial Prototypes of Disgust

In examining the results related to prototypes of disgust (D1, D2, D3, D4, D5, and D6), the two-way interaction between Perceiver Culture and Facial Prototype was also significant, see Table 2. To initially decompose this effect, we analyzed the impact of Facial Prototype for Canadian and Chinese cultures separately, see Fig. 2. To adjust for the pairwise

Table 2 Mixed-design ANOVA for prototypes of anger and disgust in Experiments 1, 2, and 3

Emotion	Effect	Experiment 1				Experiment 2				Experiment 3			
		F	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2
Anger	Perceiver culture (PC)	67.32	(1, 368)	< .001	0.155	100.63	(1, 196)	< .001	0.339	76.50	(1, 197)	< .001	0.280
	Target race (TR)	0.84	(1, 368)	.359	0.002	0.47	(1, 196)	.495	0.002	6.79	(1, 197)	.010	0.033
	Facial prototype (FP)	8.25	(3, 1104)	< .001	0.022	65.09	(3, 588)	< .001	0.249	99.06	(3, 591)	< .001	0.335
	PC × TR	6.50	(1, 368)	.011	0.017	2.17	(1, 196)	.142	0.011	22.23	(1, 368)	< .001	0.101
	PC × FP	3.60	(3, 1104)	.013	0.010	21.60	(3, 588)	< .001	0.099	11.70	(3, 591)	< .001	0.056
	TR × FP	1.29	(3, 1104)	.275	0.004	0.12	(3, 588)	.945	0.001	1.59	(3, 591)	.190	0.008
	PC × TR × FP	1.25	(3, 1104)	.290	0.003	4.40	(3, 588)	.004	0.022	0.74	(3, 591)	.528	0.004
Disgust	Perceiver culture (PC)	25.67	(1, 368)	< .001	0.065	41.65	(1, 196)	< .001	0.175	12.25	(1, 197)	.001	0.059
	Target race (TR)	4.42	(1, 368)	.036	0.012	0.05	(1, 196)	.822	0.000	8.83	(1, 197)	.003	0.043
	Facial prototype (FP)	80.56	(5, 1840)	< .001	0.180	2.61	(5, 980)	.023	0.013	5.05	(5, 985)	< .001	0.025
	PC × TR	15.70	(1, 368)	< .001	0.041	14.10	(1, 196)	< .001	0.067	36.66	(1, 368)	< .001	0.157
	PC × FP	52.94	(5, 1840)	< .001	0.126	9.92	(5, 980)	< .001	0.048	15.86	(5, 985)	< .001	0.075
	TR × FP	3.38	(5, 1840)	.005	0.009	5.07	(5, 980)	< .001	0.025	1.34	(5, 985)	.245	0.007
	PC × TR × FP	2.17	(5, 1840)	.055	0.006	1.49	(5, 980)	.191	0.008	3.54	(5, 985)	.004	0.018

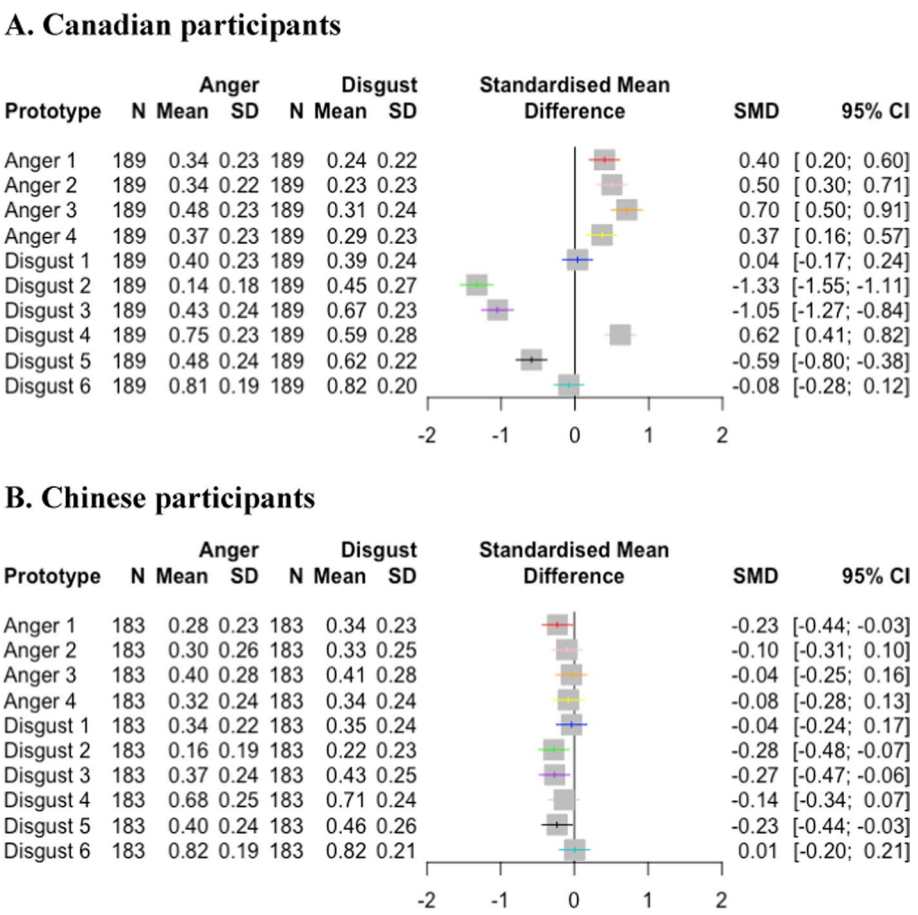

Fig. 2 Differences between anger and disgust scores for facial prototypes generated by FaceGen in Experiment 1. *Note*. Positive scores related to standardized mean differences (*SMD*, Cohen's *d*) reflect higher ratings of anger than disgust while negative scores of *SMD* reflect higher ratings of disgust than anger. For further details on results related to the targets of different races, see Figure S1

comparisons across the six facial prototypes within each cultural group, we used a Bonferroni corrected $p = .003$ (0.05/15).

For Canadian participants, pairwise comparisons indicated that D2 and D3 (with no significant difference between them, $p = .040$, Cohen's $d = 0.19$) were significantly higher than D5, $ps < .001$, Cohen's $ds \geq 0.36$, while D5 was significantly higher than D1 and D6 (with no significant difference between them, $p = .291$, Cohen's $d = 0.08$), $ps < .001$, Cohen's $ds > 0.40$. Furthermore, *t*-tests were conducted to compare the difference score for each facial prototype with zero, with a significance threshold Bonferroni adjusted to .008 (0.05/6). The results indicated that the difference scores for D2, D3, and D5 were significantly greater than zero, $ps < .001$, while the difference score for D4 was significantly less than zero, $p < .001$. The difference scores for D1 and D6 did not significantly differ from zero, $ps > .359$, see Fig. 2 for Cohen's *d*s. Together these findings suggest that Canadian participants primarily perceived D2, D3, and D5 as disgust, with D2 and D3 being the most effective in conveying this emotion. In contrast, D4 was predominantly associated

Fig. 3 Facial prototypes of anger and disgust of Asian and White targets generated by MetaHuman creator in Experiments 2 and 3

with anger rather than disgust, while D1 and D6 were perceived as a mixture of anger and disgust.

For Chinese participants, however, no pairwise comparisons were significantly different among the six disgust prototypes, $ps > .006$, Cohen's $ds < 0.29$. Furthermore, *t*-tests comparing the difference score for each facial prototype with zero indicated that the difference scores for D2, D3, and D5 were significantly greater than zero, $ps < .001$, while the difference scores for D1, D4, and D6 did not significantly differ from zero, $ps > .013$, see Fig. 2 for Cohen's *d*s. Together these findings suggest that Chinese participants primarily perceived D2, D3, and D5 as disgust and perceived D1, D4, and D6 as a mixture of anger and disgust.

We also decomposed the two-way interaction between Perceiver Culture and Facial Prototype by Facial Prototype. The significance threshold was Bonferroni adjusted to $p = .008$ (0.05/6) due to the comparisons between two cultures across six disgust prototypes. For D2, D3, and D5, the difference scores for Canadian participants were significantly higher than for Chinese participants, $ps < .006$, Cohen's $ds \geq 0.36$. For D4, the difference score for Canadian participants was significantly lower than that for Chinese participants, with $p < .001$, Cohen's $d = 0.75$. However, for D1 and D6, difference scores between the two cultural groups were not significantly different, $ps > .377$, Cohen's $ds < 0.09$. As depicted in Fig. 2, Chinese compared to Canadian participants perceived more mixed emotions for D2,

D3, D4, and D5, with difference scores closer to 0. However, the two cultural groups did not differ in the perception of mixed emotions for D1 and D6.

Discussion

In summary, the results of Experiment 1 suggest some prototypes of anger and disgust were more effective in conveying emotions than others, but this pattern varied across cultures. In general, the effect size of Facial Prototype was small for anger prototypes ($\eta_p^2 = 0.022$), but large for disgust prototypes ($\eta_p^2 = 0.180$). Perceiver Culture had a large effect size for anger prototypes ($\eta_p^2 = 0.180$) and a medium effect size for disgust prototypes ($\eta_p^2 = 0.065$). These results suggest that both facial prototypes and the cultural background of the perceivers play a significant role in emotion perception. The magnitude of their respective influences depends on the specific emotion being examined.

Specifically, among Canadian participants, all anger prototypes were predominantly perceived as anger, with A3 being the most effective in conveying this emotion. In contrast, Chinese participants did not identify any anger prototype as depicting more anger than disgust, with A1 perceived as more disgust than anger and the remaining 3 prototypes perceived as a blend of anger and disgust. Regarding the disgust prototypes, Canadian participants primarily recognized D2, D3, and D5 as disgust, with D2 and D3 being the most effective in communicating this emotion. They perceived D4 as more anger than disgust, while they perceived and D1 and D6 as a mixt of anger and disgust. Similarly, Chinese participants primarily identified D2, D3, and D5 as disgust, while interpreting D1, D4, and D6 as a mix of anger and disgust. Furthermore, our findings suggest that individuals from distinct cultural backgrounds discern varying degrees of mixed emotions from identical prototypes, with Chinese compared to Canadian participants detecting more mixed emotions in all anger prototypes and in four disgust prototypes (D2, D3, D4, and D5).

Experiment 2

As depicted in Fig. 1, the emotional faces generated with FaceGen may not accurately reflect the nuances of natural facial expressions, particularly with regards to muscle movements (e.g., wrinkles). Therefore, it is vital to investigate whether the findings of Experiment 1 replicate with more naturalistic emotional facial stimuli. While all faces generated using FaceGen were bald, male targets in Experiment 1 (as explained in footnote 1), we created more authentic and lifelike emotional faces of both male and female targets with MetaHuman in Experiment 2. The primary objective of this study was to explore how participants from Eastern and Western cultures perceive emotions from four anger prototypes and six disgust prototypes with these alternative set of stimuli.

Method

Participants

The final sample in this experiment included 96 (19 men and 77 women; $M_{age} = 19.75$, $SD = 6.45$) participants born and raised in Canada and recruited from a university in Ontario and 102 (21 men and 81 women; $M_{age} = 19.06$, $SD = 0.91$) participants born and raised in China and recruited from a university in Zhejiang Province. Canadian and

Chinese participants were compensated with either course credits or money, respectively. A sensitivity analysis using G*Power indicated that our sample could detect a small effect size, Cohen's $d = 0.20$, for differences between anger and disgust judgments (with $\alpha = .05$).

Stimuli and Procedure

Using 4 Asian models (two male and two female) and 4 White models (two men and two women) from MetaHuman, we created 10 facial prototypes (four related to anger and six related to disgust) for each of the eight models, resulting in a total of 80 faces, see Fig. 3. Although in Experiment 2, we used MetaHuman rather than FaceGen faces, the rest of the procedure was identical to Experiment 1.

Results

In accordance with the analytic strategy in Experiment 1, we conducted a 2 (Perceiver Culture: Canadian, Chinese) × 2 (Target Race: White, Asian) × 4 (Facial Prototype: A1, A2, A3, A4) mixed-design ANOVA for prototypes of anger, and a 2 (Perceiver Culture: Canadian, Chinese) × 2 (Target Race: White, Asian) × 6 (Facial Prototype: D1, D2, D3, D4, D5, D6) mixed-design ANOVA for prototypes of disgust. In contrast to Experiment 1, Target Race and Facial Prototype were within-subject variables, whereas Perceiver Culture was a between-subjects variable, see Table 2.

Facial Prototypes of Anger

In our first set of analyses that examined anger prototypes, the two-way interaction between Perceiver Culture and Facial Prototype was significant, see Table 2. To decompose this effect, we first analyzed the impact of Facial Prototype for Canadian and Chinese cultures separately, see Fig. 4. Pairwise comparisons, with an adjusted significance threshold of $p = .008$, as in Experiment 1, indicated that for Canadian participants, the difference score for A2 and A3 (with no significant difference between the two, $p = .632$, Cohen's $d = 0.05$) was significantly higher than for A1 and A4, $ps < .001$, Cohen's $ds > 0.59$. In addition, the difference score for A1 was significantly greater than for A4, $p < .001$, Cohen's $d = 0.67$. Moreover, *t*-tests comparing the difference score for each facial prototype with zero, with an adjusted significance threshold of $p = .013$, revealed that the difference scores for A1, A2, and A3 were greater than zero, $ps < .001$, while A4 did not significantly differ from zero, $p = .837$, see Fig. 4 for Cohen's *ds*. Together these findings suggest that A1, A2, and A3 were primarily perceived as anger by Canadian participants, with A2 and A3 being more effective than A1 in conveying this emotion. A4, alternatively was primarily perceived as a mixture of anger and disgust.

For Chinese participants, pairwise comparisons indicated that although the difference score for A3 was greater than for A4, $p < .001$, Cohen's $d = 0.38$, the difference scores when comparing any other anger prototypes were not significant, $ps > .007$, Cohen's $ds \leq 0.31$. Moreover, *t*-tests comparing the difference score for each facial prototype with zero revealed that the difference scores for A1 and A4 were significantly less than zero, $ps < .001$, indicating that these prototypes were perceived as disgust rather than anger. In contrast, the difference scores for A2 and A3 were not significantly different from zero, $ps > .140$, suggesting that these prototypes were primarily perceived as a mixture of anger and disgust, see Fig. 4 for Cohen's *ds*.

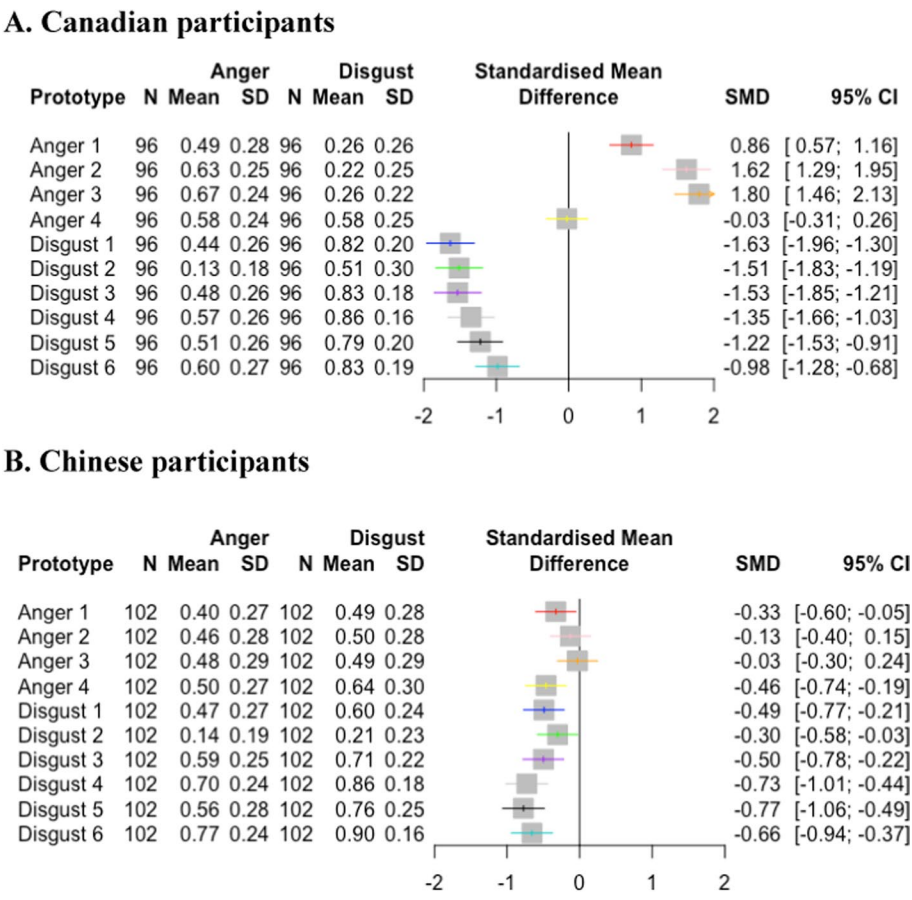

Fig. 4 Differences between anger and disgust scores for facial prototypes generated by MetaHuman in Experiment 2. *Note*. Positive scores related to standardized mean differences (SMD, Cohen's *d*) reflect higher ratings of anger than disgust while negative scores of SMD reflect higher ratings of disgust than anger. For further details on the results related to targets of different races, see Figure S2

We also decomposed the two-way interaction between Perceiver Culture and Facial Prototype by Facial Prototype. Pairwise comparisons adjusted to a significance threshold of $p = .013$ indicated that for A1, A2, and A3, the difference scores for Canadian compared to Chinese participants were significantly higher, $ps < .001$, Cohen's $ds \geq 1.04$. There were no cultural differences, however, for A4, $p = .025$, Cohen's $d = 0.44$. As depicted in Fig. 4, Chinese compared to Canadian participants perceived more mixed emotions for A1, A2, and A3, with difference scores closer to 0. However, the two cultural groups did not differ in the perception of mixed emotions for A4.

Facial Prototypes of Disgust

In the second set of analyses that examined disgust prototypes, the two-way interaction between Perceiver Culture and Facial Prototype was also significant, see Table 2. To decompose this effect, we analyzed again the impact of Facial Prototype for Canadian and

Chinese cultures separately, see Fig. 4. Pairwise comparisons with significance thresholds adjusted to $p = .003$ indicated that for Canadian participants, the difference scores for D1, D2, and D3 were significantly higher than for D6, $ps < .001$, Cohen's $ds \geq 0.36$. The difference scores between all other pairs of disgust prototypes, however, were not significant, $ps > .004$, Cohen's $ds \leq 0.30$. Furthermore, *t*-tests comparing the difference score for each facial prototype with zero, with corrected significance thresholds adjusted to $p = .008$, indicated that the difference scores for all disgust prototypes were significantly greater than zero, $ps < .001$, see Fig. 4 for Cohen's *ds*. Together these findings suggest that Canadian participants perceived all disgust prototypes as disgust, with D1, D2, and D3 being more effective than D6 in conveying this emotion.

For Chinese participants, pairwise comparisons revealed that the difference score for D5 was significantly higher than for D2, $p < .001$, Cohen's $d = 0.45$, but the difference scores between all other pairs of disgust prototypes were not significant, $ps > .003$, Cohen's $ds \leq 0.30$. Furthermore, *t*-tests comparing the difference score for each facial prototype with zero indicated that the difference scores for all disgust prototypes were significantly greater than zero, $ps < .001$, see Fig. 4 for Cohen's *ds*. Together these findings suggest that Chinese participants also perceived all disgust prototypes as disgust, with D5 being more effective than D2 in conveying this emotion.

We also decomposed the two-way interaction between Perceiver Culture and Facial Prototype by Facial Prototype. Pairwise comparisons with significance thresholds adjusted to $p = .008$ indicated that for D1, D2, D3, and D4, the difference scores for Canadian compared to Chinese participants were significantly higher, $ps < .006$, Cohen's $ds > 0.50$. No difference, however, was found between the two cultural groups for D5 and D6, $ps > .206$, Cohen's $ds \leq 0.33$. As depicted in Fig. 4, Chinese compared to Canadian participants showed a higher perception of mixed for D1, D2, D3, and D4, with difference scores closer to 0. The perceived level of mixed emotions for D5 and D6, however, was similar between the two cultural groups.

Discussion

In summary, the findings of Experiment 2 indicate that prototypes of anger and disgust differ in the extent to which they convey the intended emotions and that the results vary across cultures. In general, the effect size of Facial Prototype was large for anger prototypes ($\eta_p^2 = 0.249$), but small for disgust prototypes ($\eta_p^2 = 0.013$). However, Perceiver Culture had a large effect size for both anger ($\eta_p^2 = 0.339$) and disgust ($\eta_p^2 = 0.175$) prototypes. These results, consistent with those of Experiment 1, highlight the significant roles of facial prototypes and the cultural background of perceivers in emotion perception, with their respective influences varying depending on the specific emotion being examined.

Specifically, in line with the findings of Experiment 1, Canadian participants primarily perceived A1, A2, and A3 as anger, and A4 as a combination of anger and disgust, whereas Chinese participants perceived A1 and A4 as disgust, and A2 and A3 as a mixture of anger and disgust. However, in contrast to Experiment 1, both Canadian and Chinese participants perceived all disgust prototypes as disgust. Moreover, Chinese compared to Canadian participants perceived a higher level of mixed emotions from three anger (except for A4) and four disgust (except for D5 and D6) prototypes.

To further examine the impact of different software programs on emotion perception, we conducted a combined analysis of Experiments 1 and 2. Notably, Target Race was treated as a between-subjects factor in Experiment 1, whereas in Experiment 2, it was treated as a

within-subject factor. We thus merged the Target Race factor when performing a combined analysis for Experiments 1 and 2. The analysis revealed a significant main effect of Software Program, as indicated in Table S1. Notably, the effect size of Software Program was small for anger prototypes ($\eta_p^2 = 0.020$) but big for disgust prototypes ($\eta_p^2 = 0.204$).

Experiment 3

In both Experiments 1 and 2, we found that some anger or disgust prototypes were perceived as mixed emotions of anger and disgust, rather than solely anger or disgust. To further investigate this phenomenon, in Experiment 3, we instructed participants to judge the intensity of anger and disgust emotions, rather than making categorical judgments. Specifically, we recruited a new group of participants from Canada and China who were presented with the same set of MH emotional faces used in Experiment 2 and this new evaluation task.

Method

Participants

The experiment included 98 (13 men and 85 women; $M_{age} = 19.31$, $SD = 3.01$) participants born and raised in Canada and recruited from a university in Ontario and 101 (16 men and 85 women; $M_{age} = 18.79$, $SD = 0.92$) participants born and raised in China and recruited from a university in Zhejiang Province. Participants from Canada and China were compensated with either course credit or money, respectively. A sensitivity analysis using G*Power indicated that our final sample could detect a small effect size of Cohen's $d = 0.20$ for the differences between anger and disgust judgments (with $\alpha = .05$).

Stimuli and Procedure

The stimuli and procedure were identical to those of Experiment 2 with one exception. Participants were presented with each face once and instructed to rate the extent to which the face displayed anger and disgust on a scale from 0 (*not at all*) to 100 (*extremely*) with the order of ratings of anger and disgust counterbalanced across participants. Each participant completed a total of 80 trials, with the order of faces randomized.

Results

We conducted a 2 (Perceiver Culture: Canadian, Chinese) × 2 (Target Race: White, Asian) × 4 (Facial Prototype: A1, A2, A3, A4) mixed-design ANOVA for prototypes of anger, and a 2 (Perceiver Culture: Canadian, Chinese) × 2 (Target Race: White, Asian) × 6 (Facial Prototype: D1, D2, D3, D4, D5, D6) mixed-design ANOVA for prototypes of disgust. For both analyses, Target Race and Facial Prototype were within-subject variables and Perceiver Culture was a between-subjects variable, see Table 2.

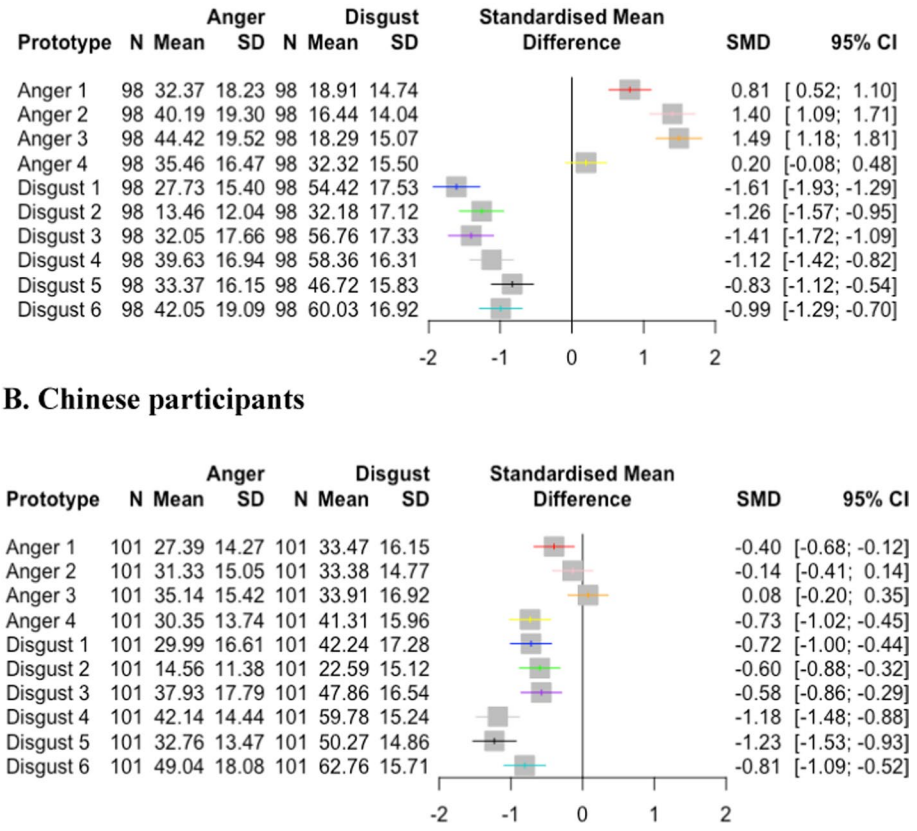

Fig. 5 Differences between anger and disgust scores for facial prototypes generated by MetaHuman in Experiment 3. *Note*. Positive scores related to standardized mean differences (SMD, Cohen's *d*) reflect higher ratings of anger than disgust while negative scores of SMD reflect higher ratings of disgust than anger. For further details on the results related to targets of different races, see Figure S3

Facial Prototypes of Anger

In the first set of analyses related to prototypes of anger, the two-way interaction between Perceiver Culture and Facial Prototype was significant, see Table 2. To decompose this effect, we first analyzed the impact of Facial Prototype for Canadian and Chinese cultures separately, see Fig. 5. Pairwise comparisons with significance thresholds adjusted to $p = .008$ revealed that for Canadian participants, the difference score for A2 and A3 (with no difference between the two, $p = .632$, Cohen's $d = 0.14$) was significantly higher for A1 and A4, $ps < .001$, Cohen's $ds \geq 0.76$. Furthermore, the difference score for A1 was higher than A4, $p < .001$, Cohen's $d = 0.62$. *T*-tests comparing the difference score for each facial prototype with zero, with significance thresholds adjusted to $p = .013$, indicated that the difference scores for A1, A2, and A3 were significantly greater than zero, $ps < .001$. A4, however, was not significantly different from zero, $p = .084$, see Fig. 5 for Cohen's *ds*. Together these results suggest that for Canadian participants, A1, A2, and A3 were perceived as

anger, with A2 and A3 being more effective than A1 in conveying this emotion. In contrast, A4 was primarily perceived as a mixture of anger and disgust.

Although for Chinese participants, no significant pairwise comparisons were found between difference scores for A2 and A3, $p = .012$, Cohen's $d = 0.26$, the difference score for A2 was higher than for A1, $p = .002$, Cohen's $d = 0.32$, and the difference score for A1 was higher than for A4, $p = .003$, Cohen's $d = 0.31$. In addition, *t*-tests comparing the difference score for each facial prototype with zero indicated that the difference scores for A1 and A4 were significantly less than zero, $ps < .001$, and the difference scores for A2 and A3 were not significantly different from zero, $ps \geq .254$, see Fig. 5 for Cohen's *ds*. Together these findings suggest that for Chinese participants, A1 and A4 were perceived as disgust, while A2 and A3 were primarily perceived as a mixture of anger and disgust.

We also decomposed the two-way interaction between Perceiver Culture and Facial Prototype by Facial Prototype. For all anger prototypes except A4, the difference scores for Canadian compared to Chinese participants were significantly higher, $ps < .001$, Cohen's $ds \geq 0.86$. As depicted in Fig. 5, while Chinese compared to Canadian participants showed a higher perception of mixed emotions for A1, A2, and A3, with difference scores closer to 0. However, Canadian compared to Chinese participants identified more mixed emotions with A4, as indicated by the difference score approaching 0.

Facial Prototypes of Disgust

In our second set of analyses related to disgust prototypes, the two-way interaction between Perceiver Culture and Facial Prototype was significant, see Table 2. To initially decompose this effect, we analyzed the impact of Facial Prototype for Canadian and Chinese cultures separately, see Fig. 5. Pairwise comparisons with significance thresholds adjusted to $p = .003$ revealed that for Canadian participants, although the difference score of D1 was not significantly different from D3, $p = .204$, Cohen's $d = 0.13$, it was significantly greater than D2, D4, D5, and D6, $ps < .001$, Cohen's $ds \geq 0.37$. The difference score of D3 was also significantly higher than D5, $p < .001$, Cohen's $d = 0.47$. No differences were observed in the other comparisons, $ps \geq .007$, Cohen's $ds \leq 0.28$. *T*-tests comparing the difference score for each facial prototype with zero, with significance thresholds adjusted to $p = .008$, demonstrated that the difference scores for all disgust prototypes were considerably higher than zero, $ps < .001$, see Fig. 5 for Cohen's *ds*. Together these findings suggest that for Canadian participants all disgust prototypes were perceived as disgust, with both D1 and D3 conveying disgust relatively stronger.

For Chinese participants, however, pairwise comparisons indicated that the difference score of D4 was not significantly different from D5 and D6, $ps > .006$, Cohen's $ds \leq 0.28$, but it was significantly greater than D1, D2, and D3, $ps < .001$, Cohen's $ds \geq 0.34$. Additionally, the difference score of D5 was significantly greater than D2 and D3, $ps < .001$, Cohen's $ds \geq 0.38$. No significant differences were found in the other comparisons, $ps \geq .003$, Cohen's $ds < 0.30$. Furthermore, *t*-tests comparing the difference score for each facial prototype with zero indicated that the difference scores for all six prototypes were significantly greater than zero, $ps < .001$, see Fig. 5 for Cohen's *ds*. Together these findings suggest that for Chinese participants all disgust prototypes were perceived as disgust, with D4, D5, and D6 conveying this emotion more effectively than D1, D2, and D3.

We also decomposed the two-way interaction between Perceiver Culture and Facial Prototype by Facial Prototype. Pairwise comparisons indicated that for D1, D2, and D3, the difference scores for Canadian compared to Chinese participants were significantly higher,

$ps < .001$, Cohen's $ds \geq 0.67$. For D4, D5 and D6, however, no differences were found between the two cultural groups, $ps > .074$, Cohen's $ds \leq 0.25$. As depicted in Fig. 5, Chinese compared to Canadian participants showed smaller difference scores for D1, D2, and D3, indicating a higher perception of mixed emotions for these disgust prototypes. The perceived level of mixed emotions for D4, D5, and D6, however, was similar between the two cultural groups.

Discussion

In summary, consistent with the findings of previous two experiments, the current experiment indicates that prototypes of anger and disgust differ in the extent to which they convey the intended emotions and that the results vary across cultures. In general, the effect size of Facial Prototype was large for anger prototypes ($\eta_p^2 = 0.335$), but small for disgust prototypes ($\eta_p^2 = 0.025$). Perceiver Culture had a large effect size for anger prototypes ($\eta_p^2 = 0.339$), but a small effect size for disgust prototypes ($\eta_p^2 = 0.059$). These results, in line with Experiments 1 and 2, underscore the significant roles of facial prototypes and the cultural background of perceivers in emotion perception, with their respective influences varying depending on the specific emotion under examination.

Specifically, Canadian participants primarily perceived A1, A2, and A3 as anger and A4 as a mixture of anger and disgust, while Chinese participants perceived A1 and A4 as disgust and A2 and A3 as a mixture of anger and disgust. Notably, both Canadian and Chinese participants perceived all disgust prototypes as disgust. Furthermore, Chinese participants perceived a higher level of mixed emotions for all anger prototypes except A4, and they also perceived a higher level of mixed emotions from D1, D2, and D3 compared to Canadian participants. In contrast, Chinese and Canadian participants did not show any cross-cultural differences in the perception of mixed emotions for D4, D5, and D6. Comparing the results of Experiments 2 and 3 suggest that despite changes in the evaluation paradigms from forced choice to ratings scales, the pattern of effects remain largely consistent to the previous findings related to MetaHuman stimuli.

General Discussion

In the field of emotion perception, researchers often use a variety of facial prototypes and materials created with different software programs without distinguishing between them. In the present research, we investigated the impact of standard prototypes and software often used to create emotional expressions on the efficacy of decoding these signals, while also considering cultural differences. In examining anger and disgust prototypes in depth, we found that decoding emotions differed depending on the specific prototypes used, the software programs employed, and the cultural background of the perceivers. These findings are consistent with previous research (Fang et al., 2018, 2019; Jack et al., 2009) and highlight the significant and relatively substantial influence of perceivers' cultural background on emotion perception. However, the magnitude of the effects of facial prototypes and software programs on emotion perception depended on the specific emotion being investigated or the prototypes being employed. These results emphasize the crucial importance of carefully selecting facial prototypes and software programs in emotion perception research, particularly in cross-cultural contexts.

Do Different Prototypes Convey the Same Emotional Signals?

When examining anger prototypes, Canadian participants primarily perceived the intended emotion in all prototypes generated by FaceGen, as well as in A1, A2, and A3 prototypes generated by MetaHuman, but a blend of anger and disgust in A4 on MH faces. In contrast, Chinese participants primarily perceived disgust rather than anger in A1 on both FG and MH faces, as well as in A4 on MH faces, but a blend of anger and disgust in all other anger prototypes. Further examination of the differences between A4 and the other anger prototypes revealed that lips apart (AU25) was the main distinguishing feature, which was present in all disgust prototypes. Therefore, it is possible that due to the frequent occurrence of the lips apart in disgust prototypes, participants were able to perceive a higher level of disgust in the anger prototype, A4.

When examining disgust prototypes, both Canadian and Chinese participants primarily perceived disgust in all disgust prototypes on MH faces. However, on FG faces, Canadian participants primarily perceived disgust in D2, D3, and D5, anger rather than disgust in D4, and a blend of anger and disgust in D1 and D6. Although Chinese participants also primarily perceived disgust in D2, D3, and D5, and a blend of anger and disgust in D1and D6, they differed from Canadian participants in perceiving a blended emotion (not anger) in D4. Notably, perceptual differences between the disgust prototypes created with the two software primarily stem from D1, D4, and D6, which all involve the nose wrinkler (AU9).

In comparing the nose wrinkler (AU9) in the two software programs, the resulting changes in FaceGen include a lowered brow, narrowed eyes, and a horizontal wrinkle across the bridge of the nose, see D1 in Fig. 1. These changes are very similar to facial changes caused by the two key AUs in anger, namely the brow lowerer (AU4) and lid tightener (AU7), see A1 in Fig. 1. Therefore, when FaceGen is used to create disgust prototypes involving the nose wrinkler (AU9), participants are more likely to perceive a blend of anger and disgust emotions. In contrast, in MetaHuman the resulting changes when manipulating the nose wrinkles (AU9) include a slight lowering of the brow, narrowed eyes, several vertical wrinkles on the bridge of the nose, and deepened nasolabial lines, see D1 in Fig. 3, which are different from the changes caused by the brow lowerer (AU4) and lid tightener (AU7), see A1 in Fig. 3. Therefore, disgust prototypes created with MetaHuman software are primarily perceived as conveying disgust.

Furthermore, across all three experiments, A3 was found to be the most effective in conveying anger for Canadian participants, as it included both upper and lower facial actions, namely the brow lowerer (AU4), lid tightener (AU7), chin raiser (AU17), and lip tightener (AU23). The addition of lower facial actions appears to enhance the perception of anger compared to upper facial actions alone. Our findings suggest that in future studies seeking to create facial expression prototypes that effectively convey anger, A3 should be considered. This could potentially improve the accuracy in identification of facial expressions.

With regard to disgust prototypes, for Canadian participants, the disgust prototype that was most effective was D3, which is comprised of the nose wrinkler (AU9), upper lip raiser (AU10), and lips apart (AU25). This prototype contains the two key action units, AU 9 and AU 10, that are essential for conveying disgust. However, for Chinese participants, the most effective disgust prototype was D5, which is comprised of the brow lowerer (AU4), lid tightener (AU7), upper lip raiser (AU10), and lips apart (AU25). These findings suggest that for Chinese participants, disgust conveyed through the AU10 was more effective than the AU9.

The Impact of Culture on Mixed Emotion Perception

Across three experiments, we demonstrated that overall Chinese participants perceived a higher degree of mixed emotions in facial expressions than Canadian participants. This finding is consistent with previous research indicating that participants from Eastern cultures such as China compared to participants from Western cultures such as the Netherlands perceived higher levels of mixed emotions in both FACS-based (Western) prototypes of emotional expressions and empirically based facial expressions from Eastern and Western cultures (Fang et al., 2019). As noted in the introduction, one possible explanation for this pattern of results is that participants from Eastern cultures, which are characterized by holistic and dialectical thinking, may be more inclined to perceive mixed emotions in facial expressions than participants from Western cultures, which are characterized by analytical and nondialectical thinking (Nisbett et al., 2001; Peng & Nisbett, 1999; Spencer-Rodgers et al., 2004).

East Asians' inclination towards mixed emotions has been found not only in the perception of facial expressions but also in the production of facial expressions (Fang et al., 2022; Wood et al., 2016). Using machine learning to categorize expressions based on patterns of facial movements, Fang et al. (2022) found that both posed and spontaneous expressions of anger and disgust by Chinese participants were less distinct than Dutch participants. This finding suggests that, like the perception of emotional facial expressions, East Asians also produce emotional expressions that overlap more than Westerners.

It is worth noting that while, overall, Chinese participants were found to perceive more mixed emotions from facial expressions compared to Canadian participants, this phenomenon did not hold true for every prototype. For example, both Canadian and Chinese participants primarily perceived disgust from D5 and D6 on MH faces, and the degree of mixed emotions perceived was similar for both groups. Therefore, if D5 or D6 were used as experimental stimuli, there may be no difference in recognition rates between East Asian and Western participants. Conversely, if D3 were used as the experimental material, Western participants may be more likely to perceive the expression as disgust compared to East Asian participants. These findings suggest that researchers should exercise caution when selecting specific facial prototypes of emotions and consider how these perceptions may vary across cultures.

Implications

The present findings contribute in important ways to research on emotion perception in general as well as across cultures. Standardized facial expressions of emotion have been widely utilized in measures of emotional intelligence in various settings such as workplaces and classrooms (Mayer et al., 1999; Zeidner et al., 2004), in investigations of children's comprehension of others' emotions and intentions (Camras & Allison, 1985; Gross & Ballif, 1991; Widen & Russell, 2003), in diagnostic tools for mental disorders including anxiety disorders and borderline personality disorder (Leible & Snell Jr., 2004; Lizeretti & Extremera, 2011), and of course in understanding how individuals decode emotions (Ekman & Cordaro, 2011). The present findings suggest that different prototypes associated with these standardized expressions do not necessarily convey emotions equally, and the selection of specific prototypes and software to create prototypes can significantly impact an individual's ability to recognize these emotions. Therefore, it is crucial for researchers

and practitioners to be aware of the potential cross-cultural variations in emotion perception and consider the influence of specific prototypes on the effectiveness of measures and interventions related to emotion recognition.

Limitations and Future Directions

The present research has several limitations that should be acknowledged. First, in the current experiments, we did not include additional filler facial expressions, such as sadness and fear, or provide more emotion options. These procedural decisions may have constrained participants' perception of the stimuli to anger and disgust exclusively. Future research should, therefore incorporate additional filler expressions and expand response options, allowing participants to choose from a wider range of expressions. By doing so, we can examine whether our results hold true under a less restrictive format.

Second, while it was challenging to precisely control each action unit of the recruited models, we synthesized a series of anger and disgust prototypes using software to conduct our investigation. Therefore, we acknowledge that uncertainty exists regarding whether these findings are applicable to real human faces. Nonetheless, comparisons between MetaHuman and FaceGen faces in Figs. 1 and 3 indicate a significant improvement in the realism and simulation of facial muscle movements achieved by MetaHuman. A productive avenue for future research is to recruit professional actors to create corresponding expression prototypes, using Photoshop techniques to meet experimental material standards.

Third, this study relied solely on static facial expressions. Although static expressions are commonly used in emotion research, they still differ from the dynamic expressions we encounter in real-life situations. Dynamic facial expressions contain more information, such as the onset, peak, and offset of facial expressions (Krumhuber et al., 2013). This information may assist individuals in better differentiating between various emotions, thereby facilitating the perception of more specific emotions from prototypes. Future research could further explore this possibility.

Conclusion

Although a large body of research has examined perceptions of standard facial expressions, little is known about how different prototypes related to a single emotion impact these perceptions and whether the prototypes generated by different software are decoded differently. Even less is known about how these prototypes are perceived across cultures. The present research provides new evidence on differences in the signal value of four major prototypes of anger and six major prototypes of disgust generated by FaceGen and MetaHuman, as well as cultural differences in the extent to which specific prototypes are successful at depicting a single emotion whether anger or disgust. Together these findings suggest that researchers should exercise caution when choosing specific facial prototypes of emotions and the facial generation software for their research and understand the impact of these choices within and across cultures.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10919-024-00453-0.

Acknowledgements The authors thank Siqi Liu and Feixuan Du for their valuable contributions in the production of experimental materials and collection of experimental data.

Author Contributions XF and KK developed the study concept and design. XF collected data. XF performed the data analysis and interpretation. XF drafted the manuscript, and KK provided critical revisions. Both authors approved the final version of the manuscript for submission.

Funding This work was supported by the Social Sciences and Humanities Research Council of Canada (435-2018-0907).

Declarations

Conflict of interest The authors declare no competing interests.

References

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724–732. https://doi.org/10.1111/j.1467-9280.2008.02148.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

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9*(2), 84–94. https://doi.org/10.1007/BF00987140

Chen, L. F., & Yen, Y. S. (2007). *Taiwanese Facial Expression Image Database*. Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan.

Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion, 18*, 75–93. https://doi.org/10.1037/emo0000302

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. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124–129. https://doi.org/10.1037/h0030377

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior, 1*(1), 56–75. https://doi.org/10.1007/BF01115465

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press.

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

Ekman, P., Friesen, W. V., O'sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology, 53*(4), 712–717. https://doi.org/10.1037/0022-3514.53.4.712

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*(3875), 86–88. https://doi.org/10.1126/science.164.3875.86

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*, 131–146. https://doi.org/10.1037/1528-3542.7.1.131

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

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

Friesen, J. P., Kawakami, K., Vingilis-Jaremko, L., Caprara, R., Sidhu, D. M., Williams, A., Hugenberg, K., Rodríguez-Bailón, R., Cañadas, E., & Niedenthal, P. (2019). Perceiving happiness in an intergroup context: The role of race and attention to the eyes in differentiating between true and false smiles. *Journal of Personality and Social Psychology, 116*(3), 375–395. https://doi.org/10.1037/pspa0000139

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

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191.

Giner-Sorolla, R., Kupfer, T., & Sabo, J. (2018). What makes moral disgust special? An integrative functional review. In *Advances in experimental social psychology* (Vol. 57, pp. 223–289). Academic Press.

Gosselin, P., & Kirouac, G. (1995). Decoding facial emotional prototypes. *Canadian Journal of Experimental Psychology, 49*(3), 313–329.

Gross, A. L., & Ballif, B. (1991). Children's understanding of emotion from facial expressions and situations: A review. *Developmental Review, 11*(4), 368–398. https://doi.org/10.1016/0273-2297(91)90019-K

Izard, C. E. (1971). *The face of emotion*. Appleton-Century-Crofts.

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

Ji, L.-J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology, 78*(5), 943–955. https://doi.org/10.1037//0O22-3514.78.S.943

Kollareth, D., & Russell, J. A. (2017). The English word disgust has no exact translation in Hindi or Malayalam. *Cognition and Emotion, 31*(6), 1169–1180. https://doi.org/10.1080/02699931.2016.1202200

Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5*(1), 41–46.

Lizeretti, N. P., & Extremera, N. (2011). Emotional intelligence and clinical symptoms in outpatients with generalized anxiety disorder (GAD). *Psychiatric Quarterly, 82*(3), 253–260. https://doi.org/10.1007/s11126-011-9167-1

Leible, T. L., & Snell, W. E., Jr. (2004). Borderline personality disorder and multiple aspects of emotional intelligence. *Personality and Individual Differences, 37*(2), 393–404. https://doi.org/10.1016/j.paid.2003.09.011

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*(4), 267–298. https://doi.org/10.1016/S0160-2896(99)00016-1

Matsumoto, D., Anguas-Wong, A. M., & Martinez, E. (2008). Priming effects of language on emotion judgments in Spanish—English bilinguals. *Journal of Cross-Cultural Psychology, 39*(3), 335–342. https://doi.org/10.1177/0022022108315489

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13*(2), 143–157. https://doi.org/10.1007/BF00992959

Matsumoto, D., & Hwang, H. C. (2017). Methodological issues regarding cross-cultural studies of judgments of facial expressions. *Emotion Review, 9*(4), 375–382. https://doi.org/10.1177/1754073916679008

Malek, N., Messinger, D., Gao, A. Y. L., Krumhuber, E., Mattson, W., Joober, R., ... & Martinez-Trujillo, J. C. (2019). Generalizing Duchenne to sad expressions with binocular rivalry and perception ratings. *Emotion, 19*(2), 234.

Mende-Siedlecki, P., Qu-Lee, J., Backer, R., & Van Bavel, J. J. (2019). Perceptual contributions to racial bias in pain recognition. *Journal of Experimental Psychology: General, 148*(5), 863–889. https://doi.org/10.1037/xge0000600

Miller, E. J., Krumhuber, E. G., & Dawel, A. (2022). Observers perceive the Duchenne marker as signaling only intensity for sad expressions, not genuine emotion. *Emotion, 22*(5), 907.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review, 5*(1), 8–15. https://doi.org/10.1177/1754073912457227

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291–230. https://doi.org/10.1037/0033-295X.108.2.291

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087–11092.

Peng, K., & Nisbett, R. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist, 54*(9), 741–754. https://doi.org/10.1037/0003-066X.54.9.741

Peng, K., Spencer-Rodgers, J., & Nian, Z. (2006). Naïve dialecticism and the Tao of Chinese thought. In U. Kim, K. S. Yang, & G. Huang (Eds.), *The handbook of indigenous and cultural psychology* (pp. 247–262). Blackwell.

Pochedly, J. T., Widen, S. C., & Russell, J. A. (2012). What emotion does the "facial expression of disgust" express? *Emotion, 12*(6), 1315. https://doi.org/10.1037/a0027998

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

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102. https://doi.org/10.1037/0033-2909/94/S3.00

Scherer, K. R. (1988). Criteria for emotion-antecedent appraisal: A review. In V. Hamilton, G. H. Bower, & N. H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 89–126). Dordrecht: Springer.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48*(4), 813–838.

Spencer-Rodgers, J., Peng, K., Wang, L., & Hou, Y. (2004). Dialectical self-esteem and East-West differences in psychological well-being. *Personality and Social Psychology Bulletin, 30*(11), 1416–1432.

Widen, S. C., Pochedly, J. T., Pieloch, K., & Russell, J. A. (2013). Introducing the sick face. *Motivation and Emotion, 37*(3), 550–557. https://doi.org/10.1007/s11031-013-9353-6

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 and Emotion, 22*(8), 1513–1541. https://doi.org/10.1080/02699930801906744

Widen, S. C., Russell, J. A., & Brooks, A. (2004, May). Anger and disgust: Discrete or overlapping categories. In *2004 APS annual convention, Boston College, Chicago, IL.*

Wood, A., Rychlowska, M., & Niedenthal, P. M. (2016). Heterogeneity of long-history migration predicts emotion recognition accuracy. *Emotion, 16*(4), 413–420. https://doi.org/10.1037/emo0000137

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

Yik, M., Widen, S. C., & Russell, J. A. (2013). The within-subjects design in the study of facial expressions. *Cognition and Emotion, 27*(6), 1062–1072. https://doi.org/10.1080/02699931.2013.763769

Yoder, A. M., Widen, S. C., & Russell, J. A. (2016). The word disgust may refer to more than one emotion. *Emotion, 16*(3), 301–308. https://doi.org/10.1037/emo0000118

Zeidner, M., Matthews, G., & Roberts, R. D. (2004). Emotional intelligence in the workplace: A critical review. *Applied Psychology, 53*(3), 371–399. https://doi.org/10.1111/j.1464-0597.2004.001