



# The India Face Set: International and Cultural Boundaries Impact Face Impressions and Perceptions of Category Membership

Anjana Lakshmi<sup>1\*</sup>, Bernd Wittenbrink<sup>2</sup>, Joshua Correll<sup>3</sup> and Debbie S. Ma<sup>4</sup>

<sup>1</sup> Department of Psychology, University of Chicago, Chicago, IL, United States, <sup>2</sup> Booth School of Business, University of Chicago, Chicago, IL, United States, <sup>3</sup> Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, United States, <sup>4</sup> Department of Psychology, California State University, Northridge, Los Angeles, CA, United States

## OPEN ACCESS

### Edited by:

Dmitry Grigoryev,  
National Research University Higher  
School of Economics, Russia

### Reviewed by:

Elena Tsankova,  
Institute for Population and Human  
Studies (BAS), Bulgaria  
Ferenc Kocsor,  
University of Pécs, Hungary

### \*Correspondence:

Anjana Lakshmi  
anjanachandran@uchicago.edu

### Specialty section:

This article was submitted to  
Personality and Social Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 09 November 2020

**Accepted:** 13 January 2021

**Published:** 11 February 2021

### Citation:

Lakshmi A, Wittenbrink B,  
Correll J and Ma DS (2021) The India  
Face Set: International and Cultural  
Boundaries Impact Face Impressions  
and Perceptions of Category  
Membership.  
Front. Psychol. 12:627678.  
doi: 10.3389/fpsyg.2021.627678

This paper serves three specific goals. First, it reports the development of an Indian Asian face set, to serve as a free resource for psychological research. Second, it examines whether the use of pre-tested U.S.-specific norms for stimulus selection or weighting may introduce experimental confounds in studies involving non-U.S. face stimuli and/or non-U.S. participants. Specifically, it examines whether subjective impressions of the face stimuli are culturally dependent, and the extent to which these impressions reflect social stereotypes and ingroup favoritism. Third, the paper investigates whether differences in face familiarity impact accuracy in identifying face ethnicity. To this end, face images drawn from volunteers in India as well as a subset of Caucasian face images from the Chicago Face Database were presented to Indian and U.S. participants, and rated on a range of measures, such as perceived attractiveness, warmth, and social status. Results show significant differences in the overall valence of ratings of ingroup and outgroup faces. In addition, the impression ratings show minor differentiation along two basic stereotype dimensions, competence and trustworthiness, but not warmth. We also find participants to show significantly greater accuracy in correctly identifying the ethnicity of ingroup faces, relative to outgroup faces. This effect is found to be mediated by ingroup-outgroup differences in perceived group typicality of the target faces. Implications for research on intergroup relations in a cross-cultural context are discussed.

**Keywords:** normed face stimuli, India and U.S., cultural differences, subjective impressions, stereotypes

## INTRODUCTION

It has been noted that psychology conducts its research largely on people from Western, educated, industrialized, rich, and democratic countries—coined WEIRD societies by Henrich et al. (2010). Social psychology, despite its focus on the importance of social context for psychological functioning, is no exception in this regard. Within the area of intergroup relations, studies on stereotyping, group attitudes, and intergroup behavior, have been conducted largely with participants from the United States and Western Europe, investigating how people perceive,

judge, and interact with social groups that are culturally relevant to these parts of the world. By comparison, studies with participants and/or target groups from non-WEIRD societies are few and far between (e.g., Jahoda, 1959; Kashima et al., 2003; Cuddy et al., 2009; Durante et al., 2017). The limited empirical scope raises questions about how research findings might generalize to other cultural contexts. And it leaves the field with missed opportunities for studying psychological determinants of intergroup relations.

Ironically, recent efforts to improve methodological practices in psychology (Kahneman, 2012; Asendorpf et al., 2013; Open Science Collaboration, 2017) carry some risk to further exacerbate this situation. For example, in order to improve experimental control and to facilitate comparisons across studies, researchers are encouraged to rely on standardized procedures and materials in their studies (Shrout and Rodgers, 2018). However, such standardization is likely to come at the expense of methodological diversity. A case in point is the Chicago Face Database (CFD; Ma et al., 2015), a collection of face images and norming data that our lab developed and made available as a free resource for use as stimulus materials in research.

The database provides easy access to face images that are uniform in terms of image quality, lighting, camera positioning, model pose, and other potentially confounding aspects of photographs. The face images come with extensive norming data that cover physical attributes (e.g., face height, width, luminance, etc.) as well as subjective impressions of the faces (e.g., perceived age, attractiveness, trustworthiness, etc.), allowing researchers to select images for particular face attributes while controlling for other factors that are extraneous to the research question. The database was inspired by the International Affective Picture System (IAPS; Lang et al., 1997), a stimulus database that has seen widespread use in research involving emotion and affect. Similar to the IAPS, the CFD was intended to facilitate and help standardize the broad variety of psychological research that involves the presentation of face stimuli to participants (e.g., impression formation, intergroup processes, stereotyping, prejudice, emotions). Since its release just 5 years ago, the database has seen rapid adoption, with more than 7,000 downloads and 700 published papers that report studies with CFD faces.

An explicit goal in developing the database was also to broaden the demographic composition of face images available to researchers. The existing image resources available include either exclusively Caucasian faces (Ekman and Friesen, 1976; Troje and Bühlhoff, 1996; Lundqvist et al., 1998), or only a relatively small number of non-Caucasian faces (Tottenham et al., 2009; Langner et al., 2010; DeBruine and Jones, 2017; see **Table 1** for a list of widely used image sets and their ethnic makeup). In contrast, the CFD now offers images and norming data for nearly 600 Asian, Black, Latino, and White males and females.

While the database makes it easier for researchers to include non-Caucasian faces in their studies, all CFD models were volunteers recruited in the U.S. As a result, the ethnic diversity represented in the database remains limited to a subset of U.S. ethnic social groups. And the composition of these groups reflects the obvious limitations of a convenience sample. For

instance, models of the database who self-identified as *Asian* are predominantly U.S.-born models with East Asian ancestry, covering only a portion of the ecological diversity of faces on the Asian continent. Likewise, the subjective norming data included in the database were collected with U.S. rater samples. They offer information on how attractive etc. the faces appear to U.S. participants, but raise the question whether these impressions might be different for perceivers of different cultural background and/or group identity.

While researchers have employed creative methods such as morphing and caricatures to generate additional non-Caucasian face stimuli from the limited number of available base faces (e.g., Byatt and Rhodes, 1998; Krumhuber et al., 2015), the focus of existing face databases on Caucasian faces has obvious methodological and conceptual implications. The reliance on U.S.-specific norms for stimulus selection may introduce experimental confounds if the norms don't generalize to non-U.S. participants. The use of face stimuli that insufficiently capture the ecological diversity of faces may adversely impact a study's external and internal validity (Wells and Windschitl, 1999) and yield incorrect effect estimates or fail to identify important moderators (Fiedler, 2011). In addition, the ready availability of certain ethnicities in the database may influence what target groups are being chosen for investigation in the first place, curtailing research on hypotheses for which materials aren't readily available.

## The Current Research

The research reported in this paper aims to address some of these issues and improve the usefulness of the database for work with non-U.S. participants and non-U.S. faces. It has three specific goals. First, we describe the development of an expansion to the

**TABLE 1** | Face image sets and their ethnic makeup.

Database	Number of models by ethnicity	
CFD-India	Indian Asian	142
	Asian	109
	Black	197
	Caucasian	183
	Latino	108
FACES (Ebner et al., 2010)	Caucasian	171
KDEF (Lundqvist et al., 1998)	Caucasian	70
Facelab London Set (DeBruine and Jones, 2017)	Asian	19
	Black	13
	Caucasian	69
	Multiethnic	1
	NimStim (Tottenham et al., 2009)	Asian
POFA (Ekman and Friesen, 1976)	Black	10
	Caucasian	25
	Latino	2
	Caucasian	14
	RaFD (Langner et al., 2010)	Caucasian
	Moroccan	18

CFD with face images of individuals recruited in India, drawing on a large non-U.S. ethnic group that accounts for approximately 18% of the world population (United Nations Department of Economic and Social Affairs Population Division, 2019). Second, we explore the extent to which subjective impressions of these faces are culturally dependent. And, third, we investigate whether differences in target face familiarity and perceived group typicality impact judgments of face ethnicity.

### The India Face Set

The new image set introduced here includes high resolution face images of 142 unique individuals, displaying a variety of facial expressions (neutral, angry, fearful, and happy). The images are standardized according to the procedures used for the CFD and, hence, can serve as stimuli side-by-side with the original U.S. face images. They are accompanied by comprehensive norming data. Beyond the physical face attributes and subjective impressions that are part of the CFD, these norms now also include self-reported background information on the models (e.g., ancestry, home state, religious affiliation, caste, and SES measures). All materials are available as a free resource at [www.chicagofaces.org](http://www.chicagofaces.org).

### Cultural Dependency of Subjective Image Norms

A second goal of the current research is to explore the extent to which the subjective rating norms are culturally dependent and the extent to which these ratings might differ for ingroup and outgroup faces. Although some studies have found impressions from faces to be consistent across culturally diverse rater samples (Wagatsuma and Kleinke, 1979; Bernstein et al., 1982; Cunningham et al., 1995) several recent studies have documented systematic cultural differences in what impressions perceivers glean from faces (Sutherland et al., 2018; Wang et al., 2019; Jones et al., 2021). Moreover, there are various theoretical arguments and related empirical findings that would suggest impressions for ingroup faces and outgroup faces to differ. For example, the mere exposure hypothesis (Zajonc, 1968) predicts that more familiar faces should be judged more positively. In fact, faces with feature sets near the population average are perceived to be more familiar (Langlois et al., 1994). And familiar faces, in turn, are judged as more likable (Zebrowitz et al., 2008), trustworthy (Lewicki, 1985), and attractive (Winkielman et al., 2006; Zebrowitz, 1996). To the extent that Indian and U.S. faces differ systematically in their feature sets, and that raters are relatively more familiar with their ingroup, one would expect ingroup faces to be viewed more positively. Theories of intergroup behavior, such as social identity theory (Tajfel et al., 1979), would similarly predict impressions to reflect ingroup favoritism, with impressions of ingroup faces to be more positive.

On the other hand, social stereotypes may also impact impressions of both ingroup and outgroup faces with regard to particular stereotypic attributes. For example, the stereotype content model suggests that groups viewed as competitors are perceived to be less warm, and groups of lower status as less competent (Fiske et al., 2002). With regard to the groups of interest to the current research, Lee and Fiske (2006) observed that U.S. participants' stereotypes of Indian Asian immigrants are similar in content and valence to the stereotypes U.S.

participants hold about their own ingroup. Also, though we are unaware of any direct data on this issue, a 2014 Pew Research Center Survey suggests that the majority of Indians hold favorable (58%) or very favorable (30%) views of the U.S. (Pew Research Center, 2014). Based on these data we might expect impression ratings to reflect mutual admiration, rather than ingroup favoritism.

To explore these possibilities, we collected subjective impression ratings in a full ingroup-outgroup design, with samples of Indian and U.S. participants each rating both Indian Asian and Caucasian face images on a variety of attributes (e.g., attractiveness, competence, etc.). The design allowed us to identify separate effects of participant and target group on face impressions, and test for evidence of stereotyping and ingroup/outgroup favoritism in these ratings.

### Judgments of Face Ethnicity

Finally, a third goal of the research was to determine whether differences in familiarity with Indian and Caucasian faces would impact participants' ability to identify face ethnicity. Across domains, stimulus familiarity has been found to impact processing efficiency (Posner and Keele, 1968; Lewellen et al., 1993) and categorization (e.g., Smith, 1967; Johnson and Mervis, 1997; Whittlesea and Leboe, 2000). In the case of faces, it has been suggested that familiar ingroup faces function as a perceptual default facilitating their processing and identification, while impeding the processing and identification of other-race faces (e.g., Goldstein and Chance, 1980; Rhodes et al., 1987; Macron et al., 2009). Hence we expected greater accuracy in judgments of familiar faces, with Indian Asian faces to be more likely classified as such by Indian raters than U.S. raters, whereas the opposite should hold for Caucasian faces.

## MATERIALS AND METHODS

### Face Stimuli

The present study used Caucasian and Indian Asian target faces as experimental stimuli. Caucasian face stimuli were randomly drawn from the existing pool of CFD images depicting Caucasian models from the U.S. (for a full list of target images, see the online **Supplementary Material**). Face stimuli for Indian Asian targets were collected at the University of Chicago Center in Delhi, India. Potential volunteers were contacted via convenience sampling, snowball sampling as well as pamphlets that were distributed to various cultural organizations with memberships from different regions in India. Volunteers were required to be between the ages of 18 and 50. Of the resultant volunteers, 53 were female and 91 were male. Self-report data about the volunteers' location within India (87 North Indian, 15 South Indian, 15 West Indian, 12 North East Indian, 7 Central Indian, 7 East Indian), religion (79 Hindu, 25 Muslim, 19 Sikh, 18 Christian, 1 Jain, 1 agnostic, 1 no religion), caste category, native language, education, employment and annual income were collected as was information about location of birth, current location of residence and ancestry.

## Photo Sessions

Upon arrival participants were each asked to carefully read an informed consent and image release form. The forms were made available in both English and Hindi, and upon request were translated on site to other Indian languages. For illiterate participants, the experimenter read aloud the consent instructions and probed for comprehension. Afterward, participants changed into a gray t-shirt (the same type of shirt worn by all models of the existing CFD image set). Next, at the participants' discretion, they removed any make-up and jewellery. If needed, they were encouraged to shave and adjust their hair so that it did not obstruct the face. We chose not to enforce compliance with these grooming preparations as they may have interfered with cultural practices. For example, some married women in India wear vermilion on the apex of their hairline and/or a traditional necklace. Tradition may prevent them from appearing in public without these signifiers of their married status. Likewise, men may grow a beard or wear a turban for religious purposes. In such instances, volunteers were photographed as is.

For the actual photo session, volunteers were then seated at a fixed distance from a digital camera. The technical setup for these sessions followed closely the procedures used for the existing CFD image set, described in detail in Ma et al. (2015). Volunteers were asked to make neutral, happy (with both open and closed mouth smile), angry and fearful expressions while also maintaining an upright and straight head position. Each volunteer completed three rounds of photographs. In the first round, they received a prompt (e.g., "make a closed mouth smile"), and when necessary, the photographer followed up with more specific directions (e.g., "Please try to engage your eyes in the smile"). The second and third round repeated the full cycle of facial expressions. Volunteers who struggled reaching credible expressions were offered illustrations taken from Ekman and Friesen (1976). This resulted in multiple photographs for each volunteer displaying each of the requested facial expressions. Sessions lasted approximately 30 min. At the conclusion, refreshments were provided and thereafter volunteers were thanked and compensated with Rs. 500.

## Image Standardization

From the resulting pool of images, we selected one neutral expression image per volunteer, based on head position (i.e., straight and upright) image quality (i.e., in focus), and how neutral the expression indeed was. Using these criteria, two (for a subset of targets, three) independent judges first rated each image and identified their top three face stimuli. Next, these top picks were used to settle on a consensual best choice for the final image selection<sup>1</sup>.

The selected images were edited using Adobe Photoshop software (version 20) following the standardization procedures described in Ma et al. (2015). RAW image files were corrected for uniform color temperature and exposure across

images, matching the existing CFD materials. Where necessary, additional corrections were made to reach a realistic skin tone. Next, we made digital modifications to select images, to remove any blemishes, markings or tattoos, facial or ear piercings, as well as any earrings, hair accessories and/or jewelry<sup>2</sup>. All images were then resized so that the size of the core facial features was more or less equivalent across all images and consistent with the existing CFD face stimuli. For this, a 796 pixels (wide) × 435 pixels (high) template was fit over the target's core facial features, adjusting the image size such that either the eyebrow-lip distance matched the template height, and/or the max. cheekbone distance matched the template width. Finally, a white background was inserted, and the image was exported to a 2,444 pixels by 1,718 pixels JPEG file (see **Figure 1** for sample images).

## Norming Data

The standardized neutral expression images serve as the basis for the norming data, which include both objective measures of physical face features and subjective ratings of face impressions. The latter are the focus of our question whether subjective image norms are culturally dependent. In contrast, the objective norms are part of the image set development and provide descriptive information on the physical attributes of the new face sample. We report them here in order to document the steps we took to capture the physical attributes of the India face set.

## Subjective Norms

Subjective ratings of the 284 target faces (142 Indian Asian, 142 Caucasian) with neutral facial expressions were obtained using two separate tasks (A and B) designed with Qualtrics Research Suite Software. Task A asked Indian and U.S. participants to rate the Indian Asian and Caucasian faces on a range of attributes. In task B a separate sample of Indian and U.S. participants was asked to rate the Indian Asian and Caucasian target faces for their group typicality. Participant recruitment and data collection for both tasks were conducted using Amazon's Mechanical Turk.

### Task A

In a within-subjects design, each participant was presented with 8 target faces—2 Caucasian male, 2 Caucasian female, 2 Indian Asian male, and 2 Indian Asian female. For each participant, these 8 faces were selected at random from the target pool, with no replacement until all of the target faces were judged once for that iteration. The entire task took approximately 15 min to complete; U.S. participants were compensated with \$3 and Indian participants with Rs. 100.

For each target, participants first saw the target pictured at the top of the computer screen followed by prompts below to estimate the target's age, race (with response options: Chinese Asian, Japanese Asian, Indian Asian, Other Asian, Black, Hispanic/Latino, White/Caucasian, and Other) and gender. Next,

<sup>2</sup>For targets with traditional signifiers, like the aforementioned vermilion head marking (sindoor), or a marriage necklace (mangal sutra), we prepared duplicate image versions where technically feasible. Both versions were processed in identical fashion, except version 1 removed the signifiers whereas version 2 kept them intact. Version 1 was used for the norming data collection. But as researchers may be interested in alternate versions of the same target, both are distributed with the CFD-India image set.

<sup>1</sup>Faces with emotional expressions are not of immediate interest to the current study. Their selection and standardization followed the procedures outlined in Ma et al. (2015).



**FIGURE 1** | Sample Stimuli from the India Face Set.

the target image remained, but the prompts were replaced, asking participants to rate their impression of the target on the following dimensions: attractive, warm, competent, trustworthy, happy, sad, disgusted, surprised, fearful/afraid, angry, threatening, masculine, feminine, baby-faced, and unusual (such that they would stand out in a crowd). For each target, these attributes were presented across two successive screens and the ordering of attributes within each screen was chosen at random. Participants responded with a Likert scale of 1 (Not at all) through 4 (Neutral) to 7 (Extremely). The next screen showed a prompt asking participants to characterize the social status of the target from 1 (Low) through 7 (High). To facilitate these ratings, the prompt was accompanied by the following explanation: People of high status are typically thought to be wealthy and well-educated, working in highly paid jobs whereas those who are of low status are thought to be poor and not well-educated (or not educated at all), typically working in low paid positions or unemployed (see Lakshmi et al., 2019). All items but for status, competence and warmth were drawn from Ma et al. (2015). Status, warmth and competence were included to assess any evidence of stereotyping as suggested by the Stereotype Content Model (Fiske et al., 2002).

In addition to these items, Indian participants received several additional prompts that were omitted for U.S. participants as the queries required more detailed knowledge of Indian culture. Specifically, Indian participants were asked to further estimate the ethnicity of each Indian Asian target (with response options: North Indian, South Indian, North East Indian, East Indian, West Indian, Anglo Indian, and Other), their caste category (upper, middle, lower and tribe) and their religious affiliation (Hindu, Muslim, Christian, Sikh, Buddhist, Jain, Jewish, Parsi/Zoroastrian, No religion, and Other). These data are not of interest to the current study but are available via the norming data distributed with the CFD-India face set.

We took several steps to ensure data quality. Participants completed a bot check (captcha) and a Geo-IP check at the start of the task. The Geo-IP check filtered for participant IP addresses to be located either in India or the United States while excluding participants connected via a Virtual Private Network (VPN) to mask their country location. Following the Bot/Geo-IP check, the actual survey began with an instructional manipulation check (IMC; Oppenheimer et al., 2009). This IMC was intended to screen out random clicking participants. It consisted of a set of instructions at the top of the screen, followed by a Likert scale with items labeled 1 through 9, and an arrow at the bottom of the screen. Instructions asked participants to advance to the next screen by clicking on the arrow and to ignore the scale items.

1,709 Indian participants and 2,937 U.S. participants offered consent and cleared the bot check. Of these, 1,226 Indian participants and 1,839 U.S. participants passed the Geo-IP test and completed their task. Of these participants, 981 (80%) Indian participants and 1,371 (75%) U.S. participants responded accurately to the attention check, suggesting similar data quality in the India and U.S. samples. Of these, 878 Indian participants (238 female, average age = 33.51, age sd = 8.48) and 900 U.S. participants (392 female, average age = 37.61, age sd = 11.39) self-reported as Asian Indian and White/Caucasian, respectively, and had no missing data in their records.

### **Task B**

For this second task, we divided the 284 target faces into four subsets along target gender and ethnicity: Indian Asian females, Indian Asian males, Caucasian females, and Caucasian males. In a between-subjects design with face subset as the between-participant factor, each participant was presented with 40 target faces chosen at random from one of these four face subsets. The entire task took about 10 min to complete; U.S. participants were compensated with \$2 and Indian participants with Rs. 70.

Depending on the experimental condition, participants were instructed that they would see pictures of Indian Asian males (Indian Asian females; Caucasian males; Caucasian females). The instructions further explained that these people would differ in terms of how much their physical features resemble the features of Indian (White) people. For example, their skin color, hair, eyes, nose, cheeks, lips, and other physical features, may be more Indian/White (i.e., typical of Indians/White people) or less Indian/White (i.e., less typical of Indians/White people). Their task would be to rate how Indian (White) looking each person's physical features were. Thereafter participants saw Indian Asian (Caucasian) male (female) targets one at a time, and rated how typical that person's physical features are of Indian (White) people. They were offered a 5-point scale (less typically Indian (White) looking, somewhat typically Indian (White) looking, fairly typically Indian (White) looking, more typically Indian (White) looking and very typically Indian (White) looking).

Participants also completed the same set of bot check (captcha), Geo-IP check, and ICM used in task A. Given the screen layout and response format stayed consistent in this task, rather than switch from screen to screen as in task A, we included a second attention check. For this check, the very last target trial displayed a female Latino target face with the word "Less" superimposed on the forehead. Instructions asked participants to select the response option that matched the word displayed on the face.

339 Indian and 594 U.S. participants offered consent and cleared the bot check. Of these, 335 Indian and 459 U.S. participants also cleared the Geo-IP check. 260 (78%) Indian and 348 (76%) U.S. participants responded accurately to the first attention check, of which 218 Indian and 276 U.S. participants also responded accurately to the second attention check, again indicating similar data quality in the India and U.S. samples. Among this participant set, 215 Indian participants (51 female, average age = 31.35, age sd = 7.46) and 207 U.S. participants (91 female, average age = 37.19, age sd = 12.02) self-reported as Asian Indian and White/Caucasian, respectively and completed the entire task.

### Objective Norms

For the Caucasian faces included in the current study, measurements of the physical features are available as part of the existing CFD norming data. For the Indian Asian face stimuli, we carried out physical measurements in accordance with the procedures described in Ma et al. (2015). **Table 1** in the **Supplementary Material** summarizes all measures and the calculations used to obtain them. In response to requests from researchers, and because the literature in some cases has used multiple definitions for a given measure, the objective norms have been expanded since the original release of the database. We included the full expanded set of physical norms in our assessment of the Indian Asian face stimuli. Specifically, the following measurements were obtained: median luminance of the face, nose width, nose length, lip thickness, face length, height and width of each eye, face width at the most prominent part of the cheek, face width at the mouth, face width at the ears, forehead length, distance between each pupil and the top of the

head, distance between each pupil and the upper lip, distance between pupils, chin length, length of cheek to chin for both sides of the face, the distance between the middle of each brow and the hairline atop that brow, face color (red, green, blue), hair color (red, green, blue), thickness of each eyebrow and eyelid. Using the CFD measurement guide (available on the database website), three coders independently completed the measurements in Adobe Photoshop. For each face and measure, the coders' average measurements were computed and individual measurements that exceeded the mean by 20% in either direction were flagged. These differences were then discussed and reconciled by the research team (consisting of the three coders, joined by A.L., and B.W.) A final set of measures was obtained based on the resulting raters' averages. The inter-rater reliability for these measurements was acceptable to high (Cronbach's alpha equaled 0.69 on face width at cheeks, and was between 0.72 and 0.99 on all other attributes).

## RESULTS

### Subjective Norms

Our analyses focus on the subjective impression and ethnic classification ratings. Specifically, these analyses address two questions with regard to how the participant sample (India vs. U.S.) may have impacted ratings of the target faces: (1) do the resulting stimulus norms for the target groups vary with the participant sample, and if so, do these differences reflect stereotyping and/or ingroup favoritism? (2) do perceptions of face ethnicity vary with participant sample, such that categorization accuracy is higher for ingroup than outgroup targets? Across analyses, participant and target group were each contrast coded (0.5 = Indian/Indian Asian, -0.5 = U.S./Caucasian)<sup>3</sup>.

### Impression Ratings

We first considered whether the subjective stimulus norms varied with participant sample and whether any of these differences varied with target group, across impression attributes. Next, we examined the specific effects on individual impression attributes. For the overall effect, we ran a linear mixed effects model using the lme4 package (Bates et al., 2015) in R with participant impression ratings as the dependent variable and participant group and target group as independent variables. In this analysis, attribute ratings were standardized within each attribute, and then averaged across attributes per participant per target. Participant and target face were included as random effects variables. The full set of results from this analysis is available in **Supplementary Material**. We focus here on the participant group main effect and the target group by participant group interaction. There was no significant main effect of participant group ( $p = 0.722$ ) but there was indeed a significant interaction effect between participant and target group [ $t_{(12188.4)} = -7.71, p < 0.001, \eta^2_p = 0.005 (0.00, 0.01)$ ]. Across attributes, both Indian participants [Caucasian

<sup>3</sup>Our analyses are based on 172 targets of Indian Asian ethnicity. In addition, the India face set includes 12 models of North East Indian Asian ethnicity that were not considered for the current analyses.

faces: Mean  $z$  score = 0.621 vs. Indian Asian faces: Mean  $z$  score =  $-0.749$ ;  $t_{(12188.4)} = -9.28$ ,  $p < 0.001$ ] and U.S. participants [Caucasian faces: Mean  $z$  score = 0.359 vs. Indian Asian faces: Mean  $z$  score =  $-0.233$ ;  $t_{(12188.4)} = -3.89$ ,  $p < 0.001$ ] gave higher impression ratings for Caucasian faces than Indian Asian faces, however this effect was significantly higher among Indian participants.

To clarify how participant group impacted each of the impression attributes, and whether any of these differences varied with target group, we conducted separate linear mixed effects models for each attribute (using the same model specifications as in the parent model, but scores were not standardized within attribute since we were not combining data across attributes for these analyses). Means and test statistics for the participant group and target group main effects are reported in **Table 2** (see the online **Supplementary Material** for a complete set of test statistics). In these analyses, participant group had a main effect on impression ratings for happiness, anger, surprise, fear, masculinity, babyface, competence and perceived status. Indian, compared to U.S. participants rated the target faces to be more happy [ $t_{(1769.9)} = 3.24$ ,  $p = 0.001$ ,  $\eta^2_p = 0.006$  (0.00, 0.01)] and less angry [ $t_{(1770.2)} = -2.89$ ,  $p = 0.004$ ,  $\eta^2_p = 0.005$  (0.00, 0.01)]. Several impression attributes showed target group main effects: Indian Asian faces were judged to be less babyfaced [ $t_{(278.8)} = -7.58$ ,  $p < 0.001$ ,  $\eta^2_p = 0.170$  (0.11, 0.24)] and more unusual [ $t_{(251.2)} = 3.10$ ,  $p = 0.002$ ,  $\eta^2_p = 0.040$  (0.01, 0.08)] than Caucasian faces. In addition, participant group by target group interactions emerged for attractiveness, competence, trustworthiness, anger, masculinity, babyfacedness, unusualness, and status. A breakdown of these interactions is reported in **Table 2**. Next, we explored whether these observed effects reflected any systematic pattern of stereotyping and/or ingroup favoritism.

### Stereotyping

Here we considered the ratings for the basic stereotype dimensions suggested by the Stereotype Content Model (SCM; Fiske et al., 2002; also see Kervyn et al., 2015), warmth, trustworthiness, and competence. In our analyses, stereotyping could be evidenced as a target group main effect, such that participants from both India and the U.S. differentiate Indian Asian from Caucasian faces in similar fashion. Alternatively, Indian and U.S. participants could stereotype their respective ingroup and outgroup differently, resulting in a participant by target group interaction. While analyses for the competence ratings showed no significant target group main effect ( $p = 0.800$ ), a significant main effect of participant group, qualified by a significant interaction effect between target group and participant group emerged [ $t_{(12199.4)} = -4.45$ ,  $p < 0.001$ ,  $\eta^2_p = 0.002$  (0.00, 0.00)]. Simple slopes analyses reveal that U.S. participants rated Indian Asian faces (Mean = 4.61, SD = 1.28) as marginally more competent than Caucasian faces [Mean = 4.55, SD = 1.32;  $t_{(12199.4)} = 1.75$ ,  $p = 0.080$ ]. On the other hand, Indian participants, rated Caucasian faces (Mean = 4.20, SD = 1.50) as significantly more competent than Indian Asian faces [Mean = 4.11, SD = 1.56;  $t_{(12199.4)} = -2.19$ ,  $p = 0.030$ ; see

**TABLE 2** | Attribute ratings by participant group and target group.

Attribute	Effect means and standard deviations			
Attractive	Participant × Target***	All Ps	India Ps**	U.S. Ps**
	All Targets***	4.07 <sub>(1.62)</sub>	4.04 <sub>(1.65)</sub>	4.09 <sub>(1.59)</sub>
	India Targets**	3.88 <sub>(1.62)</sub>	3.79 <sub>(1.65)</sub>	3.96 <sub>(1.58)</sub>
	U.S. Targets	4.26 <sub>(1.60)</sub>	4.30 <sub>(1.61)</sub>	4.21 <sub>(1.60)</sub>
Warm	Participant × Target	All Ps	India Ps	U.S. Ps
	All Targets	3.88 <sub>(1.58)</sub>	3.87 <sub>(1.58)</sub>	3.89 <sub>(1.58)</sub>
	India Targets	3.86 <sub>(1.58)</sub>	3.84 <sub>(1.58)</sub>	3.88 <sub>(1.58)</sub>
	U.S. Targets	3.90 <sub>(1.57)</sub>	3.91 <sub>(1.57)</sub>	3.90 <sub>(1.57)</sub>
Competent	Participant × Target***	All Ps***	India Ps*	U.S. Ps
	All Targets	4.37 <sub>(1.43)</sub>	4.15 <sub>(1.53)</sub>	4.58 <sub>(1.30)</sub>
	India Targets**	4.36 <sub>(1.45)</sub>	4.11 <sub>(1.56)</sub>	4.61 <sub>(1.29)</sub>
	U.S. Targets**	4.38 <sub>(1.42)</sub>	4.20 <sub>(1.50)</sub>	4.55 <sub>(1.32)</sub>
Trustworthy	Participant × Target*	All Ps	India Ps	U.S. Ps
	All Targets	4.33 <sub>(1.42)</sub>	4.32 <sub>(1.47)</sub>	4.34 <sub>(1.38)</sub>
	India Targets	4.34 <sub>(1.42)</sub>	4.30 <sub>(1.46)</sub>	4.37 <sub>(1.38)</sub>
	U.S. Targets	4.32 <sub>(1.42)</sub>	4.33 <sub>(1.47)</sub>	4.32 <sub>(1.38)</sub>
Happy	Participant × Target	All Ps**	India Ps**	U.S. Ps*
	All Targets**	3.37 <sub>(1.75)</sub>	3.47 <sub>(1.73)</sub>	3.28 <sub>(1.77)</sub>
	India Targets*	3.25 <sub>(1.73)</sub>	3.33 <sub>(1.70)</sub>	3.17 <sub>(1.75)</sub>
	U.S. Targets**	3.50 <sub>(1.77)</sub>	3.60 <sub>(1.74)</sub>	3.40 <sub>(1.79)</sub>
Angry	Participant × Target*	All Ps**	India Ps	U.S. Ps
	All Targets	2.85 <sub>(1.78)</sub>	2.76 <sub>(1.72)</sub>	2.93 <sub>(1.83)</sub>
	India Targets**	2.87 <sub>(1.77)</sub>	2.76 <sub>(1.70)</sub>	2.97 <sub>(1.83)</sub> <sup>c</sup>
	U.S. Targets*	2.83 <sub>(1.78)</sub>	2.76 <sub>(1.73)</sub>	2.90 <sub>(1.83)</sub> <sup>c</sup>
Sad	Participant × Target	All Ps	India Ps*	U.S. Ps
	All Targets*	3.20 <sub>(1.82)</sub>	3.23 <sub>(1.81)</sub>	3.17 <sub>(1.82)</sub>
	India Targets	3.31 <sub>(1.82)</sub>	3.35 <sub>(1.82)</sub>	3.26 <sub>(1.82)</sub>
	U.S. Targets	3.09 <sub>(1.80)</sub>	3.11 <sub>(1.79)</sub>	3.08 <sub>(1.82)</sub>
Disgusted	Participant × Target	All Ps	India Ps	U.S. Ps*
	All Targets	2.76 <sub>(1.75)</sub>	2.77 <sub>(1.69)</sub>	2.76 <sub>(1.82)</sub>
	India Targets	2.79 <sub>(1.76)</sub>	2.79 <sub>(1.68)</sub>	2.80 <sub>(1.83)</sub>
	U.S. Targets	2.74 <sub>(1.75)</sub>	2.76 <sub>(1.69)</sub>	2.71 <sub>(1.81)</sub>
Surprised	Participant × Target	All Ps***	India Ps**	U.S. Ps**
	All Targets***	2.71 <sub>(1.77)</sub>	2.91 <sub>(1.72)</sub>	2.52 <sub>(1.81)</sub>
	India Targets**	2.65 <sub>(1.75)</sub>	2.83 <sub>(1.70)</sub>	2.47 <sub>(1.79)</sub>
	U.S. Targets**	2.78 <sub>(1.79)</sub>	2.98 <sub>(1.73)</sub>	2.57 <sub>(1.83)</sub>
Fearful	Participant × Target	All Ps***	India Ps	U.S. Ps*
	All Targets	2.81 <sub>(1.75)</sub>	2.95 <sub>(1.71)</sub>	2.68 <sub>(1.79)</sub>
	India Targets**	2.86 <sub>(1.76)</sub>	2.98 <sub>(1.71)</sub>	2.74 <sub>(1.79)</sub>
	U.S. Targets**	2.77 <sub>(1.75)</sub>	2.92 <sub>(1.71)</sub>	2.62 <sub>(1.78)</sub>
Threatening	Participant × Target	All Ps	India Ps	U.S. Ps
	All Targets	2.96 <sub>(1.79)</sub>	2.98 <sub>(1.75)</sub>	2.93 <sub>(1.84)</sub>
	India Targets	2.96 <sub>(1.79)</sub>	2.98 <sub>(1.74)</sub>	2.94 <sub>(1.84)</sub>
	U.S. Targets	2.95 <sub>(1.8)</sub>	2.98 <sub>(1.76)</sub>	2.93 <sub>(1.83)</sub>
Masculine	Participant × Target*	All Ps***	India Ps	U.S. Ps
	All Targets	3.84 <sub>(2.07)</sub>	3.75 <sub>(2.08)</sub>	3.93 <sub>(2.06)</sub>
	India Targets**	3.89 <sub>(2.10)</sub>	3.78 <sub>(2.11)</sub>	4.01 <sub>(2.09)</sub>
	U.S. Targets*	3.80 <sub>(2.04)</sub>	3.73 <sub>(2.05)</sub>	3.86 <sub>(2.02)</sub>

(Continued)

TABLE 2 | Continued

Attribute	Effect means and standard deviations			
Feminine	Participant × Target	All Ps	India Ps	U.S. Ps
	All Targets	3.81 <sub>(2-17)</sub>	3.81 <sub>(2-21)</sub>	3.80 <sub>(2-13)</sub>
	India Targets	3.70 <sub>(2-19)</sub>	3.71 <sub>(2-24)</sub>	3.70 <sub>(2-13)</sub>
	U.S. Targets	3.91 <sub>(2-15)</sub>	3.91 <sub>(2-17)</sub>	3.91 <sub>(2-13)</sub>
Babyfaced	Participant × Target**	All Ps***	India Ps**	U.S. Ps**
	All Targets***	3.14 <sub>(1-84)</sub>	2.94 <sub>(1-81)</sub>	3.34 <sub>(1-85)</sub>
	India Targets**	2.92 <sub>(1-81)</sub>	2.69 <sub>(1-74)</sub>	3.16 <sub>(1-84)</sub>
	U.S. Targets**	3.37 <sub>(1-85)</sub>	3.20 <sub>(1-84)</sub>	3.53 <sub>(1-84)</sub>
Unusual	Participant × Target***	All Ps	India Ps*	U.S. Ps**
	All Targets**	3.09 <sub>(1-79)</sub>	3.13 <sub>(1-75)</sub>	3.04 <sub>(1-82)</sub>
	India Targets	3.14 <sub>(1-78)</sub>	3.09 <sub>(1-75)</sub>	3.19 <sub>(1-82)</sub>
	U.S. Targets**	3.04 <sub>(1-79)</sub>	3.18 <sub>(1-76)</sub>	2.90 <sub>(1-82)</sub>
Status	Participant × Target***	All Ps***	India Ps**	U.S. Ps**
	All Targets***	4.42 <sub>(1-31)</sub>	4.58 <sub>(1-30)</sub>	4.27 <sub>(1-29)</sub>
	India Targets**	4.17 <sub>(1-32)</sub>	4.26 <sub>(1-33)</sub>	4.09 <sub>(1-31)</sub>
	U.S. Targets**	4.67 <sub>(1-24)</sub>	4.89 <sub>(1-19)</sub>	4.45 <sub>(1-26)</sub>

All cell values are Means<sub>(SD)</sub>; Main Effects indicated at "All Targets"/"All Participants" cells (for Target and Participant main effects respectively); Interaction effects indicated at "Participant × Target" cells; Simple effects indicated at "India Ps"/"U.S. Ps"/"India Targets"/"U.S. Targets" cells; \* $p \leq 0.05$ . \*\* $p \leq 0.01$ . \*\*\* $p \leq 0.001$ .

**Table 2].** Trustworthiness ratings showed no significant main effect of target group or participant group, but again, yielded a significant interaction effect between target group and participant group, with the respective outgroup faces being seen as more trustworthy (Indian participants: Mean = 4.33,  $SD = 1.47$ ; U.S. participants: Mean = 4.37,  $SD = 1.38$ ) than the ingroup (Indian participants: Mean = 4.30,  $SD = 1.46$ ; U.S. participants: Mean = 4.32,  $SD = 1.38$ ) ratings [ $t_{(12185-9)} = -2.20$ ,  $p = 0.028$ ,  $\eta^2_p = 0.0004$  (0.00, 0.00); see **Table 2**]. Simple slopes analyses for comparing the target group means within participant group were not significant (all  $ps > 0.270$ ). No significant effects emerged for perceived warmth (all  $ps > 0.141$ ).

Group perceptions of competence have reliably been found to correlate with and be informed by perceived social status (Fiske et al., 1999; Caprariello et al., 2009), suggesting that ratings of perceived social status should parallel our results for competence. In fact, the analyses for perceived status do yield this target group by participant group interaction [ $t_{(12170-4)} = -8.28$ ,  $p < 0.001$ ,  $\eta^2_p = 0.006$  (0.00, 0.01)]. However, the pattern of means deviates somewhat from the results for perceived competence. Simple slopes analyses indicate that, while both participant groups rated Caucasian faces higher in status, this effect was greater among Indian participants [Caucasian faces: Mean = 4.89,  $SD = 1.19$  vs. Indian Asian faces: Mean = 4.26,  $SD = 1.33$ ;  $t_{(12170-4)} = -11.41$ ,  $p < 0.001$ ] than U.S. participants [Caucasian faces: Mean = 4.45,  $SD = 1.26$  vs. Indian Asian faces: Mean = 4.09,  $SD = 1.31$ ;  $t_{(12170-4)} = -6.34$ ,  $p < 0.001$ ; see **Table 2**]. Given this pattern, correlations between perceived competence and status remain modest ( $r = 0.31$ ).

In summary, our analyses for participants' ratings of warmth, competence, and trustworthiness show no overall target group differences for warmth, competence, or trustworthiness.

However, we do observe differentiation in the impressions of Indian and U.S. Caucasian faces between the two participant groups. For competence and trustworthiness, both participant groups rated the respective outgroup somewhat higher than their own ingroup.

### Ingroup favoritism

The second question we posed regarding the impression ratings is whether participants would see ingroup targets overall more favorably than outgroup targets. The results for perceived competence and trustworthiness we just summarized would suggest that if anything the current data show the reverse pattern, with outgroup faces receiving more favorable ratings than ingroup faces, on these attributes. In order to address this question more systematically, we calculated two scores to capture the favorability of the impressions: a positivity score using the ratings from all positively valenced impression attributes (attractive, warm, competent, trustworthy, happy; Cronbach's  $\alpha = 0.81$ ) and a negativity score with the ratings of all negatively valenced attributes (angry, sad, disgusted, fearful, threatening; Cronbach's  $\alpha = 0.86$ ). We calculated a difference score (positivity score—negativity score) as an indicator of impression favorability (Wittenbrink, 2007). We then analyzed these favorability scores in a linear mixed effects model using the lme4 package in R to analyze the data with the favorability scores as the dependent variable and target group and participant group as independent variables. We employed random intercepts for participant and target face stimulus.

The full set of results from this analysis is available in the online **Supplementary Material**. We focus here on ingroup favoritism, which is represented by the target group and participant group interaction. The effect was small but significant [ $t_{(12171-5)} = -2.24$ ,  $p = 0.025$ ,  $\eta^2_p = 0.0004$  (0.00, 0.00)]. For U.S. participants, impressions of ingroup faces were marginally more favorable than their impressions of outgroup faces [Ingroup faces: Mean = 1.23,  $SD = 1.73$ ; Outgroup faces: Mean = 1.05,  $SD = 1.71$ ;  $t_{(12171-5)} = -1.81$ ,  $p = 0.070$ ]. For Indian participants on the other hand this pattern reversed. Impressions of outgroup faces were significantly more favorable than their impressions of ingroup faces [Ingroup faces: Mean = 0.90,  $SD = 1.67$ ; Outgroup faces: Mean = 1.16,  $SD = 1.76$ ,  $t_{(12171-5)} = -2.87$ ,  $p < 0.001$ ].

### Typicality

A final impression item asked participants to rate the target faces in terms of group typicality. We first examined the effects of target group and participant group on perceived target typicality. Analyses of these typicality ratings employed the same mixed effects model used for all other impression attributes. With regard to effects involving participant group, these analyses yielded a significant main effect—Indian participants (Mean = 3.46,  $SD = 1.18$ ) rated typicality overall higher than U.S. participants [Mean = 3.24,  $SD = 1.28$ ;  $t_{(15902-2)} = 12.87$ ,  $p < 0.001$ ,  $\eta^2_p = 0.010$  (0.01, 0.01)]. This main effect was qualified by a significant interaction effect between participant group and target group [ $t_{(15902-2)} = 12.84$ ,  $p < 0.001$ ,  $\eta^2_p = 0.010$  (0.01, 0.01)]. Simple slopes analyses indicate that Caucasian faces (Mean = 3.41,  $SD = 1.29$ ) were perceived as significantly more typical than



Indian Asian faces (Mean = 3.04,  $SD = 1.24$ ) by U.S. participants [ $t_{(15902.2)} = -6.15, p < 0.001$ ] but the same difference did not emerge for Indian participants (Caucasian faces: Mean = 3.41,  $SD = 1.14$ ; Indian Asian Faces: Mean = 3.51,  $SD = 1.22$ ;  $p = 0.140$ ). Unrelated to our question of interest, there was also a significant main effect of target group on perceived typicality—Caucasian faces (Mean = 3.41,  $SD = 1.22$ ) received higher typicality ratings overall than Indian Asian faces [Mean = 3.28,  $SD = 1.25$ ;  $t_{(270.3)} = -2.45, p = 0.015, \eta^2_p = 0.020 (0.00, 0.06)$ ].

## Face Categorization

Our second primary research question concerned perceptions of face ethnicity for the two target groups and whether they would vary with participant sample. Because of greater familiarity with ingroup faces, we expected participants to more accurately identify the ethnicity of their respective ingroup faces.

### Categorization Accuracy

To address this question we calculated for each target face the probability of accurate categorization as a proportion of the number of times the target face was categorized correctly (i.e., an Indian Asian face identified as Asian Indian, and a Caucasian face judged to be Caucasian), relative to the number of times it was categorized at all. The resulting accuracy score served as the dependent variable in a binomial generalized linear model using the lme4 package (Bates et al., 2015) in R with target group and participant group as independent variables. Weights were added to the model based on categorization count; i.e., the number of times each target was categorized at all.

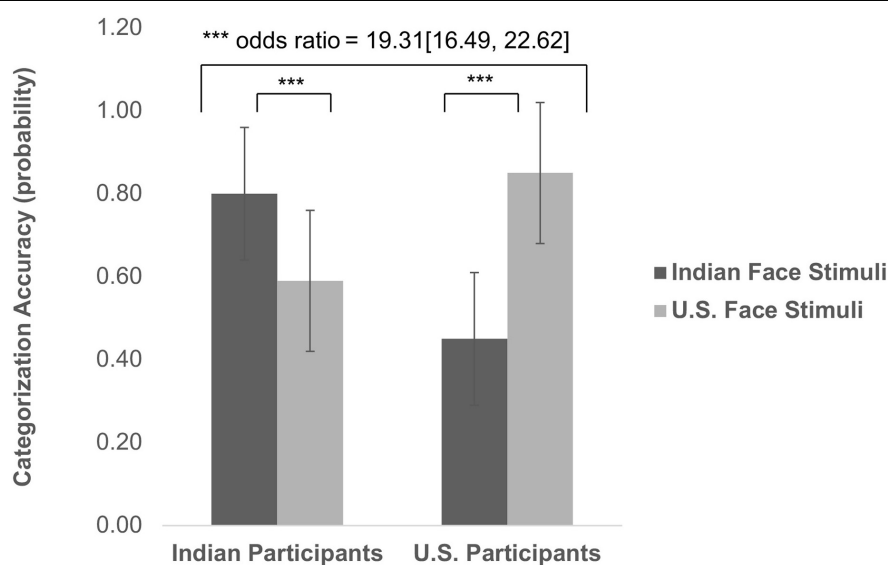
Consistent with the expected ingroup accuracy advantage, there was a significant interaction effect between target group and participant group [ $t_{(543)} = 36.69, p < 0.001$ , odds ratio = 19.31 (16.49, 22.62); see **Figure 2**]. Simple slopes analyses of the interaction between target group and participant

group indicate that among U.S. participants, the probability of accurate categorization was significantly higher for Caucasian faces (Mean = 0.85,  $SD = 0.17$ ) than for Indian Asian faces [Mean = 0.45,  $SD = 0.16$ ;  $t_{(543)} = -34.23, p < 0.001$ ]. Indian participants showed a similar ingroup accuracy bias. For them, the probability of accurate categorization was significantly higher for Indian Asian faces (Mean = 0.80,  $SD = 0.17$ ) than for Caucasian faces [Mean = 0.59,  $SD = 0.16$ ;  $t_{(543)} = 17.09, p < 0.001$ ]. Unrelated to our primary question, there was a significant main effect of target group [ $t_{(543)} = -13.39, p < 0.001$ , odds ratio = 0.58 (0.54, 0.63)]: Categorization accuracy was overall higher for Caucasian faces (Mean = 0.72,  $SD = 0.21$ ) than Indian Asian faces (Mean = 0.62,  $SD = 0.24$ ).

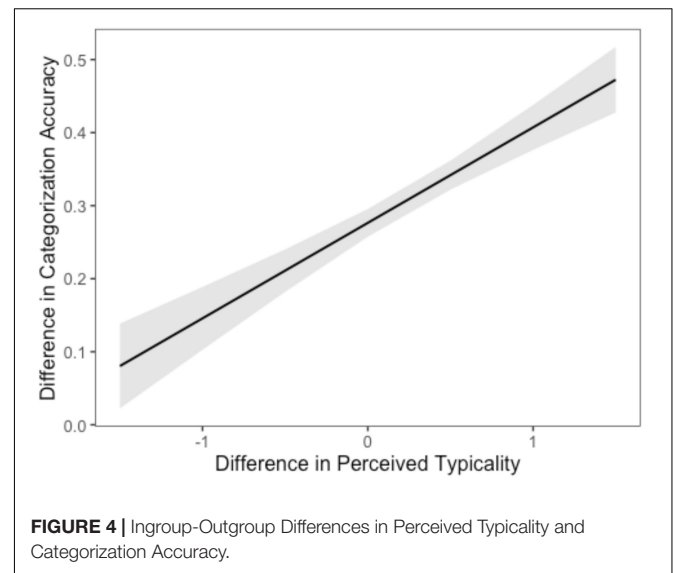
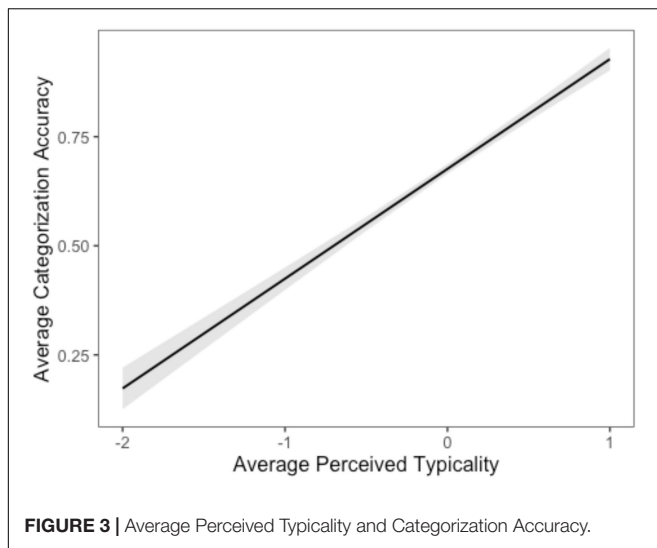
### Typicality

Another factor that might impact the categorization of faces is their ethnic typicality. That is, one might expect faces that are seen to be more typically Indian in appearance to be more readily classified as Indian Asian. In fact, such effects of typicality on categorization are well established. A robin is more readily recognized as a bird than an ostrich (Rosch, 1973). Face categorization, including categorization by ethnicity, is no exception and is similarly sensitive to typicality effects (Maddox and Gray, 2002; Locke et al., 2005). We therefore used the typicality impression ratings we already reported earlier to test whether the observed ingroup advantage in categorization accuracy is mediated by perceptions of typicality prevalent in the two participant groups.

For ingroup advantage in categorization accuracy to be mediated by perceived typicality, two conditions have to be met: (1) typicality should affect categorization accuracy (a test of the link between perceived typicality and categorization accuracy); and (2) ingroup-outgroup differences in typicality should affect ingroup outgroup differences in categorization accuracy (a test



**FIGURE 2** | Ingroup-Outgroup Differences in Categorization Accuracy.



of the link between ingroup advantage, perceived typicality and categorization accuracy; see Judd et al., 2001).

For each target, we calculated a mean typicality rating and categorization accuracy score for ingroup participants, as well as a mean typicality rating and categorization accuracy score for outgroup participants. Using these measures, we obtained four values for each target: average typicality rating (across ingroup and outgroup), average categorization accuracy (across ingroup and outgroup), difference in typicality rating (ingroup-outgroup) and difference in percentage accuracy (ingroup-outgroup). Using these scores, we set up two linear models to test the influence of group membership and typicality ratings on categorization accuracy.

The first linear model used average categorization accuracy as dependent variable and mean centered average typicality as independent variable. There was a significant intercept [Mean = 0.68,  $t_{(270)} = 119.30$ ,  $p < 0.001$ ] suggesting that on average, categorization accuracy was significantly above zero controlling for typicality. Average typicality added significantly to categorization accuracy [ $t_{(270)} = 21.24$ ,  $p < 0.001$ ,  $\eta^2 = 0.630$  (0.57, 0.67)], with categorization accuracy improving with higher perceived average typicality (see **Figure 3**). In other words, average typicality did affect categorization accuracy.

In the second linear model, difference in categorization accuracy served as the dependent variable and the ingroup/outgroup difference in average typicality ratings served as the independent variable. There was a significant intercept Mean = 0.28,  $t_{(270)} = 28.55$ ,  $p < 0.001$ , suggesting that on average, ingroup/outgroup difference in categorization accuracy was significantly above zero, controlling for ingroup/outgroup difference in average typicality. Ingroup-outgroup difference in mean typicality ratings added significantly to the ingroup-outgroup difference in categorization accuracy. As the ingroup-outgroup difference in mean rating increased, so did the ingroup-outgroup difference in categorization accuracy [ $t_{(270)} = 7.96$ ,  $p < 0.001$ ,  $\eta^2 = 0.190$  (0.12, 0.26); see **Figure 4**].

Thus, these analyses suggest that the effect of familiarity (as determined by group membership) on categorization accuracy was mediated significantly albeit not fully, by perceived typicality.

### Miscategorization

Finally, we explored what categories were used in error when Indian Asian faces were not identified as Indian Asian, and Caucasian faces not judged to be Caucasian. Toward this, we selected all instances of inaccurate categorizations and identified the two most common ethnicities participants chose in these instances, Middle Eastern and Hispanic/Latino, accounting for 54.65% of all erroneous categorizations. For each target face, we then generated percentages of inaccurate categorization as (1) Middle Eastern and (2) as Latino, separate for Indian participants and U.S. participants, respectively. For example, to calculate the percentage of inaccurate categorization as Middle Eastern, the number of times a target was inaccurately categorized as Middle Eastern was divided by the number of times it was inaccurately categorized at all. As with accurate categorizations, we employed a binomial generalized linear model using the lme4 package (Bates et al., 2015) in R to analyze the data with probability of inaccurate categorization as Middle Eastern and Latino, respectively, as the dependent variables and target group and participant group as independent variables, weighted by target categorization count.

For inaccurate categorization into Middle Eastern, there was a significant effect of target group such that Indian Asian faces (Mean = 0.31,  $SD = 0.26$ ) were inaccurately categorized as Middle Eastern with higher probability than Caucasian faces were [Mean = 0.21,  $SD = 0.22$ ;  $t_{(504)} = 1.97$ ,  $p = 0.049$ , odds ratio = 1.18 (1.00, 1.38)]. However, this main effect was qualified by a significant interaction between target group and participant group [ $t_{(504)} = -5.32$ ,  $p < 0.001$ , odds ratio = 0.42 (0.30, 0.58)]: For both participant groups, errors made for outgroup faces were more likely to be misjudged as Middle Eastern, compared to ingroup faces with errors. Simple slopes analyses indicate that

that U.S. participants inaccurately categorized Indian Asian faces (Mean = 0.42,  $SD = 0.23$ ) as Middle Eastern at a significantly higher probability than they did Caucasian faces [Mean = 0.22,  $SD = 0.28$ ;  $t_{(504)} = 5.33$ ,  $p < 0.001$ ]. Indian participants on the other hand, categorized Caucasian faces (Mean = 0.21,  $SD = 0.15$ ) as Middle Eastern at a significantly higher probability than they did with Indian Asian faces [Mean = 0.18,  $SD = 0.23$ ;  $t_{(504)} = -2.29$ ,  $p = 0.020$ ].

For inaccurate categorization into Latino, there was a significant main effect of target group such that Caucasian faces (Mean = 0.46,  $SD = 0.30$ ) were inaccurately categorized as Latino more often than were Indian Asian faces [Mean = 0.19,  $SD = 0.20$ ;  $t_{(504)} = -14.12$ ,  $p < 0.001$ , odds ratio = 0.31 (0.27, 0.37)]. The main effect was again qualified by a significant target group and participant group interaction [ $t_{(504)} = -2.47$ ,  $p = 0.014$ , odds ratio = 0.67 (0.48, 0.92)]. Indian participants inaccurately categorized Caucasian faces (Mean = 0.36,  $SD = 0.20$ ) as Latino at a significantly greater probability than they did Indian Asian faces [Mean = 0.10,  $SD = 0.16$ ;  $t_{(504)} = -10.51$ ,  $p < 0.001$ ]. U.S. participants as well, inaccurately categorized Caucasian faces (Mean = 0.58,  $SD = 0.35$ ) as Latino at a significantly greater probability than they did Indian Asian faces, but his effect was smaller than among Indian participants [Mean = 0.27,  $SD = 0.20$ ;  $t_{(504)} = -9.47$ ,  $p < 0.001$ ].

Interestingly, and related to our main research questions with respect to the effect of participant group, we also observed a significant main effect of participant group such that U.S. participants (Mean = 0.32,  $SD = 0.28$ ) inaccurately categorized faces as Middle Eastern at a significantly higher probability than Indian participants did [Mean = 0.20,  $SD = 0.19$ ;  $t_{(504)} = -7.71$ ,  $p < 0.001$ , odds ratio = 0.53 (0.45, 0.62)]. Also, U.S. participants (Mean = 0.42,  $SD = 0.32$ ) inaccurately categorized faces as Latino at a significantly higher probability than Indian participants did [Mean = 0.24,  $SD = 0.22$ ;  $t_{(504)} = -12.83$ ,  $p < 0.001$ , odds ratio = 0.35 (0.30, 0.41)].

## DISCUSSION

Human faces are an important factor in social life. Perceivers use them for a wide range of social inferences about emotions, personal identity, social category membership, traits, preferences, and even culpability in legal cases (e.g., Ekman et al., 1972; Blair et al., 2004; Zebrowitz and Montepare, 2008; Todorov et al., 2015). As a result, a good part of social psychological research involves the presentation of face stimuli. The Chicago Face Database (CFD) is a frequently used resource for this type of work. Since its release just 5 years ago, the database materials have been retrieved by over 7,000 researchers worldwide and some 700 published papers have reported studies with CFD faces. Yet, as is the case with psychological research in general, the database materials remain limited in their cultural and ethnic diversity. Not only by name, the database to-date is U.S.-centric. It contains the faces of volunteers recruited in the U.S., and its stimulus norms are based on U.S. rater samples.

With the current research we set out to broaden the scope of the database and improve its usefulness for work with non-U.S. participants and non-U.S. faces. To this effect, we introduce a new set of face stimuli representing a 142 individuals from a large non-U.S. ethnic group, Indian Asians. We report the development and standardization of these stimulus materials, which follow the established procedures of the database, so that the new Indian Asian images can be used interchangeably with the full set of CFD stimuli. With the new image set, we also provide extensive norming data that cover both the physical face attributes as well as subjective impressions of the faces. Finally, in addition to the neutral expression images relevant to the current research questions, the India face set also includes images of models making a variety of emotional expressions.

The empirical part of the current research then focused on the subjective face impressions included in the norming data. First, we asked whether the resulting face norms are culturally dependent and will vary with the participant sample. To address this issue, we collected impression ratings in a full ingroup-outgroup design with samples of Indian and U.S. participants, for both Indian Asian and Caucasian face images. Results show that impression ratings indeed varied significantly with participant group. Compared to U.S. raters, Indian participants judged faces to be more happy, surprised, fearful, and of higher social status, but less angry, masculine, babyfaced, and competent. The current results add to evidence from other recent studies that impressions from faces are to some extent culturally specific (Sutherland et al., 2018; Wang et al., 2019; Jones et al., 2021). Possibly of greater consequence for the use of these impression norms in selecting or weighting study materials, the differences between participant groups depended on the target group. For example, Indian and U.S. participants significantly differed in their ratings of Indian Asian and Caucasian faces on perceived trustworthiness. Consequently, a study among Indian participants with both Indian Asian and Caucasian faces that relied on U.S. image norms in selecting faces of similar trustworthiness would run the risk of confounding trustworthiness and face ethnicity. Hence, the current findings highlight the importance of obtaining local stimulus norms for research with non-U.S. participant samples.

We further explored whether the differences we observed between Indian and U.S. raters followed systematic patterns of ingroup favoritism and stereotyping. With regard to ingroup favoritism, we observed that U.S. participants reported marginally more favorable impressions for faces of their ingroup, compared to outgroup faces. However, Indian participants' ratings, in contrast, showed outgroup favoritism. Their impressions of outgroup faces were significantly more favorable than their impressions of ingroup faces. The result highlights the importance of conducting research on intergroup relations across diverse cultural and international settings. While the literature has generated a long history of findings demonstrating general ingroup favoritism in social judgment (Brewer, 2007), our results for the Indian participant sample clearly deviate from this established effect.

With regard to stereotyping, we focused on face ratings of warmth, competence, and trustworthiness, following the SCM by Fiske and colleagues (Fiske et al., 2002; Kervyn et al., 2015).

Overall, the results show no target group differences along the basic stereotype dimensions of the SCM. However, we did observe differentiation between the two participant groups. For competence and trustworthiness, both participant groups rated the respective outgroup somewhat higher than their own ingroup. The results may reflect the fact that, for both target groups, outgroup stereotypes are overlapping with ingroup stereotypes. Consistent with this interpretation, Lee and Fiske (2006) found U.S. stereotypes about Indian immigrants living in the U.S. to be largely similar to ingroup stereotypes. In a cluster analysis of stereotype content, Indian immigrants appeared in the same cluster as various ingroups (e.g., college students). Arguably, this study investigated a specific subset of Indian Asians, Indian immigrants in the U.S. However, we should note that in our study the impression rating task (task A) made no reference to the targets' nationality, ethnicity, or any other social category for that matter. Participants merely saw faces. Without mentioning an international context, it seems quite likely that our U.S. participants considered both the Indian Asian faces as well as the Caucasian faces to represent individuals living in the U.S. Similarly, we suspect our Indian participants considered the Indian Asian faces to depict individuals from their immediate environment, India. Caucasian faces, in contrast, are considerably less prevalent in Indian society and may be more readily assumed to be non-Indian foreigners by our Indian participants. Possibly, differences in attributions between the participant groups with regard to the targets' background may account for our results for perceived status. Our data deviate somewhat from prior findings, which generally show substantive correlations between perceived group status and perceptions of competence. However, the existing research here has generally focused on status differences within a given society (e.g., Durante et al., 2017).

The relative prevalence of stereotypic impressions for the two target groups may have been further impacted by participants misclassifying face ethnicity. As a matter of fact, as we had predicted categorization accuracy differed significantly for ingroup and outgroup faces. Moreover, the categories chosen most frequently in error differed for ingroup and outgroup faces. In these instances of misclassification, where, for example, an Indian Asian target is seen to be Middle Eastern, we would expect different stereotypes to impact the impression ratings. The observed misclassification of outgroup faces may have considerable real-world consequences, for example in forensic settings where law enforcement officers may use either explicitly or implicitly a suspect's ethnicity. Likewise, some research suggests that, post 9/11, South Asians living in the United States experienced misclassification as Middle Eastern, resulting in identity threat, stereotyping, and prejudice (Joshi, 2006; Bhatia, 2008; Poolokasingham et al., 2014).

Arguably, there is considerable value in research on group-level stereotypes; research that investigates the content and the dynamics of beliefs about entire groups. And, given the scarcity of data on the stereotypes Indians hold about people from the U.S. and vice versa, we wish more of this kind of group-level research was conducted in an international context.

Our finding that perceptions of face ethnicity depended on the raters' own group membership has both methodological as well as conceptual implications. Methodologically, our data show that what may serve as a typical Indian Asian face in a study with both U.S. and Indian participants is not an equally typical face for both participant groups. Similarly, the manipulation of target group membership or ethnicity through the use of faces (e.g., Krumbhuber et al., 2015) may be compromised if the faces end up being misclassified.

Conceptually, our finding that perceptions of face ethnicity depended on the raters' own group membership is consistent with well-documented effects of familiarity on categorization speed and accuracy (e.g., Smith, 1967; Johnson and Mervis, 1997). However, in the face perception literature, few studies have directly investigated the role of familiarity for the categorization of faces by social group or ethnicity.

Indirect evidence comes from work on the "other-race-effect," whereby own-race faces are more readily and accurately identified than other-race faces (ORE, see Meissner and Brigham, 2001). One explanation for the ORE holds that face processing occurs along face dimensions that effectively differentiate among the types of faces frequently encountered (Goldstein and Chance, 1980; Valentine, 1991). As a result, more familiar own-race faces function as a perceptual default facilitating their processing and identification, while impeding the processing and identification of other-race faces.

The ORE, thus, is consistent with our finding that ethnicity can be inferred more accurately for ingroup than outgroup faces. However, ORE studies do not directly assess categorization accuracy of the stimulus faces. In fact, studies that do ask participants to classify faces by race, have found the opposite effect, showing that classification is *faster* and *more* accurate for other-race than own-race faces, an effect labeled other-classification race advantage (ORCA; see Levin, 1996; Zhao and Bentin, 2011). Yet, a notable difference between these demonstrations and our current study is that our participants chose from a list of eight ethnic categories, whereas ORCA studies use a category-verification task with a binary choice option (e.g., Asian, Caucasian). In category verification, participants see an array of faces and have to decide whether the face is either Asian or Caucasian. Such a binary choice task is likely to increase the salience of features that differentiate between the two groups used in the task (see Wang et al., 2016).

At times, social interactions may require such a binary differentiation. But often interactions lack explicit group identifiers. With considerable frequency, we encounter people not knowing their ethnic origin, whether they are from the U.S., Europe, India, the Middle East, or any other part of the world. Our data capture the kinds of face impressions people form under these circumstances. We believe research on face impression and stereotyping will benefit from considering a cross-cultural and international context in which the origin of a face is not immediately determined by a small set size of stimulus attributes. We hope the India Face set helps facilitate such research.

The data and materials for this research are available at [www.chicagofaces.org](http://www.chicagofaces.org).

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by authors, without undue reservation, to any qualified researchers with verifiable credentials, who do not pose a risk to confidentiality and safety of our participants.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board, University of Chicago. The patients/participants provided their written/online informed consent to participate in this study. Written/online informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

AL and BW designed and supervised image collection and standardization, designed and conducted the norming study, analyzed data, and co-wrote the manuscript. JC analyzed

the data and co-wrote the manuscript. DM co-wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

Support for this work was provided by the University of Chicago Booth School of Business.

## ACKNOWLEDGMENTS

We thank Lucia Agajanian, Japneet Bhatia, Solomon Lister, Roseleen Kaur, Neha Menon, Naomi Nero, Jenny Ni, Chloe Roske, and the University of Chicago Delhi Center staff for their contributions to this research.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.627678/full#supplementary-material>

## REFERENCES

- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., et al. (2013). Recommendations for increasing replicability in psychology. *Eur. J. Pers.* 27, 108–119.
- Bates, D., Maechler, M., Bolker, B., Walker, S., and Haubo Bojesen Christensen, R. (2015). *lme4: Linear Mixed-Effects Models Using Eigen and S4. R Package Version 1.1-7*. 2014.
- Bernstein, I. H., Tsai-Ding, L., and McClellan, P. (1982). Cross- vs. within-racial judgments of attractiveness. *Percept. Psychophys.* 32, 495–503. doi: 10.3758/bf03204202
- Bhatia, S. (2008). 9/11 and the Indian diaspora: narratives of race, place and immigrant identity. *J. Intercult. Stud.* 29, 21–39. doi: 10.1080/07256860701759923
- Blair, I., Judd, C. M., and Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychol. Sci.* 15, 674–679. doi: 10.1111/j.0956-7976.2004.00739.x
- Brewer, M. B. (2007). “The social psychology of intergroup relations: social categorization, ingroup bias, and outgroup prejudice,” in *Social Psychology: Handbook of Basic Principles*, eds A. W. Kruglanski and E. T. Higgins (New York, NY: Guilford Press), 695–715.
- Byatt, G., and Rhodes, G. (1998). Recognition of own-race and other-race caricatures: implications for models of face recognition. *Vision Res.* 38, 2455–2468. doi: 10.1016/s0042-6989(97)00469-0
- Caprariello, P. A., Cuddy, A. J., and Fiske, S. T. (2009). Social structure shapes cultural stereotypes and emotions: a causal test of the stereotype content model. *Group Process. Intergroup Relat.* 12, 147–155. doi: 10.1177/1368430208101053
- Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., et al. (2009). Stereotype content model across cultures: towards universal similarities and some differences. *Br. J. Soc. Psychol.* 48, 1–33. doi: 10.1348/014466608x314935
- Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., and Wu, C.-H. (1995). “Their ideas of beauty are, on the whole, the same as ours”: consistency and variability in the cross-cultural perception of female physical attractiveness. *J. Pers. Soc. Psychol.* 68, 261–279. doi: 10.1037/0022-3514.68.2.261
- DeBruine, L., and Jones, B. (2017). *Face Research Lab London Set*. Dataset. Figshare.
- Durante, F., Fiske, S. T., Gelfand, M. J., Crippa, F., Suttora, C., Stillwell, A., et al. (2017). Ambivalent stereotypes link to peace, conflict, and inequality across 38 nations. *PNAS* 114, 669–674. doi: 10.1073/pnas.1611874114
- Ebner, N., Riediger, M., and Lindenberger, U. (2010). FACES – a database of facial expressions in young, middle-aged, and older women and men: development and validation. *Behav. Res. Methods* 42, 351–362. doi: 10.3758/brm.42.1.351
- Ekman, P., and Friesen, W. V. (1976). *Pictures of Facial Affect*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., and Ellsworth, P. (1972). *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York, NY: Pergamon Press.
- Fiedler, K. (2011). Voodoo correlations are everywhere – not only in social neurosciences. *Perspect. Psychol. Sci.* 6, 163–171. doi: 10.1177/1745691611400237
- Fiske, S. T., Cuddy, A. J. C., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902. doi: 10.1037/0022-3514.82.6.878
- Fiske, S. T., Xu, J., Cuddy, A. C., and Glick, P. (1999). (Dis) respecting versus (dis) liking: status and interdependence predict ambivalent stereotypes of competence and warmth. *J. Soc. Issues* 55, 473–489. doi: 10.1111/0022-4537.00128
- Goldstein, A. G., and Chance, J. E. (1980). Memory for faces and schema theory. *J. Psychol.* 105, 47–59. doi: 10.1080/00223980.1980.9915131
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83. doi: 10.1017/s0140525x0999152x
- Jahoda, G. (1959). Nation Preferences and National stereotypes in Ghana before independence. *J. Soc. Psychol.* 50, 165–174. doi: 10.1080/00224545.1959.9921990
- Johnson, K., and Mervis, C. (1997). Effects of varying levels of expertise on the basic-level of categorization. *J. Exp. Psychol. Gen.* 126, 248–277. doi: 10.1037/0096-3445.126.3.248

- Jones, B., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., and Peters, K. O. (2021). To which world regions does the valence-dominance model of social perception apply? *Nat. Hum. Behav.* 5, 159–169.
- Joshi, K. Y. (2006). The racialization of Hinduism, Islam, and Sikhism in the United States. *Equity Excell. Educ.* 39, 211–226. doi: 10.1080/10665680600790327
- Judd, C. M., Kenny, D. A., and McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychol. Methods* 6:115. doi: 10.1037/1082-989x.6.2.115
- Kahneman, D. (2012). A proposal to deal with questions about priming effects. *Open Lett.* 2012:26.
- Kashima, Y., Kashima, E. S., Gelfand, M., Goto, S., Takata, T., Takemura, K., et al. (2003). War and peace in East Asia: Sino-Japanese relations and National stereotypes. *Peace Confl.* 9, 259–276. doi: 10.1207/s15327949pac0903\_5
- Kervyn, N., Fiske, S., and Yzerbyt, V. (2015). Forecasting the primary dimension of social perception. *Soc. Psychol.* 46, 36–45. doi: 10.1027/1864-9335/a000219
- Krumhuber, E. G., Swiderska, A., Tsankova, E., Kamble, S. V., and Kappas, A. (2015). Real or artificial? Intergroup biases in mind perception in a cross-cultural perspective. *PLoS One* 10:e0137840. doi: 10.1371/journal.pone.0137840
- Lakshmi, A. J., Venezia, S. A., Mattan, B. D., Cloutier, J., and Kubota, J. T. (2019). *Understanding Status and Race Perception Using Reverse Correlation Image Classification*. Available online at: osf.io/9px4n (accessed January 22, 2020).
- Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1997). “Motivated attention: affect, activation, and action,” in *Attention and Orienting: Sensory and Motivational Processes*, eds P. J. Lang, R. F. Simons, and M. Balaban (Mahwah, NJ: Erlbaum), 97–135.
- Langlois, J. H., Roggman, L. A., and Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychol. Sci.* 5, 214–220. doi: 10.1111/j.1467-9280.1994.tb00503.x
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., and van Knippenberg, A. (2010). Presentation and validation of the radboud Faces database. *Cogn. Emot.* 24, 1377–1388. doi: 10.1080/02699930903485076
- Lee, T. L., and Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: immigrants in the stereotype content model. *Int. J. Intercult. Relat.* 30, 751–768. doi: 10.1016/j.ijintrel.2006.06.005
- Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1364–1382. doi: 10.1037/0278-7393.22.6.1364
- Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., and Greene, B. G. (1993). Lexical familiarity and processing efficiency: individual differences in naming, lexical decision, and semantic categorization. *J. Exp. Psychol.* 122, 316–330. doi: 10.1037/0096-3445.122.3.316
- Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *J. Pers. Soc. Psychol.* 48, 563–574. doi: 10.1037/0022-3514.48.3.563
- Locke, V., Macrae, C. N., and Eaton, J. L. (2005). Is person categorization modulated by exemplar typicality? *Soc. Cogn.* 23, 417–428. doi: 10.1521/soco.2005.23.5.417
- Lundqvist, D., Flykt, A., and Öhman, A. (1998). *The Karolinska Directed Emotional Faces*. Stockholm: Karolinska Institute.
- Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: a free stimulus set of faces and norming data. *Behav. Res. Methods* 47, 1122–1135. doi: 10.3758/s13428-014-0532-5
- Macron, J. L., Susa, K. J., and Meissner, C. A. (2009). Assessing the influence of recollection and familiarity in memory for own- versus other-race faces. *Psychonom. Bull. Rev.* 16, 99–103. doi: 10.3758/pbr.16.1.99
- Maddox, K. B., and Gray, S. A. (2002). Cognitive representations of Black Americans: Reexploring the role of skin tone. *Pers. Soc. Psychol. Bull.* 28, 250–259. doi: 10.1177/0146167202282010
- Meissner, C. A., and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychol. Public Policy, Law* 7, 3–35. doi: 10.1037/1076-8971.7.1.3
- Open Science Collaboration (2017). “Maximizing the reproducibility of your research,” in *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*, eds S. O. Lilienfeld and I. D. Waldman (New York, NY: Wiley), 3–21.
- Oppenheimer, D. M., Meyvis, T., and Davidenko, N. (2009). Instructional manipulation checks: detecting satisficing to increase statistical power. *J. Exp. Soc. Psychol.* 45, 867–872. doi: 10.1016/j.jesp.2009.03.009
- Pew Research Center (2014). *Indians Reflect on their Country & the World*. March, 2014. Washington, DC: Pew Research Center.
- Poolokasingham, G., Spanierman, L. B., Kleiman, S., and Houshmand, S. (2014). “Fresh off the boat?” racial microaggressions that target South Asian Canadian students. *J. Div. High. Educ.* 7, 194. doi: 10.1037/a0037285
- Posner, M. I., and Keele, S. W. (1968). On the genesis of abstract ideas. *J. Exp. Psychol.* 77(3 Pt.1), 353–363.
- Rhodes, G., Brennan, S., and Carey, S. (1987). Identification and ratings of caricatures: implications for mental representations of faces. *Cogn. Psychol.* 19, 473–497. doi: 10.1016/0010-0285(87)90016-8
- Rosch, E. H. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350.
- Shrout, P., and Rodgers, J. (2018). Research and statistical practices that promote accumulation of scientific findings. *Ann. Rev. Psychol.* 69, 487–510.
- Smith, E. E. (1967). Effects of familiarity on stimulus recognition and categorization. *J. Exp. Psychol.* 74, 324–332. doi: 10.1037/h0021274
- Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., and Young, A. W. (2018). Facial first impressions across culture: data-driven modeling of Chinese and British perceivers’ unconstrained facial impressions. *Pers. Soc. Psychol. Bull.* 44, 521–537. doi: 10.1177/0146167217744194
- Tajfel, H., Turner, J. C., Austin, W. G., and Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational Identity: A reader*, 56:9780203505984-16.
- Todorov, A., Olivola, C. Y., Dotsch, R., and Mende-Siedlecki, P. (2015). Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Ann. Rev. Psychol.* 66, 519–545. doi: 10.1146/annurev-psych-113011-143831
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006
- Troje, N. F., and Buülhoff, H. H. (1996). Face recognition under varying poses: the role of texture and shape. *Vision Res.* 36, 1761–1771. doi: 10.1016/0042-6989(95)00230-8
- United Nations Department of Economic and Social Affairs Population Division (2019). *World Population Prospects 2019: Comprehensive Tables (ST/ESA/SER.A/426)*, Vol. 1. New York, NY: United Nations, Department of Economic and Social Affairs, Population Division.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Q. J. Exp. Psychol.* 43, 161–204. doi: 10.1080/14640749108400966
- Wagatsuma, E., and Kleinke, C. L. (1979). Ratings of facial beauty by Korean-American and caucasian females. *J. Soc. Psychol.* 109, 299–300. doi: 10.1080/00224545.1979.9924207
- Wang, H., Han, C., Hahn, A., Fasolt, V., Morrison, D., and Jones, B. C. (2019). A data-driven study of Chinese participants’ social judgments of Chinese faces. *PLoS One* 14:e0210315. doi: 10.1371/journal.pone.0210315
- Wang, X., Tao, Y., Tempel, T., Xu, Y., Li, S., Tian, Y., et al. (2016). Categorization method affects typicality effect: ERP evidence from a category-inference task. *Front. Psychol.* 7:184. doi: 10.3389/fpsyg.2016.00184
- Wells, G. L., and Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Pers. Soc. Psychol. Bull.* 25, 1115–1125. doi: 10.1177/01461672992512005
- Whittlesea, B. W. A., and Leboe, J. P. (2000). The heuristic basis of remembering and classification: fluency, generation, and resemblance. *J. Exp. Psychol.* 129, 84–106. doi: 10.1037/0096-3445.129.1.84
- Winkielman, P., Halberstadt, J., Fazendeiro, T., and Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychol. Sci.* 17, 799–806. doi: 10.1111/j.1467-9280.2006.01785.x
- Wittenbrink, B. (2007). “Measuring attitudes through priming,” in *Implicit Measures of Attitudes*, eds B. Wittenbrink and N. Schwarz (New York, NY: Guilford Press), 17–58.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol. Monogr. Suppl.* 9 (Pt. 2), 1–27. doi: 10.1037/h0025848
- Zebrowitz, L. A. (1996). “Physical appearance as a basis of stereotyping,” in *Foundations of Stereotypes and Stereotyping*, eds N. MacRae,

- M. Hewstone, and C. Stangor (New York, NY: Guilford Press), 79–120.
- Zebrowitz, L. A., and Montepare, J. M. (2008). Social psychological face perception: why appearance matters. *Soc. Pers. Psychol. Compass* 2, 1497–1517. doi: 10.1111/j.1751-9004.2008.00109.x
- Zebrowitz, L. A., White, B., and Wieneke, K. (2008). Mere exposure and racial prejudice: exposure to other-race faces increases liking for strangers of that race. *Soc. Cogn.* 26, 259–275. doi: 10.1521/soco.2008.26.3.259
- Zhao, L., and Bentin, S. (2011). The role of features and configural processing in face-race classification. *Vision Res.* 51, 2462–2470. doi: 10.1016/j.visres.2011.10.001

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2021 Lakshmi, Wittenbrink, Correll and Ma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*