Authors' version of manuscript, as accepted for publication in *Personality and Social Psychology Bulletin*. This version includes the online supplement and participants' instructions. It is not the copy of record and may not exactly replicate the final, authoritative version of the article.

Date of acceptance: 27th October 2017.

# Facial first impressions across culture: data-driven modelling of Chinese and British perceivers' unconstrained facial impressions

Clare A.M. Sutherland[1,2*], Xizi Liu[1], Lingshan Zhang[1], Yingtung Chu[1], Julian A. Oldmeadow[3] & Andrew W. Young[1]

[1] Department of Psychology, University of York, York, UK, YO10 5DD.

[2] ARC Centre of Excellence in Cognition and its Disorders, School of Psychological Science, University of Western Australia, Crawley, Australia, WA 6009.

[3] Department of Psychological Science, Swinburne University of Technology, Melbourne, Australia VIC 3122.

*Corresponding author. Email: clare.sutherland@uwa.edu.au. Phone: +61 8 6488 3240.

**Acknowledgements**

The research was supported by an ESRC doctoral scholarship (ES/I900748/1), postdoctoral funding from the ARC Centre of Excellence in Cognition and its Disorders (CE110001021) to CS, and an ARC Discovery grant (DP170104602) to CS and AY. Funding sources had no influence on the research. Portions of the research appeared in CS's doctoral thesis and in LZ's Masters thesis, both at the University of York. We are very grateful to Mark Hymers for programming help, and Alice Towler, Xin Kang, Xiaoqian Yan, Rosa Kwok, Hollie Byrne, Josh Spowage, Jemma Collova and Nichola Burton for research assistance.

Abstract

People form first impressions from facial appearance rapidly, and these impressions can have considerable social and economic consequences. Three dimensions can explain Western perceivers' impressions of Caucasian faces: approachability, youthful-attractiveness and dominance. Impressions along these dimensions are theorized to be based on adaptive cues to threat detection or sexual selection, making it likely that they are universal. We tested whether the same dimensions of facial impressions emerge across culture by building data-driven models of first impressions of Asian and Caucasian faces derived from Chinese and British perceivers' unconstrained judgments. We then cross-validated the dimensions with computer-generated average images. We found strong evidence for common approachability and youthful-attractiveness dimensions across perceiver and face race, with some evidence of a third dimension akin to capability. The models explained ~75% of the variance in facial impressions. In general, the findings demonstrate substantial cross-cultural agreement in facial impressions, especially on the most salient dimensions.

*Keywords:* "impression formation" "face perception" "person perception" "social cognition" "cross cultural"

# Facial first impressions across culture: data-driven modelling of Chinese and British perceivers' unconstrained facial impressions

## Are Facial Impressions Universal?

When meeting someone for the first time, one of the most salient sources of information we have is their face (Bruce & Young, 2012; Calder, Rhodes, Johnson, & Haxby, 2011). A stranger's face can offer reliable cues to their gender, ethnicity and age (Bruce & Young, 2012), but perceivers typically go further than these relatively objective judgments, and also readily infer attractiveness and character traits from facial cues (Todorov, Mandisodza, Goren, & Hall, 2005). These facial impressions predict critical real-world decisions, such as whether to lend money to the target (Duarte, Siegel, & Young, 2012), whether to allow them to win a court case (Zebrowitz & McDonald, 1991) and even whether to elect them to political office (Rule et al., 2010; Todorov et al., 2005). Given these important real-life consequences, it is vital that we have a clear theoretical understanding of how people form these impressions. This aim is especially timely, since impressions from facial photographs are increasing in importance with the rise of global online communication.

Recently, researchers have characterized the key dimensions underlying facial impressions for Western perceivers (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Walker & Vetter, 2009). In an influential study, Oosterhof and Todorov (2008) used a principal components analysis to reduce a wide variety of facial trait ratings into key underlying dimensions of trustworthiness and dominance (Oosterhof & Todorov, 2008). Critically, facial ratings were selected by sampling from participants' unconstrained judgments, thereby building a data-driven model of facial impressions. Oosterhof and Todorov (2008) theorized that trustworthiness functions as an assessment of the target's intentions (good or bad), and that trustworthy inferences are based on an overgeneralization of

facial cues resembling emotional expression. Dominance impressions function to predict a target's ability to carry out these intentions, and are based on an overgeneralization of facial cues to physical strength. Together, these two dimensions represent the evaluation of threat, theorized to have a long evolutionary background due to the importance of threat perceptions in our ancestral survival (Oosterhof & Todorov, 2008). Recently, Sutherland and colleagues (2013) extended this model by using highly variable images of faces, finding an additional dimension, 'youthful-attractiveness', which linked the perception of increasing age with decreasing attractiveness (see also Wolffhechel et al., 2014). The authors theorized that this dimension could serve sexual selection functions, also potentially with a long evolutionary history.

These dimensional models have formed an influential new theoretical framework for research on facial impressions, as well as stimulating considerable interdisciplinary research spanning visual face perception and social cognition (for a review, see Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). However, these models have so far only been built from impressions shown to be important for Western perceivers: there is not a model of facial impressions derived from non-Western perceivers. This is a serious omission, because the hypothesized evolutionary basis of these models implies that they represent universal dimensions of facial judgment, in turn constituting a powerful pan-cultural aspect of social cognition. This assumption has never been empirically verified.

In support of the suggestion that these impressions may form universal dimensions of social cognition, studies of conceptual (non-facial) impressions have found considerable cross-cultural universality. In particular, the concepts of morality, competence and attractiveness appear in lexicons across distinct language groups, strongly suggesting that person perception attributes are universal (Saucier, Thalmayer, & Bel-Bahar, 2014). Likewise, the semantic differential model of human attitudes consists of three key dimensions

which bear striking resemblance to the three dimensions of facial impressions: evaluation (cf. approachability or trustworthiness), potency (cf. dominance), and activity (cf. youthful-attractiveness) and these conceptual dimensions have also been replicated across cultures (Osgood, 1964; Saucier et al., 2014). There is also high cross-cultural agreement in which attributes emerge as important for judging social groups, with warmth and competence dimensions appearing across a large number of cultures (Fiske, Cuddy, & Glick, 2007). Together, these findings suggest that these dimensions represent a fundamental aspect of human social cognition that appears in every cultural group studied so far. It is plausible that similar dimensions will be found across cultures for *facial* impressions of real people.

However, the evidence for cross-cultural agreement for facial impressions is currently debated, even for more basic judgments of emotional expression from faces (Elfenbein & Ambady, 2002; Jack, Garrod, Yu, Caldara, & Schyns, 2012; Jack & Schyns, 2017). A number of studies have found considerable cross-cultural agreement in facial trait judgments (e.g. Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Secord & Bevan, 1956; Walker, Jiang, Vetter, & Sczesny, 2011; Zebrowitz et al., 2012). Particularly striking is the finding that American perceivers generally agreed in their facial impressions with perceivers from the Tsimane' people, who live in isolation in the Bolivian rainforest (Zebrowitz et al., 2012). Yet, recent studies have claimed that facial judgments of emotional expressions are not culturally universal, with Asian perceivers having different mental representations of facial emotional expression than Western perceivers (Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Jack et al., 2012; although see Elfenbein & Ambady, 2002; Yan, Andrews, & Young, 2016). These more recent findings cast doubt on the claim that facial impressions are universal, because these judgements depend on emotional expression to a large extent (e.g. Oosterhof & Todorov, 2008). Moreover, even the *same* emotional expressions may lead to different impressions across culture, depending on local cultural norms. For example, smiling is seen

as intelligent in Germany and China, but unintelligent in Iran (Krys, Hansen, Xing, Szarota, & Yang, 2013).

Crucially, there has not yet been a direct test of the universality of dimensions of facial impressions. This test is missing because previous cross-cultural studies were designed to target specific hypotheses about pre-specified traits or facial cues. It is clearly informative that perceivers across cultures can agree on their facial impressions if directly asked. However, this approach does not tackle the more fundamental claim that the key dimensions found in studies of Westerners are also the most important dimensions found in other cultures.

To address this question, we used a data-driven approach to provide the first strong test of whether the dimensions of facial first impressions are culturally universal, by building the first model of non-Western facial impressions. To achieve this, we sampled unconstrained first impressions of own-race faces by perceivers from a non-Western culture, and then used these traits to derive our models, following Oosterhof and Todorov's (2008) original approach with American perceivers. Data-driven approaches are increasingly being used to answer fundamental questions in human social perception (Adolphs, Nummenmaa, Todorov, & Haxby, 2016; Jack & Schyns, 2017; Todorov, Dotsch, Wigboldus, & Said, 2011), but this is the first time they have been applied to understand facial impressions across cultures. A data-driven approach is critical to answering the question of which dimensions subserve impressions in non-Western cultures. Otherwise, research will necessarily prevent facial impression dimensions *other* than the Western dimensions from emerging.

**Building Facial Impression Models for Chinese Perceivers**

To test whether the dimensions of facial first impressions are culturally universal, we built models of Chinese perceivers' unconstrained impressions of own- (Asian) and other-race (Caucasian) faces. We also created models of British perceivers' impressions of the same

faces, allowing us to compare the models derived from traits used by participants in each culture.

We decided to examine Chinese perceivers' facial impressions for two reasons. First, examining Chinese impressions is intrinsically interesting since China is the world's largest country by population, with an estimated 19% of the world's population; more than the USA, Oceania and Europe together (World Population Clock, 2014). Second, examining Chinese perceivers offers a strong test of the potential universality of facial impressions because there are substantial relevant cultural differences between East Asia and the West (Hofstede, 1980; Oyserman, Coon, & Kemmelmeier, 2002). Specifically, East Asian perceivers, especially from China, are characterized as being more collectivist (having interdependent values) than Western participants, who are characterized as being more individualistic (having independent values: Hofstede, 1980; Oyserman et al., 2002). Importantly, these East Asian and Western cultural differences have been found to affect face perception and resulting social judgments (Chen, Jing, & Lee, 2012; Jack et al., 2012; Wheeler & Kim, 1997). Moreover, since East Asian cultures promote perception based on social or situational information rather than individuating information (Maass, Karasawa, Politi, & Suga, 2006; Morris & Peng, 1994; Nisbett, Peng, Choi, & Norenzayan, 2001), forming impressions of faces based on individual traits may simply be less important to East Asian perceivers. We therefore examined whether Chinese and Western perceivers form impressions of traits to the same extent.

To establish which facial impressions are important in a non-Western population, in Study 1 we collected unconstrained impressions of own-race faces from Chinese and British perceivers. We then selected the most frequently mentioned Chinese and British facial impressions to build Chinese and British models of own- and other-race facial impressions in Study 2. We then used computer-averaged images to cross-validate our models in Study 3.

## Study 1

### Methods 1

Twenty Chinese participants (mean age: 21.5 years, 10 male) were tested in Guangdong, China, and 20 British participants (mean age: 22.5 years, 10 male) in York, UK. Chinese participants had not lived in any Western countries (including the UK) for longer than a year, and likewise for British participants and East Asian countries (including China). In all studies, sample size was chosen beforehand, based on previous research (Oosterhof & Todorov, 2008; Sutherland, Young, Mootz, & Oldmeadow, 2015). Participants were students and they provided informed consent to procedures approved by the University of York Psychology Department Ethics Committee.

#### Face Images 1

Thirty female and 30 male Caucasian faces were randomly selected from an existing database of 1,000 highly variable, naturalistic images of faces taken from the internet ('ambient images'; Santos & Young, 2005; see Jenkins, White, Van Montfort, & Burton, 2011). For the current study, we collected a further 30 female and 30 male adult Asian face photographs from the internet using Chinese browsers (150 pixels in height, with preserved aspect ratio). We used ambient images so that we could index the range of potential photographic facial cues that perceivers would be exposed to (for example, while browsing online), and to compare to the previous finding that three dimensions underlie British perceivers' impressions of ambient images (Sutherland et al., 2013; Vernon, Sutherland, Young, & Hartley, 2014). Face sets were deliberately allowed to vary on many potential facial cues to social impressions, including pose, expression, lighting and facial accessories (as in Santos & Young, 2005; Sutherland et al., 2013; Vernon et al., 2014).

Seven Asian colleagues (from China, Singapore and Korea) screened the new Asian image set (also in Study 2) so that images depicted non-famous adults, plausibly represented Chinese faces, and varied as much as the Caucasian set. For simplicity, we refer to these faces as "Asian" since we cannot confirm the nationality or ethnic background of the people depicted; they were simply chosen as appearing Chinese to people from the broader region. The Caucasian set was similarly screened (Santos & Young, 2005).

**Procedure 1**

Participants were tested in quiet locations using a laptop running custom Python code to display and collect Simplified Chinese script. Each set of 60 faces was divided into five smaller sets of 12, to ensure that participants were not fatigued and to keep responses genuine. Face images from one set were shown to participants sequentially in random order with a blank text box underneath. Two additional faces were shown as a practice, and in a second block, participants saw 12 other faces with a context label for a separate study (not analysed). Participants were only shown own-race faces in Study 1, for comparability to previous models (Oosterhof & Todorov, 2008), and so that Study 2 models would be based on within-culture impressions.

Across all three studies, Chinese perceivers were tested in Mandarin by Asian experimenters and British perceivers in English by Caucasian experimenters; written Chinese materials were translated into Simplified Chinese by a native speaker and then back-translated into English by a second native speaker to ascertain equivalence of meaning (see supplementary instructions). Participants were told that this study was examining first impressions and were asked to type in anything that came to mind on viewing the face, no matter how silly or socially inappropriate. Testing took around 30mins.

**Categorising facial impressions**

Our data were participants' descriptions of the own-race faces (in Simplified Chinese or in English), split into single words or phrases (e.g., "not friendly"). Chinese descriptions were translated into English by two native Chinese speakers. Where Chinese concepts mapped onto multiple potential English concepts, or where translators disagreed, we used a slash (e.g. "热情" as "passionate/enthusiastic"). Where Chinese concepts formed compound words, we used a tilde (e.g. "和善" as "kind~and~gentle").

Two native Chinese colleagues independently categorised the content of the (untranslated) Chinese data and two native British colleagues independently categorised the British data, with the first author supervising the groups for consistency (table 1). Traits referred to a description of long-lasting character or personality (e.g. *"intelligent"*) and emotions referred to feeling states (e.g. *"angry").* Appearance words included any description of what the target looked like (e.g. *"haggard")*. Sex and age included words that unambiguously indicated the faces' sex (e.g. *"male", "she", "housewife"*) or age (e.g. *"old", "20-30s", "retired").* Categories were not mutually exclusive. There was high inter-rater consistency across the three coders for the Chinese and British data (all pairwise kappa values ≥ .67, except for Chinese emotion, which was >.42). Coders afterwards resolved any discrepancies.

## Results and Discussion 1

Chinese perceivers described the Asian faces with 601 words and phrases in total (2,913 Simplified Chinese characters), whereas British perceivers described the Caucasian faces with 1,178 words and phrases (3,295 English words). Note that Chinese descriptions were not necessarily less rich than the British descriptions, since Chinese concepts often held compound English meanings.

Strikingly, the Chinese and British participants produced very similar profiles of facial impressions (table 1), with both cultures focusing mostly on targets' traits, sex and age, and finally their appearance (with 49% of the Chinese participants' and 41% of the British participants' appearance descriptions focusing on attractiveness).

Table 1

*Number and type of unconstrained impressions of own-race faces for Chinese and British perceivers. Coded categories are not mutually exclusive. 'Total' represents total coded.*

| | Overall | Trait | Appearance | Sex | Age | Emotion | Total |
|---|---|---|---|---|---|---|---|
| Chinese | 601 | 30% | 20% | 14% | 10% | 3% | 66% |
| British | 1178 | 31% | 14% | 16% | 8% | 4% | 66% |

Figure 1 offers a visual display of the (translated) trait descriptions produced by the Chinese and British participants (see table S1-2 for unique participant frequencies, figure S1 for a visual depiction of all categories). Higher frequency descriptions are depicted in larger font (Sutherland, Young, et al., 2015). While the British descriptions cluster around a few main words ("friendly", "kind", "intelligent", and "warm"), the Chinese descriptions are more variable. However, common themes emerge, with both cultures frequently mentioning traits relating to approachability or interpersonal warmth, including "cheerful/outgoing", "benevolent", "kind" and "friendly" (with 59% of both British and Chinese trait descriptions focusing on warmth). This pattern supports the suggestion that warmth traits may be perceivers' primary concern when judging others (Fiske et al., 2007). Both cultures also mentioned competence-related traits, such as "capable", "capable/experienced", "intelligent" and "wise" (with 34% of British and 27% of Chinese trait descriptions focusing on competence). Interestingly, neither British nor Chinese participants spontaneously mentioned dominance (one Chinese participant did mention 'strong'), despite this trait's importance in

previous models (Oosterhof & Todorov, 2008; Sutherland et al., 2013; see the General Discussion). Both groups frequently mentioned attractiveness, coded as an appearance description (figure S1), and the Chinese participants mentioned the word “wretched”, meaning someone who is unattractive in character and appearance (Baidu.com, 2014). In summary, although perceivers used culture-specific words, similar approachability, competence and attractiveness concepts appeared from unconstrained impressions across culture.


*Figure 1.* Unconstrained trait impressions of own-culture faces for A) Chinese participants (translated into English) and B) British participants. Larger font size represents more frequent descriptions. Only single words and short phrases (less than five words) are included so key concepts can emerge. Word clouds from www.wordle.net.

## Study 2

### Methods 2

Since Study 1 showed that Chinese perceivers spontaneously make trait inferences when asked for unconstrained impressions, in Study 2 we modelled the key dimensions underlying Chinese facial impressions. We asked new Chinese participants to rate Asian and Caucasian faces on the most frequently mentioned attributes from Study 1. We then used factor analysis to reduce the Chinese perceivers' ratings into key dimensions, thereby building models of their impressions of Asian and Caucasian faces. Although we were open to unique Chinese dimensions emerging, we predicted that a warmth (approachability) dimension would appear, given Study 1 results and the centrality of warmth in theories of person perception (Abele & Bruckmüller, 2011; Fiske et al., 2007; Oosterhof & Todorov, 2008).

We also wanted to compare how similar Chinese facial impressions are to Western (British) impressions. Therefore, we also built equivalent models of British participants' impressions of Asian and Caucasian faces. Finally, we quantified cross-cultural similarity by correlating face scores on trait dimensions across British and Chinese perceiver models.

#### Participants 2

120 Chinese participants (mean age: 23.6 years, 60 female) and 120 British participants (mean age: 20.6 years, 60 female) were tested at the University of York. Chinese participants had been in the UK for an average of 1.47 years (none >6 years). British participants were born and had lived in the UK for most of their lives (none had visited China). Two additional Chinese participants were tested but excluded before analysis (one was not raised in China; another was distracted while rating). Twelve additional British participants were tested but excluded before analysis (five had mostly lived outwith the UK,

five were not Caucasian and two did not finish due to computer error). Testing took around an hour.

**Face Images 2**

We randomly selected 500 Caucasian faces (250 male) from an existing set of 1,000 highly varied ambient image photographs (Santos & Young, 2005). We also collected a further 440 Asian faces from the internet using Chinese browsers to create a full set of 500 Asian ambient image faces (250 male; database screening as Study 1).

**Procedure 2**

The 1,000 faces were rated on the attributes most frequently mentioned by own-race perceivers in Study 1. Frequencies were calculated by counting together positive and negative occurrences of the same root word (e.g. “not intelligent” and “intelligence” were counted as instances of “intelligent”) but did not include multiple occurrences of the same word from the same participant, to avoid biasing rating choices from idiosyncratic trait use (tables S1-S2). Ratings of age, masculinity and attractiveness were also collected, since these were also frequently mentioned by participants from both cultures in Study 1 (table 1).

British participants rated all faces from 1 (not very) to 7 (very) on either: friendly, kind, intelligent, nice, warm, quiet, shy, funny, sweet, attractive, age (young to old), or masculinity (from very feminine to very masculine). Chinese participants rated all faces from 1 (not very) to 7 (very) on either: 开朗 (cheerful/outgoing), 严肃（serious), 慈祥 (benevolent), 和蔼 (affable), 和善 (kind~and~gentle), 热情 (passionate/enthusiastic), 干练 (capable/experienced), 圆滑 (diplomatic), 猥琐 (wretched), 吸引力 (attractive), age (年轻人 (young) to 老年人(old)), or masculinity (very 女性化的 (feminine) to very 男性化的

(masculine)). Chinese traits came directly from Study 1, were rated in Simplified Chinese, and are only translated into English here for convenience.

Participants were tested in a quiet room on a PC with PsychoPy (version 1.76: Peirce, 2007) and were told that they were taking part in a study of first impressions (see supplementary instructions). On each trial, participants viewed one face with the rating scale underneath. Participants pressed the number key that corresponded with their rating and the next face photograph appeared after an ISI of 750ms. Participants rated own-race faces in a first block, and then other-race faces in a second block. We deliberately blocked trials in Study 2, since our main aim was to examine Chinese impressions of own-race faces. This design offered the best test of genuine Chinese first impressions to Asian faces, because participants were unaware that other-race faces would be rated later, or that the study was cross-cultural. It also allowed us to directly compare the own-race Chinese model with previous models without a cross-cultural aspect (e.g. Oosterhof & Todorov, 2008; Sutherland et al., 2013; note that Study 3 intermixed face race for generalizability). Participants only rated one attribute, to avoid carryover effects (Rhodes, 2006).

## Results and Discussion 2

### Modelling facial impressions

We decided *a priori* to collect data from ten participants for each trait, rather than increasing the sample size until the reliability was acceptable, in order to compare impression agreement across perceiver groups, and because increasing the number of items will increase alpha without necessarily increasing quality (Cortina, 1993). Reliabilities were good for the majority of traits for both face and participant groups (alphas above .7; tables S1-S2). There was also high cross-cultural agreement across British and Chinese participants at the trait level (figure S2).

For all combinations of face and participant culture, Bartlett's test of sphericity indicated that a factor analysis was appropriate: all $X^2 > 4{,}847$, all $p < .001$. To determine dimensionality, four criteria were utilized: a scree test and Kaiser's criterion on the unreduced correlation matrices (tables S3-S4; figure S5; figure S6 presents scree tests using the reduced matrix for comparison; Fabrigar & Wegener, 2012; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Kline, 1994), a conservative parallel analysis using the 95% percentile (table S5), and a MAP test; see Sutherland et al., 2013; Vernon et al., 2014). Where criteria disagreed, we followed Kaiser's criterion to best compare to leading Western models (e.g. Oosterhof & Todorov, 2008; Walker & Vetter, 2009), unless dimensions proved unstable across analyses. We used principal factor analysis with direct oblimin rotation to build the final models, and used the structure matrix to interpret the dimensions, ignoring loadings below .3 (based on Kline, 1994).

We verified that the same structure emerged for all models when a principal components analysis (PCA) with orthogonal varimax rotation was used (given that leading Western models use PCA: Oosterhof & Todorov, 2008; Walker & Vetter, 2009). We also used PCA to estimate the variance explained by the dimensions, which is not possible with an oblique factor analysis. For ease of comparison, and because factor direction is arbitrary, we always described dimensions in the same direction across models (e.g. as youthful-attractiveness).

**British facial impressions models**

We first built British facial impression models (table 2; figure 2 visualizes the dimensions through computer-averaging; see figure S3 for original colour image). For both the Caucasian and Asian faces, Kaiser's criterion found three dimensions, the scree test found two or possibly four dimensions, and the parallel analysis and MAP analysis found two

dimensions (figures S5-S6, table S5). We followed Kaiser's criterion for comparison with previous models; however, a three-dimensional solution for the Asian faces produced a third dimension that was not stable across analyses, so we did not interpret this solution further. We therefore built a three-dimensional British model for the Caucasian faces and a two-dimensional British model for the Asian faces. Orthogonal PCA models were highly comparable (table S6) and explained most of the variance in the original British impressions: the three Caucasian and two Asian face dimensions explained 80% and 70% of the variance respectively.

Table 2

*Dimensions of British impressions of Caucasian and Asian faces (principal axis factor analysis, structure matrices). These can be interpreted as akin to correlations between the factors and variables. Factor loadings ≥ .3 appear in bold.*

| | Caucasian face dimensions | | | Asian face dimensions | |
|---|---|---|---|---|---|
| | Approach. | Youth-Attract | Capability | Approach. | Youth-Attract |
| Friendly | **.93** | .19 | **.53** | **.97** | .13 |
| Nice | **.88** | **.31** | **.59** | **.93** | .23 |
| Warm | **.88** | .26 | **.55** | **.91** | .23 |
| Kind | **.86** | .22 | **.61** | **.88** | .13 |
| Sweet | **.84** | **.38** | **.62** | **.88** | .26 |
| Quiet | **-.92** | -.04 | -.16 | **-.81** | -.13 |
| Funny | **.74** | **-.34** | .24 | **.71** | **-.34** |
| Shy | **-.55** | -.17 | -.08 | **-.43** | .01 |
| Age | .11 | **-.77** | **.49** | -.00 | **-.79** |
| Attractive | .16 | **.85** | -.11 | .22 | **.95** |
| Masculine | -.15 | **-.62** | -.09 | **-.31** | **-.70** |
| Intelligent | .21 | -.16 | **.53** | -.09 | -.04 |
| Variance explained, varimax PCA | 49% | 20% | 11% | 48% | 22% |

The British model for the Caucasian faces showed three dimensions of approachability, youthful-attractiveness and capability (table 2), broadly agreeing with previous findings (Sutherland et al., 2013; Wolffhechel et al., 2014). The British model for

the Asian faces showed two dimensions: approachability and youthful-attractiveness (table 2). Overall, the first two dimensions for the British participants demonstrate strong similarity across face race and with previous Western dimensions, with the approachability dimension including warmth and trustworthiness-related traits (e.g. friendliness) and the youthfulness-attractiveness dimension, including youth, attractiveness and femininity (table 2). However, the British capability dimension for the Caucasian faces differs from the dominance dimension found in previous research (Oosterhof & Todorov, 2008; Sutherland et al., 2013), as it includes intelligence as well as social attributes (e.g. kindness). Model dimensions were not highly inter-correlated (table S8; all $r < .36$). Finally, the alternative two-dimensional Caucasian solution showed largely identical approachability and youthful-attractiveness dimensions, with intelligence failing to load on either (< .30).

**Chinese facial impressions models**

We then built the first Chinese models of facial impressions (table 3; figure 2 visualizes the dimensions; see figure S3 for original colour image). For the Caucasian faces, the scree test found one or three dimensions, while all other criteria found three dimensions; for the Asian faces, all criteria found four dimensions (figure S5-S6, table S5). We therefore built a three-dimensional Chinese model for the Caucasian faces and a four-dimensional Chinese model for the Asian faces. Orthogonal PCA models were highly comparable and explained most of the variance in the original Chinese impressions (table S7): the three Caucasian face and four Asian face dimensions explained 77% and 84% of the variance respectively.

Table 3

*Dimensions of Chinese impressions of Asian and Caucasian faces (principal axis factor analysis, structure matrices). These can be interpreted as akin to correlations between the factors and variables. Factor loadings ≥ .3 appear in bold.*

| | | Asian face dimensions | | | | Caucasian face dimensions | | |
|---|---|---|---|---|---|---|---|---|
| Chinese ratings | Translations | Approach. | Youth | Attract. | Capability | Approach. | Youth-Attract | Capability |
| 热情 | Passion./enthusiastic | **.95** | -.12 | .29 | .10 | **.95** | .15 | .07 |
| 开朗 | Cheerful/outgoing | **.94** | -.06 | .22 | .11 | **.94** | .10 | .08 |
| 严肃 | Serious | **-.93** | -.08 | -.21 | .03 | **-.92** | -.12 | .03 |
| 和善 | Kind~and~gentle | **.86** | -.20 | **.47** | .08 | **.93** | .21 | .25 |
| 和蔼 | Affable | **.75** | **-.41** | **.52** | .08 | **.90** | .21 | **.37** |
| 慈祥 | Benevolent | **.49** | **-.78** | **.30** | -.02 | **.72** | -.17 | **.62** |
| 年轻人/老年人 | Age | -.13 | **-.87** | -.20 | -.19 | .14 | **-.66** | **.58** |
| 猥琐 | Wretched | -.17 | .09 | **-.87** | -.21 | **-.41** | **-.63** | **-.49** |
| 女性化的/男性化的 | Masculine | **-.31** | -.20 | **-.53** | .10 | -.19 | **-.51** | -.03 |
| 吸引力 | Attractive | .13 | **.49** | **.51** | **.44** | .04 | **.68** | .16 |
| 干练 | Capable/Experienced | -.11 | .13 | .05 | **.82** | .02 | .13 | **.49** |
| 圆滑 | Diplomatic | **.32** | .09 | .20 | **.74** | **.58** | .01 | **.54** |
| Variance explained, varimax PCA | | 37% | 17% | 15% | 15% | 44% | 17% | 16% |

Chinese models for both Asian and Caucasian faces showed three dimensions of approachability, youthful-attractiveness and (social) capability that were very similar to the British Caucasian model, demonstrating overall strong cross-cultural similarity (table 3). However, the Chinese model for Asian faces was more differentiated, with four dimensions instead of three (table 3).

The first Chinese dimension for the Asian and Caucasian faces was clearly approachability, including traits such as passionate/enthusiastic, cheerfulness-outgoing, kindness, and affability (table 3). The second Chinese dimension for the Caucasian faces was clearly youthful-attractiveness, including attributes such as (decreasing) wretchedness, youth, femininity, and attractiveness. However, for the Asian faces, age and attractiveness split into separate dimensions (table 3), with the second dimension including age and benevolence (a trait linked to age in China: Baidu.com, 2015), and the third dimension including

attractiveness as opposed to wretchedness (i.e. decreased interpersonal attractiveness: Baidu.com, 2014). These culturally-specific concepts may have led to the second and third dimensions separating for Chinese impressions of own-race faces. Finally, the last Chinese dimension for the Caucasian and Asian faces looked similar to the British capability dimension, including diplomacy and capability/experience (table 3). Again, this dimension bears only slight resemblance to the dominance dimension found previously (Oosterhof & Todorov, 2008; Sutherland et al., 2013), with high loadings from pro-social attributes (e.g. benevolence, especially for Caucasian faces). Model dimensions were not highly inter-correlated (table S8; all $r < .33$). Finally, an alternative Asian three-dimensional solution showed approachability, youthful-attractiveness and capability dimensions. Alternative two-dimensional solutions showed approachability and youthful-attractiveness dimensions, with capability either only weakly cross-loading on both dimensions (Asian faces < .40) or failing to load (Caucasian faces < .30).

**Visualizing facial cues to Chinese and British impressions**

An advantage to using highly varied face images is that image averaging techniques can be used to visualize the cues that are consistently present in faces that differ on the underlying dimensions (Sutherland, Rhodes, & Young, in press). Figure 2 visualizes the facial cues subserving the Chinese and British dimensions by averaging together the twenty highest and lowest scoring faces on each dimension using Psychomorph (Tiddeman, Burt, & Perrett, 2001; Sutherland, 2015 presents detailed guidance). Strong visual similarities appear across face race and perceiver culture, especially for approachability and youthful-attractiveness dimensions (figure 2; figure S3). Facial approachability cues clearly include smiling and femininity. Youthful-attractiveness cues include decreased age, femininity, and skin smoothness. However, capability cues diverge across face race: Caucasian facial cues include

increased age and darker skintone, while Asian facial cues include decreased age, lighter skintone and masculinity (figure 2).

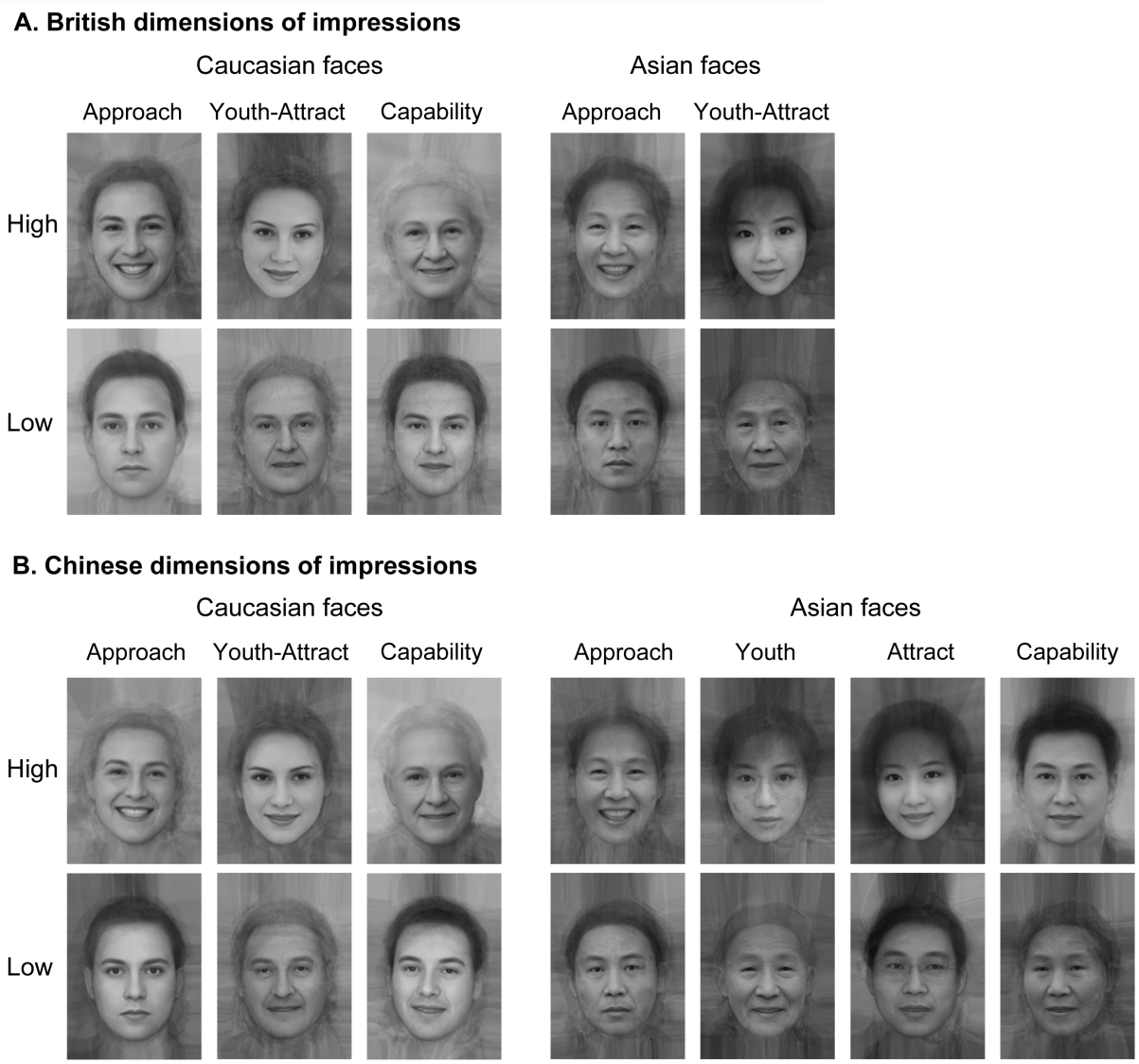


*Figure 2.* Faces created to lie high or low on the A) British and B) Chinese models of facial impressions of Asian and Caucasian faces. Each face is an average of the 20 highest or lowest scoring Asian and Caucasian face photographs on each of the dimensions in the Chinese and British models. See figure S3 in the Online Supplement for the original colour version.

**Cross-cultural model similarity**

The approachability and youthful-attractiveness dimensions created by our data-driven trait sampling method were stable across face race and perceiver culture, and there was some

evidence for a third dimension across culture, capability, although this dimension varied more. To test these claims, we quantified the cross-cultural similarity between the Chinese and British perceiver models. We calculated factor scores for each face on each dimension, and then correlated these scores across models, at the face level. These correlations demonstrate significant and high consistency in the underlying dimensions across perceiver culture (table 4; figures S7-S8 visualise the facial cues). This consistency is impressive given that the faces were rated on completely different traits, in different languages, and since factor scores themselves only approximate dimensions. In particular, the correlation for the Chinese and the British approachability dimensions is close to ceiling (both $r > .93$, $p < .001$).

Table 4

*Agreement between Chinese and British dimensions, Caucasian and Asian faces (as measured by Pearson's r correlations across factor scores, regression method). The highest cross-cultural correlation between dimensions (i.e. in each row) is highlighted in bold.*

| | Caucasian faces | | | Asian faces | | | |
|---|---|---|---|---|---|---|---|
| | Chinese Approach | Chinese Youth-Attract | Chinese Capability | Chinese Approach | Chinese Youth | Chinese Attract | Chinese Capability |
| British Approach. | **.94**** | .07 | .22** | **.93**** | -.23** | .37** | .03 |
| British Youth-Attract | .18** | **.87**** | -.15** | .20** | **.70**** | .52** | .26** |
| British Capability | .51** | -.07 | **.69**** | - | - | - | - |

** $p < .01$, $n = 500$

**Pancultural dimensions**

Given the overall correspondence across perceiver culture, we included the Chinese and British impressions together to form pancultural models for Asian, Caucasian and all faces together. For the pancultural models, scree tests returned two or four dimensions; Kaiser' criterion and the parallel analysis returned four dimensions, and the MAP analysis returned five or six dimensions (figures S5-S6, table S5). However, the fourth dimension

mainly relied on high loadings from masculinity, so we refrained from interpreting this dimension further (Kline, 1994).

The three-dimensional pancultural models formed clear dimensions of approachability, youthful-attractiveness and capability (figure 3 visualizes the facial cues subserving these dimensions; see figure S4 for colour version; table S9 presents the full models; table S8 presents inter-dimension correlations, all $r < .24$). Alternative two-dimensional solutions produced approachability and youthful-attractiveness dimensions, without capable-experienced or intelligent loading (both < .32).

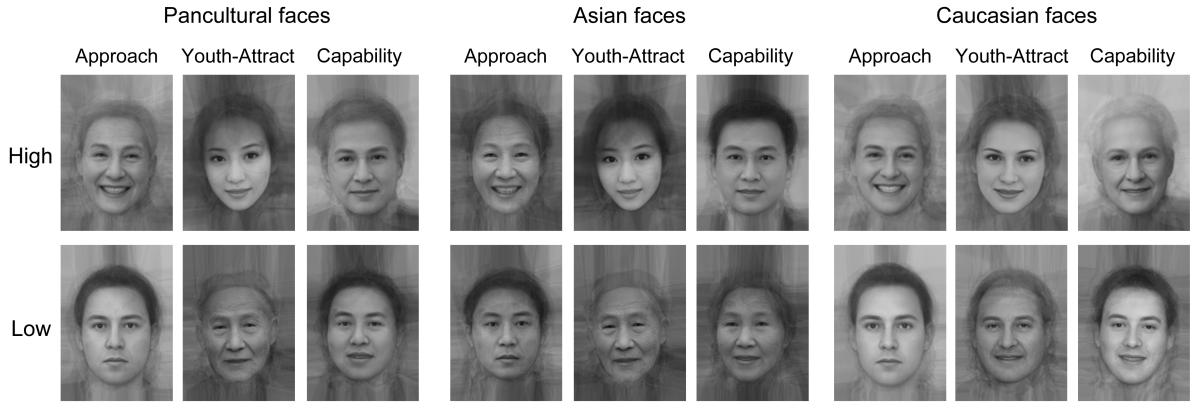


*Figure 3.* Pancultural, Asian and Caucasian face averages made from the 20 faces which scored most and least highly on pancultural dimensions of facial impressions. See figure S4 in the Online Supplement for original colour version.

## Study 3

### Method 3

Study 3 aimed to cross-validate the Study 2 dimensions by collecting ratings of the high and low average faces derived from these (figures 2-3), from new participants recruited in China and in the UK. We collected ratings of these average faces on approachability, age,

attractiveness and capability, to index the proposed factor labels. We intermixed rather than blocked face race, in order to generalize to a different design.

**Participants 3**

Forty Chinese participants (mean age: 23.0 years, 23 female) in Chengdu, Sichuan region of China, and 44 British participants (mean age: 20.2 years, 23 female) were tested in York, UK. Chinese participants had not lived in any Western countries (including the UK) for longer than a year, and likewise for the British participants and East Asian countries (including China). Participants were recruited via educational networks and tested online with Qualtrics (Provo, UT, 2017). Before analyses, we excluded 26 additional participants who dropped out, 8 participants who were not Chinese/British Caucasian, 5 participants who had lived for longer than a year outside China/the UK, and 8 participants who asked us not to use their data.

**Face Images 3**

Participants saw 42 average faces, pairs of which indexed the high and low ends of the nine pancultural, seven Chinese, and five British model dimensions from Study 2 (taken from figures 2-3). Participants first saw two neutral practice faces (one of each race, created by averaging all Asian or Caucasian average images).

**Procedure 3**

Participants rated all faces on four impressions chosen to index proposed labels for the dimensions: from 1 (not very) to 7 (very) approachable (平易近人), attractive (吸引力), capable (干练) or from young (年轻人) to old (老年人; see supplementary instructions). Faces were presented one at a time in random order, and participants rated all faces on one

trait within a block. Face order within a block and block order were randomised across participants. Testing took around 25 minutes.

## Results and Discussion 3

We reversed the age ratings to 'youth' to align with attractiveness. For simplicity and to reduce the number of comparisons, we calculated the difference between ratings given to pairs of high and low morphed faces on each dimension, for each participant and trait. Cross-cultural agreement between Chinese and British perceivers at the face level was high (all $r < .67$, $p < .001$, $n = 18$), except for capability for Caucasian faces, which did not show cross-cultural agreement ($r = .41$, $p = .089$, $n = 18$; table S10).

We tested both criterion and divergent validity, focusing on the pancultural models because these allowed interactions between participant culture and face race to be directly tested, unlike the individual cultural models, which differed in dimensionality; and because the pancultural models were most likely to be stable, being based on the largest number of traits and participants. However, analyses of the four culture-specific models produced essentially identical conclusions (see Online Supplement text, figure S9, table S11).

### Criterion validity

We examined whether the high and low average faces on each dimension differed on the predicted traits, using one-sample t-tests against zero. For example, we tested the approachability face dimension using the approachability ratings, and so on. We examined British and Asian participants and the three face groups separately (pancultural faces, Asian faces, Caucasian faces).

All comparisons were significant except for the capability dimension for the Caucasian faces with British participants (table 5). A sign test showed that the overall pattern

of results was significantly different from chance ($p < .0001$). Thus, there was strong overall criterion validity, except for Caucasian faces on capability for British participants. Tests of the four cultural models gave identical conclusions (see Online Supplement text, table S11).

Table 5

*Mean differences in trait ratings across pairs of high and low average faces on each of three dimensions.*

| British participants | Faces | Mean high – low | SD high – low | *d* |
|---|---|---|---|---|
| Approachability Dimension: *approachability ratings* | Pancultural | 3.05** | 1.75 | 1.74 |
| | Asian | 3.20** | 1.72 | 1.86 |
| | Caucasian | 3.02** | 1.50 | 2.01 |
| Youth-Attract Dimension: *youth ratings* | Pancultural | 5.32** | 1.12 | 4.75 |
| | Asian | 5.27** | 1.30 | 4.05 |
| | Caucasian | 3.25** | 0.92 | 3.53 |
| Youth-Attract Dimension: *attractiveness ratings* | Pancultural | 3.82 ** | 1.70 | 2.25 |
| | Asian | 4.14** | 1.61 | 2.57 |
| | Caucasian | 4.30** | 1.32 | 3.26 |
| Capability Dimension: *capability ratings* | Pancultural | 1.34** | 1.84 | 0.73 |
| | Asian | 1.64** | 2.20 | 0.75 |
| | Caucasian | -0.25 | 1.93 | 0.13 |
| Chinese participants | | | | |
| Approachability Dimension: *approachability ratings* | Pancultural | 3.58** | 1.93 | 1.85 |
| | Asian | 3.13** | 1.87 | 1.67 |
| | Caucasian | 3.00** | 1.89 | 1.59 |
| Youth-Attract Dimension: *youth ratings* | Pancultural | 5.00** | 1.22 | 4.10 |
| | Asian | 5.15** | 1.63 | 3.16 |
| | Caucasian | 3.00** | 1.89 | 1.59 |
| Youth-Attract Dimension: *attractiveness ratings* | Pancultural | 2.28** | 1.87 | 1.22 |
| | Asian | 2.33** | 1.76 | 1.32 |
| | Caucasian | 2.68** | 1.61 | 1.66 |
| Capability Dimension: *capability ratings* | Pancultural | 1.73** | 1.91 | 0.91 |
| | Asian | 1.78** | 1.90 | 0.94 |
| | Caucasian | 0.90* | 1.96 | 0.46 |

** $p < .001$, * $p < .01$, British $n = 44$, Chinese $n = 40$.

**Divergent validity**

We then tested whether the three dimensions differed *most* on the predicted traits relative to the other traits (i.e. whether the approachability dimension differed most on

approachability, and so on). We ran a four-way mixed ANOVA, with dimension (1-3), trait (approachability, youth, attractiveness, capability), and face race (pancultural, Asian, Caucasian) as within-subjects factors, and participant culture (British, Chinese) as a between subject factor. Huynh-Feldt corrections were used where applicable.

The critical two-way interaction between trait rating and dimension was significant, showing that the trait impressions differed across faces on the different dimensions, as predicted: $F(4.6,378.3) = 205.13$, $p < .001$, $\eta_p^2 = .714$. This pattern was moderated by significant three-way interactions between trait, dimension and participant culture: $F(4.6,378.3) = 5.95$, $p < .001$, $\eta_p^2 = .068$; as well as between trait, dimension and face race: $F(9.7, 798.5) = 45.49$, $p < .001$, $\eta_p^2 = .357$. The four-way interaction was not significant: $F(9.7,798.5) = 1.47$, $p = .128$, $\eta_p^2 = .018$.

We therefore examined each participant culture separately (figure 4): the critical two-way interaction between trait and dimension was significant for both Chinese and British participants, over all faces: both $F(4.7,181.8) = 96.57$, $p < .001$, $\eta_p^2 > .712$. We used planned contrasts to test each dimension by comparing predicted traits against other traits, for both Chinese and British participants (figure 4). As predicted, the approachability dimension differed significantly more on approachability than on the other traits, for Chinese: all $F(1,39) > 72.16$, $p < .001$, $\eta_p^2 > .649$; and British participants: all $F(1,43) > 116.32$, $p < .001$, $\eta_p^2 > .730$. Also as predicted, the youthful-attractiveness dimension differed significantly more on youth and attractiveness ratings than on the other traits, for Chinese: all $F(1,39) > 12.31$, $p < .01$, $\eta_p^2 > .240$; and British participants: all $F(1,43) > 76.36$, $p < .001$, $\eta_p^2 > .640$. Finally, the capability factor significantly differed more on capability than on approachability and youth for Chinese: both $F(1,39) > 6.33$, $p < .05$, $\eta_p^2 > .140$ ; and British participants: both $F(1,43) > 9.13$, $p < .05$, $\eta_p^2 > .175$. However, there was no difference on the capability factor between capability and attractiveness for either Chinese: $F(1,39) = 0.39$, $p = .536$, $\eta_p^2 = .010$; or British

participants: $F(1,43) = 2.24$, $p = .142$, $\eta_p^2 = .050$. Thus, there was excellent divergent validity on the first two factors and some, incomplete divergent validity for the last factor (figure 4). Tests of the other four cultural models agreed with this overall conclusion (Online Supplement text, figure S9).

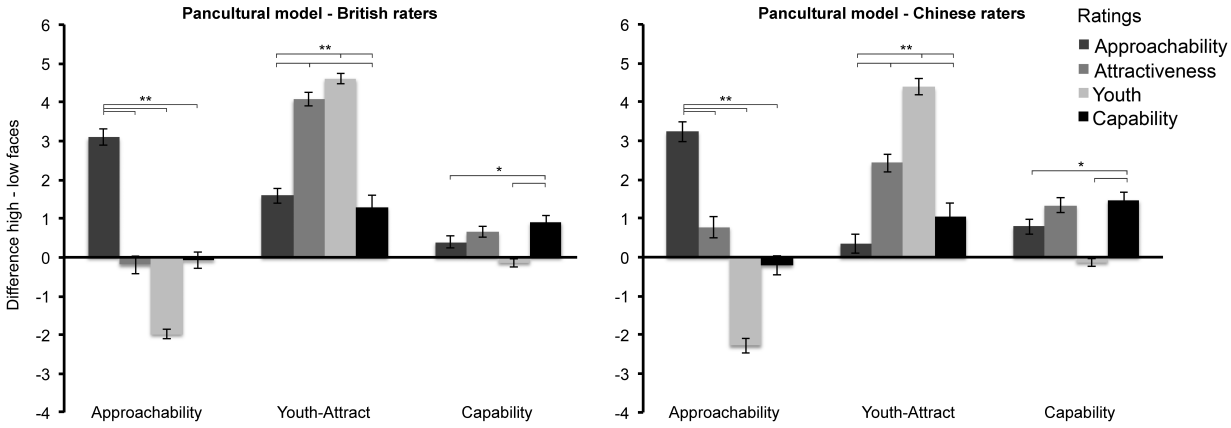


*Figure 4.* Average difference in approachability, attractiveness, youth and capability between high and low average faces on each pancultural dimension (shown on the x-axis), for British ($n = 44$) and Chinese participants ($n = 40$). ** $p < .01$, * $p < .05$. Error bars depict ± SEM.

Critically, we were able to cross-validate the approachability and youthful-attractiveness dimensions, with these dimensions showing both criterion and divergent validity. Impressions along these dimensions also appeared highly consistent across participant culture. The last (capability) dimension was less clearly cross-validated, especially for British participants judging Caucasian faces.

Of course, latter dimensions in factor analysis are inherently more variable, and our data-driven approach also necessitated that these traits were less sampled (see also Oosterhof & Todorov, 2008). When traits are sampled evenly using a top-down approach, a distinct dimension representing dominance or competence clearly emerges for Western European participants judging Caucasian faces (cf. Sutherland, Oldmeadow, & Young, 2016; Walker &

Vetter, 2009). However, a growing body of research is now demonstrating that traits along the capability or dominance dimension appear most variable; whether in terms of lower reliability (Sutherland et al., 2013); higher perceiver idiosyncrasy (Hehman, Sutherland, Flake, & Slepian, in press), or greater face gender differences (Sutherland et al., 2016; Sutherland, Young, et al., 2015). Here, we show that the greatest cross-cultural differences also exist along this dimension, across both faces and participants (figures 4 and S9, tables 5 and S11). In particular, across Studies 2-3, the last dimension for Caucasian faces appears more socially-oriented and driven by increased facial age, perhaps reflecting social ability or wisdom, whereas the last dimension for Asian and pancultural faces more clearly represents capability, driven by facial youth.

## General Discussion

We investigated non-Western (Chinese) perceivers' unconstrained facial impressions for the first time (Study 1), and then used these unconstrained descriptions to build data-driven models of Chinese impressions (Study 2). The same procedures were used to create models of British participants' impressions as a comparison. We then cross-validated these models in Study 3. Overall, we found substantial cross-cultural similarity in the frequency, type and underlying structure of facial impressions.

### Cross-cultural similarity

Study 1 showed that Chinese perceivers make facial impressions of enduring traits, emotions, age, sex and appearance in similar proportions to British perceivers. Moreover, Chinese trait attributions, although in a different language, closely echoed British descriptions. Strikingly, both cultures focused on traits relating to approachability or warmth. Study 2 found that dimensions of approachability, youthful-attractiveness and (sometimes)

capability emerged in models of impressions from both British and Chinese perceivers, supporting previous Western models (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Walker & Vetter, 2009; Wolffhechel et al., 2014). Study 3 cross-validated the dimensions with new British and Chinese perceivers, finding strong criterion and divergent validity for the approachability and youthful-attractiveness dimensions, with less clear cross-validation for the capability dimension.

In general, the current results demonstrate substantial consistency across culture (China and the UK) and face race (Asian and Caucasian) and provide initial support for the claim that universal dimensions underpin facial impressions across cultures, at least for approachability and youthful-attractiveness. It is worth reiterating that the data-driven sampling method offered a strong empirical test of this question. Although the Chinese dimensions which emerged were not radically different to the Western dimensions identified by previous research (e.g. Oosterhof & Todorov, 2008; Sutherland et al., 2013), this result arose from the participants' own impressions, in the absence of *a priori* labels chosen by the researchers. Together, the results demonstrate the value of a data-driven approach.

The approachability dimension showed the strongest cross-cultural similarity (across perceivers and faces), supporting initial predictions based on research using verbally presented targets. For example, there is greater cross-cultural stability in the meaning of social traits such as approachability, compared to skills-based traits (Ybarra et al., 2008). A large body of work in social psychology has also found evidence that warmth or morality is the primary dimension of social cognition (Fiske et al., 2007), which is further supported by the current findings. Our results indicate that judgments made from visual facial information may share a similar psychological structure with abstract conceptual impressions (Oldmeadow, Sutherland, & Young, 2013). Interestingly, it is currently debated whether warmth and morality fall under distinct conceptual dimensions (Abele & Wojciszke, 2007; Brambilla,

Rusconi, Sacchi, & Cherubini, 2011; Landy, Piazza, & Goodwin, 2016). Here we find both types of traits loading on the first dimension (e.g. friendliness, kindness) along with extraversion-related traits (e.g. not shy), consistent with previous face perception research (Sutherland, Rowley, et al., 2015; Walker & Vetter, 2016). When people form first impressions of strangers from face images, they may primarily form an overall impression of approachability, using emotional expression and social category cues (although the last dimension may also represent social ability, see below). We suggest that impressions along this face dimension function across culture to adaptively judge whether a target holds positive or negative intentions (following trustworthiness and warmth dimensions in other face models: Oosterhof & Todorov, 2008; Walker et al., 2011).

We also found clear evidence for a youthful-attractiveness dimension across all perceiver and face models. This finding fits with a long history of research on the importance of facial attractiveness as a mechanism of sexual selection (Rhodes, 2006). Indeed, attractiveness is well-known to be judged reliably across culture (e.g. Cunningham et al., 1995; see Rhodes, 2006 for a review). The current research extends this body of work by showing that these perceptions are also similar in their relative importance across face race and perceiver culture. Across all models, age or attractiveness always emerged as the second dimension of facial impressions.

**Cross-cultural differences**

While our discussion has thus far focused on the considerable cross-cultural similarities, there were also some interesting cross-cultural differences. First, perceivers' impressions of own-race faces showed higher dimensionality than impressions of other-race faces. This pattern could reflect a form of own-race advantage (e.g. Hugenberg, Young, Bernstein, & Sacco, 2010), representing more differentiated impressions of one's own ethnic

group. Alternatively, the findings might reflect fatigue, as own-race faces were always rated first, to avoid influencing own-culture impressions by immediately revealing the aims of the study. In either case, other-race faces may have been processed more superficially or with less motivation, raising the intriguing possibility that the dimensionality of the models is more flexible than previously described. The split between age and attractiveness dimensions particular to the Chinese model may also reflect the positive emphasis that Chinese culture places on old age (Chung & Lin, 2012; Levy & Langer, 1994 but see Chan et al., 2012). Future work could further investigate these interesting cultural or contextual differences.

Our models also diverged from previous studies with Western perceivers in that we did not find strong evidence for a facial dominance dimension. In these previous studies, the dominance dimension has reflected impressions of aggressiveness, power, masculinity and to some extent, even untrustworthiness (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Walker & Vetter, 2009; but see Sutherland et al., 2016). In Study 1, neither British nor Chinese perceivers spontaneously mentioned dominance when asked to give their first impressions. Interestingly, other studies of unconstrained facial impressions also do not find that dominance is frequently mentioned (Oosterhof & Todorov, 2008), even when dominant-looking faces are used (Sutherland, Young, et al., 2015). Instead, perceivers mention capability-related traits, and a dimension akin to capability appeared here across most models, although less consistently than the other dimensions and with the least cross-cultural agreement. Rather than reflecting physical dominance or aggressiveness, this dimension included intelligence and capability traits, and (especially for Caucasian faces) also reflected pro-social aspects of competence as well (i.e. social skills, status, or the ability to give resources to others).

Therefore, we suggest that capability may be a better way to represent this dimension across cultures, as well as corresponding better with prominent theoretical models in social

psychology which also focus on capability rather than physical power or dominance (including the Stereotype Content Model: Fiske et al., 2007). Interestingly, the capability dimension here also appeared similar to the sociability dimension described in other functionalist models, in the sense that sociability reflects a person's ability to recruit allies to accomplish shared goals (cf. Landy et al., 2016; note that this concept is distinct from morality or positive intentionality, described here as approachability). Construing this last dimension as capability more broadly also draws attention to the positive aspects of this dimension. Rather than simply reflecting threat, facial impressions likely also serve to highlight opportunities provided by conspecifics.

**The functionality of facial impressions**

We suggest that these dimensions of facial impressions derive either directly or indirectly from mechanisms for judging the opportunities or threat afforded by others, across face race and perceiver culture. To this extent, they should appear across cultures, as found here. However, based on our current results, we also predict that the specific attributes which form these dimensions and the facial cues used to judge them will vary between cultural or social groups, depending on their utility for the context at hand (Oldmeadow et al., 2013). In particular, we expect that the capability dimension will appear more variable across culture, given the current results and since one's capability is necessarily a function of the task being carried out. Ultimately, there should be a match between facial impression dimensions and the kinds of opportunities and threats in one's social environment. Any analysis of the key dimensions of facial first impressions should therefore explicitly allow for ecologically adaptive contextual differences in facial cues. This suggestion is similar to the concept of 'variform universality' in cross-cultural theories (Den Hartog, House, Hanges, Ruiz-

Quintanilla, & Dorfman, 1999) and with 'cultural dialect' theory in facial emotion research (Elfenbein & Ambady, 2002).

In future, studies can further test this account by manipulating the local or global context while measuring perceivers' facial impressions along these dimensions. For example, it would be interesting to compare model dimensions emerging when faces are blocked versus intermixed by social group (i.e. making the social group salient or not), or when multiple social groups are simultaneously manipulated. We predict that the approachability dimension will show the clearest translation across contexts, followed by the youthful-attractiveness dimension. It would also be worth ascertaining whether our current results generalize to other cultural groups, since we have only modelled perceptions from two possible groups (British and Chinese). This focus was deliberate, as the differences between these cultural groups offers a good initial test of potential universality. Nevertheless, an important goal for future research in facial impressions should be to understand target and perceiver variation in these impressions from diverse social groups across the world.

**Conclusions**

In three studies, we develop the first data-driven non-Western (Chinese) model of facial impressions. We find evidence for substantial cross-cultural similarity, with dimensions of approachability, youthful-attractiveness and (sometimes) capability emerging as important for both British and Chinese perceivers. The approachability and youthful-attractiveness dimensions showed the highest stability, both across the current Chinese and British models, and when compared to previous research. Although impressions may be influenced by cultural-specific facial cues, dimensions underlying these facial judgments likely function similarly across culture to judge the opportunities or threats afforded by others. Whether in

Yorkshire or in Guangdong, observers form very similar first impressions of a stranger, simply from seeing their face.

## References

Abele, A. E., & Bruckmüller, S. (2011). The bigger one of the "Big Two"? Preferential processing of communal information. *Journal of Experimental Social Psychology*, *47*, 935–948. https://doi.org/http://doi.org/10.1016/j.jesp.2011.03.028

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Adolphs, R., Nummenmaa, L., Todorov, A., & Haxby, J. (2016). Data-driven approaches in the investigation of social perception. *Phil. Trans. R. Soc. B*, *371*, 20150367. https://doi.org/10.1098/rstb.2015.0367

Baidu.com. (2014). 猥琐（汉语词语）编辑[wěi suǒ]. Retrieved from http://baike.baidu.com/subview/67302/11149339.htm

Baidu.com. (2015). 慈祥. Retrieved from http://baike.baidu.com/view/596886.htm

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, *41*(2), 135–143. https://doi.org/10.1002/ejsp.744

Bruce, V., & Young, A. W. (2012). *Face perception*. London; New York: Psychology Press.

Calder, A. J., Rhodes, G., Johnson, M., & Haxby, J. V. (Eds.). (2011). *Oxford Handbook of Face Perception*. Oxford: Oxford University Press.

Chan, W., McCrae, R. R., De Fruyt, F., Jussim, L., Löckenhoff, C. E., De Bolle, M., … Terracciano, A. (2012). Stereotypes of age differences in personality traits: Universal and accurate? *Journal of Personality and Social Psychology*, *103*(6), 1050–1066. https://doi.org/10.1037/a0029712

Chen, F. F., Jing, Y., & Lee, J. M. (2012). "I" value competence but "we" value social competence: The moderating role of voters' individualistic and collectivistic orientation in political elections. *Journal of Experimental Social Psychology*, *48*(6), 1350–1355. https://doi.org/10.1016/j.jesp.2012.07.006

Chung, C., & Lin, Z. (2012). A cross-cultural examination of the positivity effect in memory: United States vs. China. *The International Journal of Aging and Human Development*, *75*(1), 31–44. https://doi.org/10.2190/AG.75.1.d

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & 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. *Journal of Personality and Social Psychology*, *68*(2), 261–279. https://doi.org/10.1037/0022-3514.68.2.261

Den Hartog, D. N., House, R. J., Hanges, P. J., Ruiz-Quintanilla, S. A., & Dorfman, P. W. (1999). Culture specific and cross-culturally generalizable implicit leadership theories: Are attributes of charismatic/transformational leadership universally endorsed? *The Leadership Quarterly*, *10*(2), 219–256. https://doi.org/10.1016/S1048-9843(99)00018-1

Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *Review of Financial Studies*, *25*(8), 2455–2484. https://doi.org/10.1093/rfs/hhs071

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. https://doi.org/10.1037//0033-2909.128.2.203

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory Factor Analysis*. OUP USA.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/http://dx.doi.org/10.1037/1082-989X.4.3.272

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Hehman, E. A., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (in press). The unique contributions of perceiver and target characteristics in person perception. *JPSP*.

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, *10*, 15–41. https://doi.org/http://dx.doi.org/10.1080/00208825.1980.11656300

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187. https://doi.org/10.1037/a0020463

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

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *PNAS*, *109*(19), 7241–7244. https://doi.org/10.1016/j.cub.2009.07.051

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual Review of Psychology*, *68*, 269–297. https://doi.org/10.1146/annurev-psych-010416-044242

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, *121*, 313–323. https://doi.org/10.1016/j.cognition.2011.08.001

Kline, P. (1994). *An easy guide to factor analysis*. Psychology Press.

Krys, K., Hansen, K., Xing, C., Szarota, P., & Yang, M. (2013). Do only fools smile at strangers? Cultural differences in social perception of intelligence of smiling individuals. *Journal of Cross-Cultural Psychology*, *45*(2), 314–321. https://doi.org/10.1177/0022022113513922

Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When it's bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and Social Psychology Bulletin*, *42*(9), 1272–1290. https://doi.org/https://doi.org/10.1177/0146167216655984

Levy, B., & Langer, E. (1994). Aging free from negative stereotypes: Successful memory in China among the American deaf. *Journal of Personality and Social Psychology*, *66*(6), 989–997. https://doi.org/10.1037/0022-3514.66.6.989

Maass, A., Karasawa, M., Politi, F., & Suga, S. (2006). Do verbs and adjectives play different roles in different cultures? A cross-linguistic analysis of person representation. *Journal of Personality and Social Psychology*, *90*(5), 734–750. https://doi.org/10.1037/0022-3514.90.5.734

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*(6), 949–971. https://doi.org/10.1037/0022-3514.67.6.949

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological Review*, *108*(2), 291–310. https://doi.org/10.1037/0033-295X.108.2.291

Oldmeadow, J. A., Sutherland, C. A. M., & Young, A. W. (2013). Facial stereotype visualization through image averaging. *Social Psychological and Personality Science*, *4*(5), 615–623. https://doi.org/10.1177/1948550612469820

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Osgood, C. E. (1964). Semantic differential technique in the comparative study of cultures. *American Anthropologist*, *66*(3), 171–200. https://doi.org/10.1525/aa.1964.66.3.02a00880

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*(1), 3–72. https://doi.org/10.1037/0033-2909.128.1.3

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

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*(1), 199–226. https://doi.org/10.1146/annurev.psych.57.102904.190208

Rule, N. O., Ambady, N., Adams Jr, R. B., Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures.

*Journal of Personality and Social Psychology*, *98*(1), 1–15. https://doi.org/10.1037/a0017673

Santos, I. M., & Young, A. W. (2005). Exploring the perception of social characteristics in faces using the isolation effect. *Visual Cognition*, *12*(1), 213–247. https://doi.org/10.1080/13506280444000102

Saucier, G., Thalmayer, A. G., & Bel-Bahar, T. S. (2014). Human attribute concepts: Relative ubiquity across twelve mutually isolated languages. *Journal of Personality and Social Psychology*, *107*(1), 199–216. https://doi.org/10.1037/a0036492

Secord, P. F., & Bevan, W. (1956). Personalities in faces: III. A cross-cultural comparison of impressions of physiognomy and personality in faces. *The Journal of Social Psychology*, *43*(2), 283–288. https://doi.org/10.1080/00224545.1956.9919224

Sutherland, C. A. M. (2015). A basic guide to Psychomorph. University of York. Retrieved from https://dlib.york.ac.uk/yodl/app/home/detail?id=york:847354&ref=browse

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A. M., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. https://doi.org/https://doi.org/10.1016/j.cognition.2016.09.006

Sutherland, C. A. M., Rhodes, G., & Young, A. W. (in press). Facial image manipulation: A tool for investigating social perception. *Social Psychological and Personality Science*.

Sutherland, C. A. M., Rowley, L., Amoaku, U., Daguzan, E., Kidd-Rossiter, K., Maceviciute, U., & Young, A. W. (2015). Personality judgements from everyday images of faces. *Frontiers in Psychology*, *6*, 1–11. https://doi.org/10.3389/fpsyg.2015.01616

Sutherland, C. A. M., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology*, *106*(2), 186–208. https://doi.org/10.1111/bjop.12085

Tiddeman, B., Burt, M., & Perrett, D. I. (2001). Prototyping and transforming facial textures for perception research. *Computer Graphics and Applications, IEEE*, *21*(5), 42–50. https://doi.org/10.1109/38.946630

Todorov, A., Dotsch, R., Wigboldus, D. H. J., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social and Personality Psychology Compass*, *5*(10), 775–791. https://doi.org/10.1111/j.1751-9004.2011.00389.x

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. https://doi.org/10.1126/science.1110589

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *PNAS*, *111*(32), E3353–E3361. https://doi.org/10.1073/pnas.1409860111

Walker, M., Jiang, F., Vetter, T., & Sczesny, S. (2011). Universals and cultural differences in forming personality trait judgments from faces. *Social Psychological and Personality Science*, *2*(6), 609–617. https://doi.org/10.1177/1948550611402519

Walker, M., & Vetter, T. (2009). Portraits made to measure: Manipulating social judgments about individuals with a statistical face model. *Journal of Vision*, *9*(11), 1–13. https://doi.org/10.1167/9.11.12

Walker, M., & Vetter, T. (2016). Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of Personality and Social Psychology*, *110*(4), 609–624. https://doi.org/10.1037/pspp0000064

Wheeler, L., & Kim, Y. (1997). What is beautiful is culturally good: The physical attractiveness stereotype has different content in collectivistic cultures. *Personality and Social Psychology Bulletin*, *23*(8), 795–800. https://doi.org/10.1177/0146167297238001

Wolffhechel, K., Fagertun, J., Jacobsen, U. P., Majewski, W., Hemmingsen, A. S., Larsen, C. L., … Jarmer, H. (2014). Interpretation of appearance: The effect of facial features on first impressions and personality. *PLoS ONE*, *9*(9), e107721. https://doi.org/10.1371/journal.pone.0107721

World Population Clock: 7 Billion People (2014) - Worldometers. (2014). Retrieved July 23, 2014, from http://www.worldometers.info/world-population/#region

Yan, X., Andrews, T. J., & Young, A. W. (2016). Cultural similarities and differences in perceiving and recognizing facial expressions of basic emotions. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 423–440. https://doi.org/10.1037/xhp0000114

Ybarra, O., Chan, E., Park, H., Burnstein, E., Monin, B., & Stanik, C. (2008). Life's recurring challenges and the fundamental dimensions: An integration and its implications for cultural differences and similarities. *European Journal of Social Psychology*, *38*(7), 1083–1092. https://doi.org/10.1002/ejsp.559

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, *15*(6), 603–623. https://doi.org/10.1007/BF01065855

Zebrowitz, L. A., Wang, R., Bronstad, P. M., Eisenberg, D., Undurraga, E., Reyes-García, V., & Godoy, R. (2012). First impressions from faces among US and culturally isolated Tsimane' people in the Bolivian rainforest. *Journal of Cross-Cultural Psychology*, *43*(1), 119–134. https://doi.org/10.1177/0022022111411386

**SUTHERLAND ET AL. - ONLINE SUPPLEMENT**

## Study 3 Cross-validation of individual cultural models

### Criterion validity

To test criterion validity, we examined whether the high and low average faces on each dimension differed on the predicted traits, using one-sample t-tests against zero (for example, we tested the approachability ratings for the approachability dimension, and so on). We tested the four models (i.e. British or Chinese perceivers, Caucasian or Asian faces) separately since they differed in dimensionality. All dimensions in all models significantly differed on their predicted traits for both British and Chinese participant groups tested in Study 3, except the capability dimension for the two Caucasian models for British participants (table S11). A sign test showed that the overall pattern of results was significantly different from chance ($p < .0001$). Thus, there was strong overall criterion validity, except for Caucasian faces on capability for British participants.

### Divergent validity

To establish divergent validity we tested whether the three dimensions differed *most* on the predicted traits relative to the other traits (i.e. the approachability dimension should differ most on approachability, and so on; figure S9). We tested each model separately using a mixed ANOVA, with model dimension (ranging from two to four) and trait (approachability, youth, attractiveness, capability) as within-subjects factors, and participant culture (Chinese, British) as a between-subjects factor (figure S9). Huynh-Feldt corrections were used where necessary.

For each of the four models, the critical two-way interaction between trait and dimension was significant: all $F(5.7, 468.5) > 73.18$, $p < .001$, $\eta_p^2 > .472$, but was moderated by a significant three-way interaction with participant culture: all $F(2.8, 231.4) > 4.40$, all $p <$

.01, $\eta_p^2 > .051$. We therefore examined each participant culture separately; using planned contrasts to test our hypotheses by comparing predicted traits with the other traits.

**British model Caucasian faces:** The critical two-way interaction between trait and dimension was significant for British and Chinese participants separately: both $F(5.4, 231.4) = 38.96$, $p < .001$, $\eta_p^2 > .475$. As predicted, the first dimension differed significantly more on approachability than on the other traits, while the second dimension differed significantly more on youth and attractiveness than on the other traits, for Chinese: all $F(1,39) > 18.26$, $p < .001$, $\eta_p^2 > .319$, and British participants: all $F(1,43) > 4.90$, $p < .05$, $\eta_p^2 > .102$. Against predictions, the last dimension significantly differed less on capability than on approachability and youth for Chinese: all $F(1,39) > 10.08$, $p < .003$, $\eta_p^2 > .205$; and British participants: all $F(1,43) > 29.57$, $p < .001$, $\eta_p^2 > .407$. There was no difference between capability and attractiveness: Chinese $F(1,39) = 0.39$, $p = .537$, $\eta_p^2 = .010$; British: $F(1,43) = 0.03$, $p = .859$, $\eta_p^2 = .001$ (figure S9).

**British model Asian faces:** The critical two-way interaction between trait and dimension was significant for British and Chinese participants separately: both $F(3,117) = 83.77$, $p < .001$, $\eta_p^2 > .682$. As predicted, the approachability dimension differed significantly more on approachability than on the other traits, while the youthful-attractiveness dimension differed significantly more on youth and attractiveness than on the other traits, for Chinese: all $F(1,39) > 12.29$, $p < .001$, $\eta_p^2 > .240$; and British participants: all $F(1,43) > 47.60$, $p < .001$, $\eta_p^2 > .525$ (figure S9).

**Chinese model Caucasian faces:** The critical two-way interaction between trait and dimension was significant for British and Chinese participants separately: both $F(6,234) = 38.05$, $p < .001$, $\eta_p^2 > .494$. As predicted, the first dimension differed significantly more on approachability than on the other traits, while the second dimension differed significantly more on youth and attractiveness than on the other traits, for Chinese: all $F(1,39) > 4.29$, $p <$

.05, $\eta_p^2 > .099$; and British participants: all $F(1,43) > 14.97$, $p < .001$, $\eta_p^2 > .258$. Against predictions, the third dimension differed significantly more on youth than on capability for Chinese: $F(1,39) > 117.70$, $p < .001$, $\eta_p^2 > .751$; and British participants: $F(1,43) > 54.60$, $p < .001$, $\eta_p^2 > .559$. This dimension also differed more on approachability than capability for British participants: $F(1,43) = 9.07$, $p < .004$, $\eta_p^2 = .174$. There were no other significant differences: all $p > .204$, $\eta_p^2 < .037$ (figure S9).

**Chinese model Asian faces:** The critical two-way interaction between trait and dimension was significant for British and Chinese participants separately: both $F(7.2, 277.7) = 39.31$, $p < .001$, $\eta_p^2 > .502$. As predicted, the first dimension differed significantly more on approachability than on the other traits, the second dimension differed significantly more on youth than on the other traits, and the third dimension differed significantly more on attractiveness than on the other traits, for Chinese: all $F(1,39) > 10.59$, $p < .002$, $\eta_p^2 > .214$; and British participants: all $F(1,43) > 32.35$, $p < .001$, $\eta_p^2 > .429$. Also as predicted, the last dimension differed significantly more on capability versus approachability for Chinese participants: $F(1,39) = 7.62$, $p < .009$, $\eta_p^2 = .163$; and on capability versus approachability and attractiveness for British participants: both $F(1,43) > 4.27$, $p < .05$, $\eta_p^2 > .090$. However, this dimension also differed more on youth than on capability for Chinese: $F(1,39) = 16.28$, $p < .001$, $\eta_p^2 = .295$; and British participants: $F(1,43) = 25.43$, $p < .001$, $\eta_p^2 = .372$ (figure S9).

In summary, across the four models, there was excellent divergent validity for the approachability and youthful-attractiveness dimensions, and good divergent validity for the last (capability) dimension for the Chinese model for Asian faces, but not for this dimension for Caucasian faces.

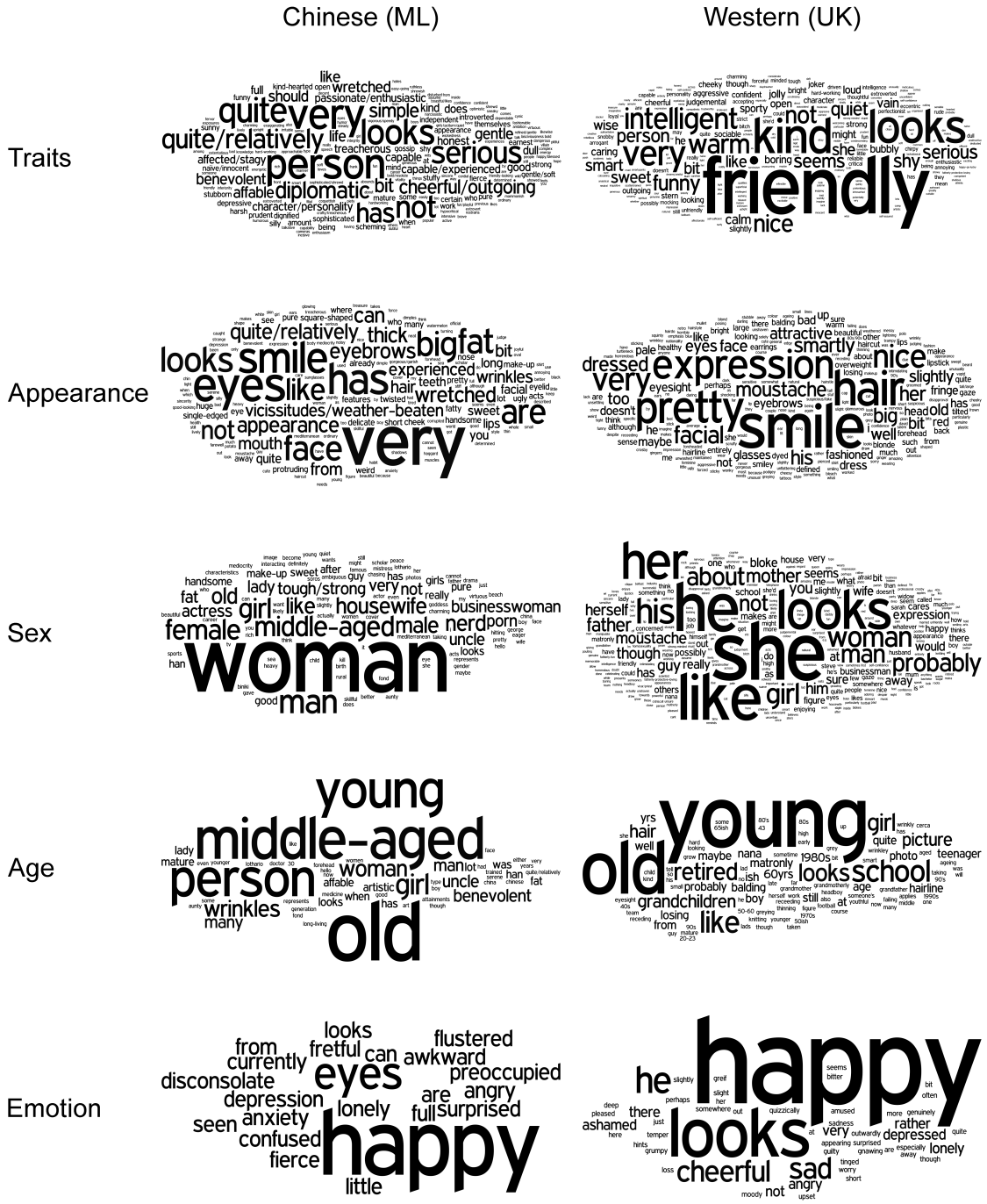


**Figure S1.** Word clouds depicting all descriptions under the five coded categories (age, appearance, emotion, sex and trait categories) for Chinese (translated only to assist readers) and British perceivers in Study 1. Note that these word clouds (unlike the main text) include all words within descriptions that were coded for each category, to give the reader an idea of the context around the key words. The only exclusions made were to bar around twenty frequent but uninformative English words from appearing (e.g. ‘a’ and ‘the’). Word clouds generated from wordle.net.

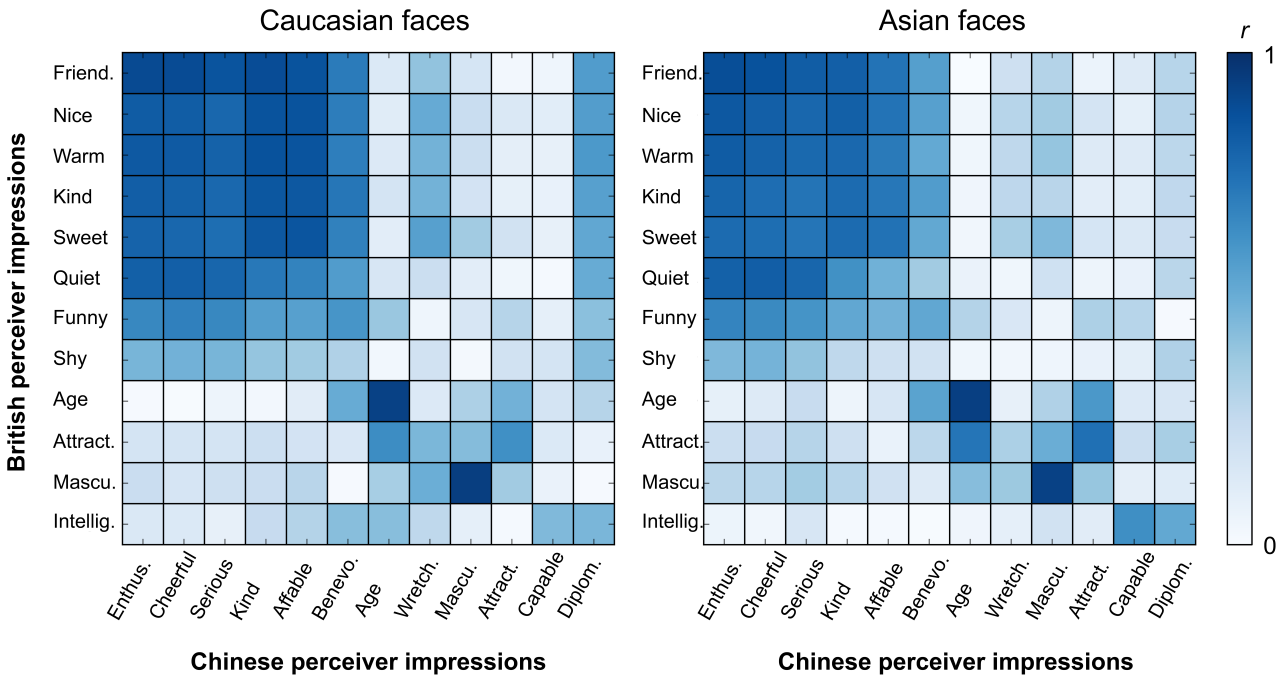


**Figure S2.** Agreement (absolute Pearson *r* of 0 = white, 1 = blue) between British (vertical axis) and Chinese impressions (horizontal axis) of Caucasian (left) and Asian faces (right) in Study 2. Although the British and Chinese participants rated the faces on culture-specific attributes, attributes are sorted by the factor solutions, so that the axis symmetry reflects cross-cultural agreement. In particular, a cluster of high agreement across cultures appears for warmth-related traits (top left), amongst age, attractiveness and masculinity (middle); and a cluster of moderate agreement amongst capability, diplomacy and intelligence (bottom right).

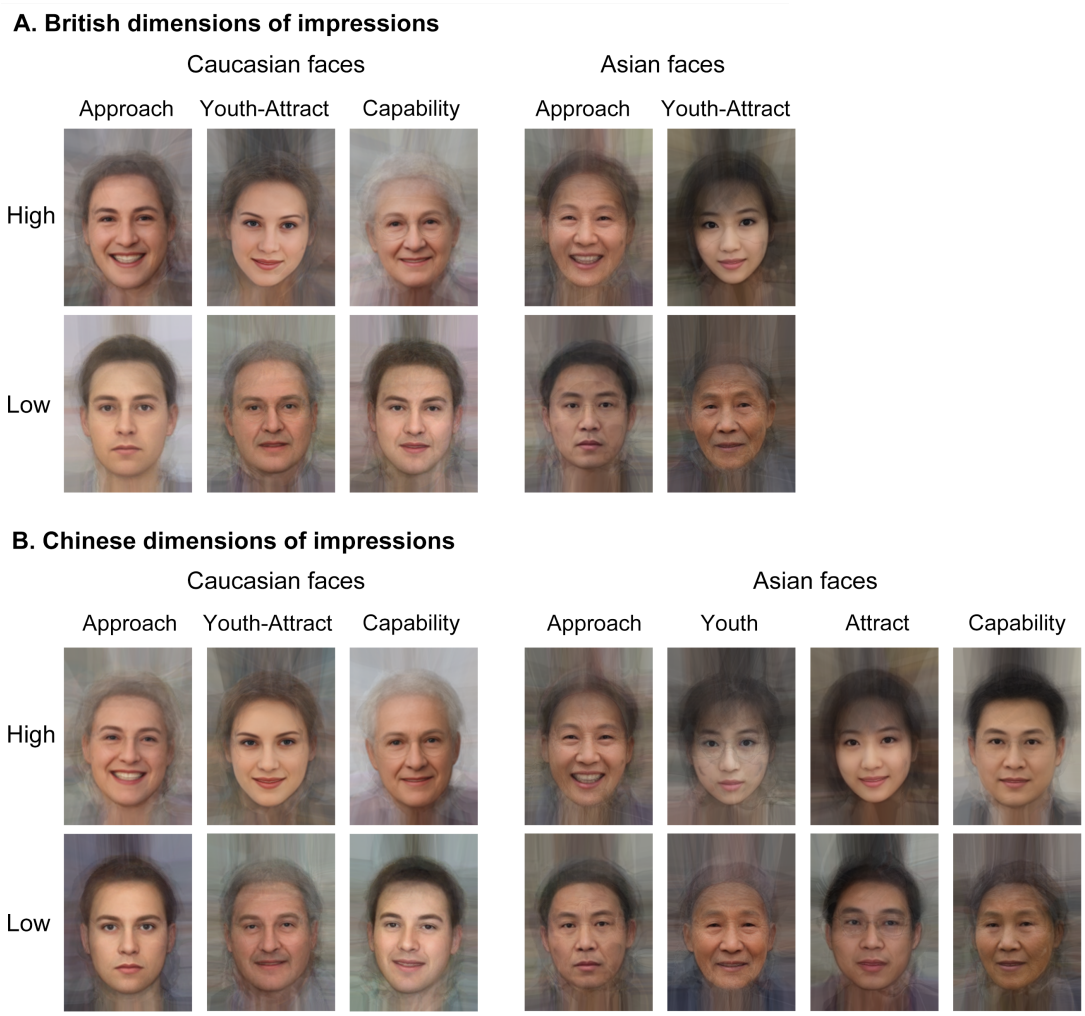


**Figure S3.** Faces created to lie high or low on the A) British and B) Chinese models of facial impressions of Asian and Caucasian faces. Each face is an average of the 20 highest or lowest scoring Asian and Caucasian face photographs on each of the dimensions in the Chinese and British models. Images used as stimuli in Study 3.

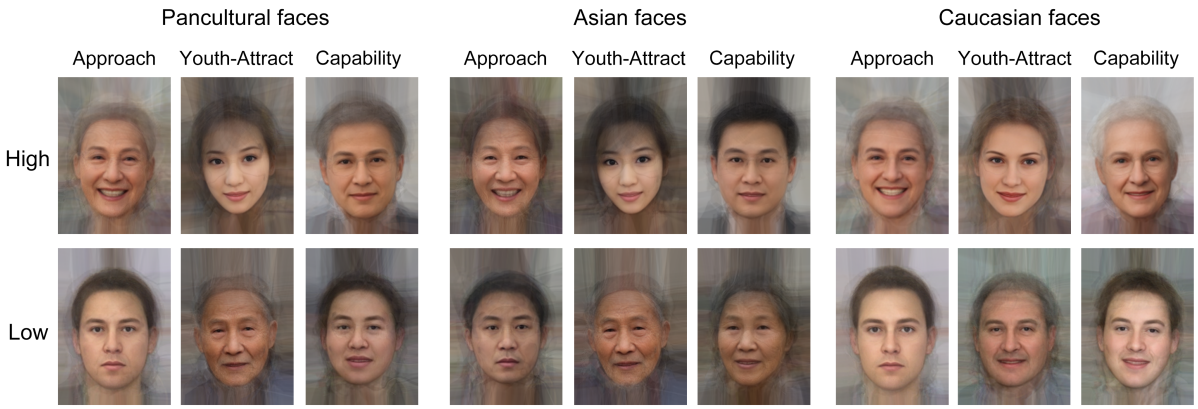


**Figure S4.** Pancultural, Asian and Caucasian face averages made from the 20 faces which scored most and least highly on pancultural dimensions of facial impressions. Images used as stimuli in Study 3.

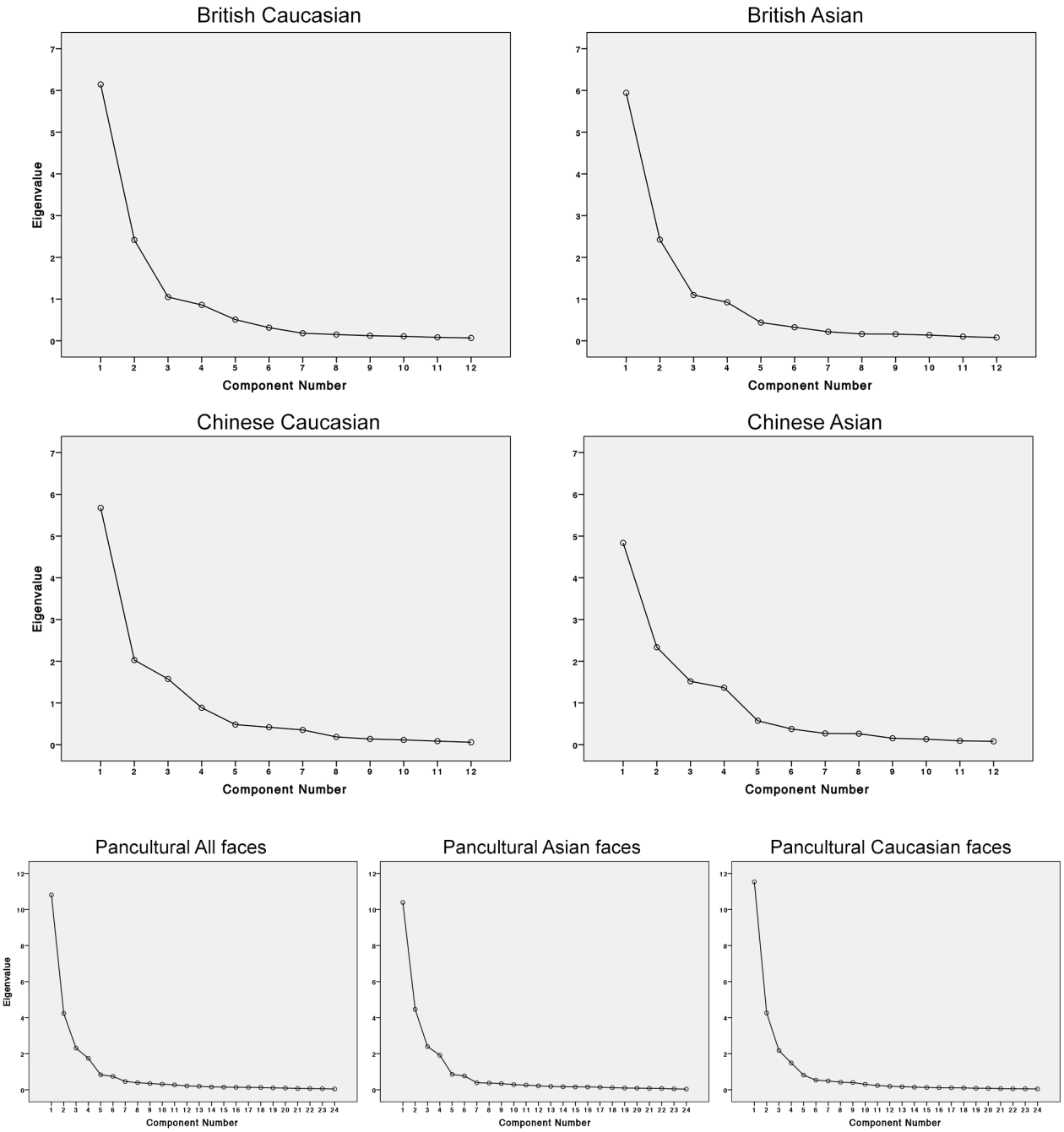


**Figure S5**. Scree plots run on the unreduced correlation matrix for each individual cultural model (British and Chinese, Caucasian and Asian faces) and for the pancultural models in Study 2.

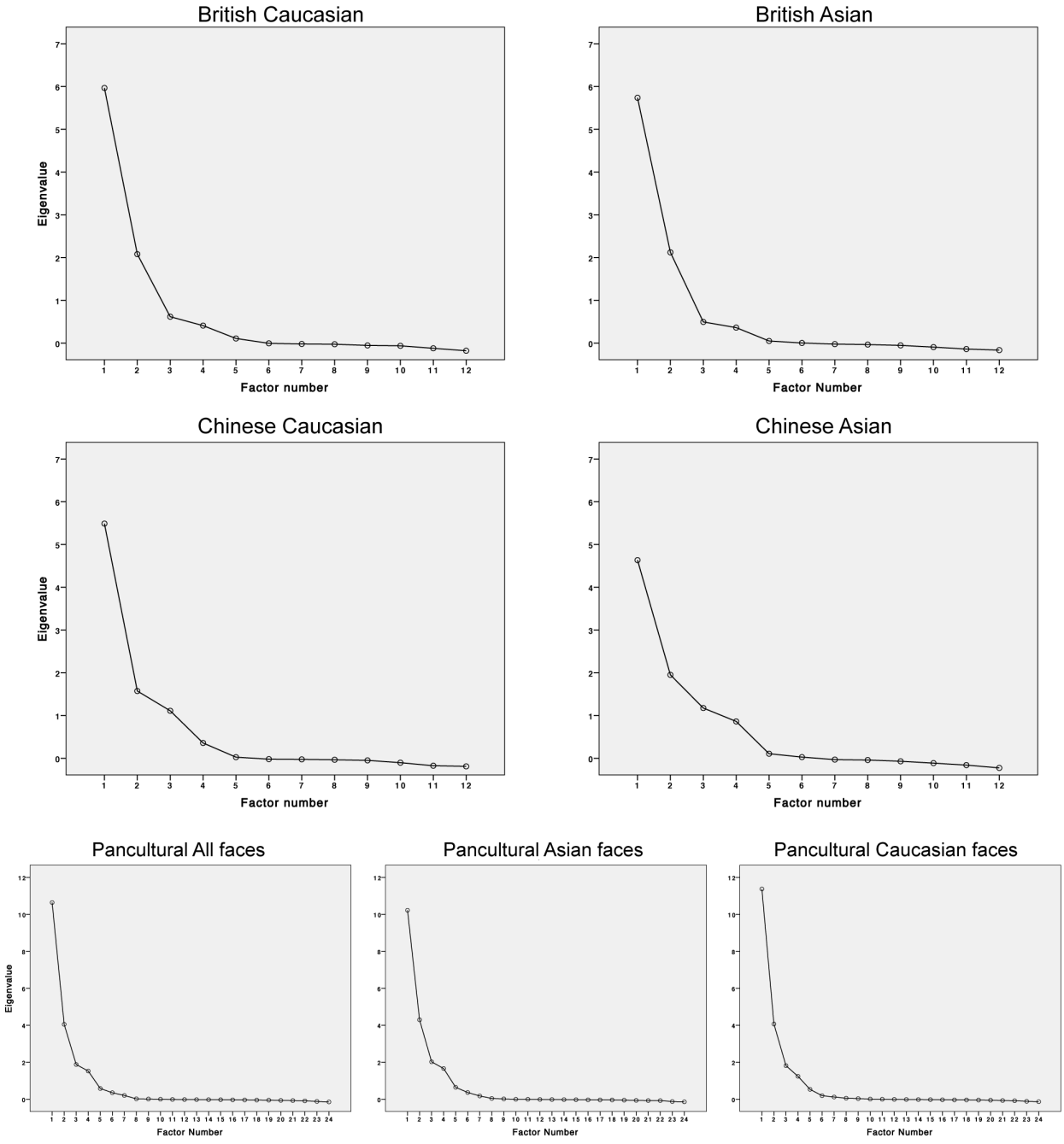


**Figure S6**. Scree plots run on the reduced correlation matrix for each individual cultural model (British and Chinese, Caucasian and Asian faces) and for the pancultural models in Study 2 (eigenvalues calculated using SPSS syntax from Fabrigar & Wegener, 2012, Exploratory Factor Analysis).


**Figure S7.** Asian and Caucasian face averages made from the 20 faces rated by British perceivers as highest and lowest on first impressions that are important to British perceivers (l-r: friendly, nice, warm, kind, sweet, shy, quiet, funny, age, attractiveness, masculinity, intelligence) in Study 2.


**Figure S8.** Asian and Caucasian face averages made from the 20 faces rated by Chinese perceivers as highest and lowest on first impressions that are important to Chinese perceivers (l-r: passionate/enthusiastic, cheerful/outgoing, serious (reversed), kind~and~gentle, affable, benevolent, age, wretched, masculinity, attractiveness, capable/experienced and diplomatic) in Study 2.

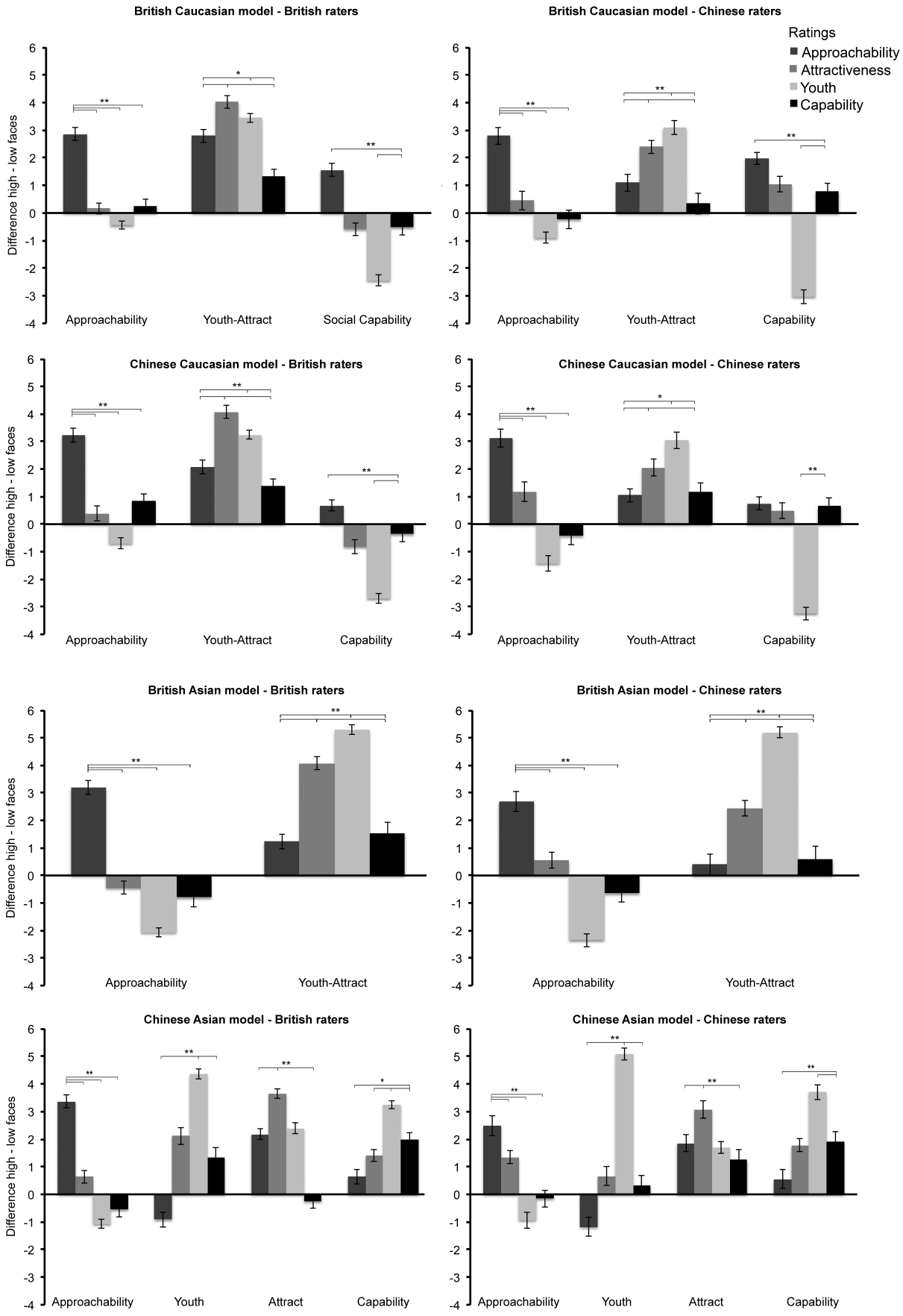
British Caucasian model - British raters
British Caucasian model - Chinese raters
Ratings
Approachability
Attractiveness
Youth
Capability
Difference high - low faces
Approachability
Youth-Attract
Social Capability
Capability
Chinese Caucasian model - British raters
Chinese Caucasian model - Chinese raters
British Asian model - British raters
British Asian model - Chinese raters
Chinese Asian model - British raters
Chinese Asian model - Chinese raters
Youth
Attract

**Figure S9.** Average difference in approachability, attractiveness, youth and capability ratings between high and low average faces on each model and dimension, for British ($n$ = 44) and Chinese participants ($n$ = 40) tested in Study 3. ** $p < .01$, * $p < .05$. Error bars depict ± SEM.

**Table S1.** Proportion of unique British participants mentioning gender, age, attractiveness and top trait words (out of $n$ = 20) in Study 1; and reliabilities (Cronbach's alpha) for own and other race facial ratings from British perceivers in Study 2 (each trait $k$ = 10, $n$ = 500).

| **Attribute** | **Proportion (frequency)** | **Own race (Caucasian faces)** | **Other race (Asian faces)** |
|---|---|---|---|
| Masculine | 0.80 (16) | .96 | .96 |
| Age | 0.75 (15) | .96 | .94 |
| Attractive | 0.70 (14) | .88 | .90 |
| Friendly | 0.65 (13) | .94 | .90 |
| Kind | 0.55 (11) | .87 | .77 |
| Intelligent | 0.40 (8) | .81 | .75 |
| Nice | 0.40 (8) | .88 | .85 |
| Warm | 0.40 (8) | .86 | .82 |
| Quiet | 0.35 (7) | .84 | .80 |
| Shy | 0.30 (6) | .63 | .45 |
| Sweet | 0.25 (5) | .89 | .78 |
| Funny | 0.25 (5) | .73 | .73 |

**Table S2.** Proportion of unique Chinese participants mentioning gender, age, attractiveness and top trait words (out of $n = 20$) in Study 1; and reliabilities (Cronbach's alpha) for own and other race facial ratings from Chinese perceivers in Study 2 (each trait $k = 10$, $n = 500$).

| | **Attribute** | **Proportion (frequency)** | **Own race (Asian faces)** | **Other race (Caucasian faces)** |
|---|---|---|---|---|
| 女性化的/男性化的 | Masculinity | 0.90 (18) | 0.95 | 0.96 |
| 年轻人/老年人 | Age | 0.85 (17) | 0.96 | 0.93 |
| 吸引力 | Attractive | 0.85 (17) | 0.83 | 0.72 |
| 严肃 | Serious | 0.30 (6) | 0.89 | 0.86 |
| 干练 | Capable/Experience | 0.25 (5) | 0.66 | 0.45 |
| 开朗 | Cheerful/Outgoing | 0.20 (4) | 0.89 | 0.88 |
| 慈祥 | Benevolent | 0.20 (4) | 0.77 | 0.77 |
| 和蔼 | Affable | 0.15 (3) | 0.81 | 0.86 |
| 圆滑 | Diplomatic | 0.15 (3) | 0.64 | 0.41 |
| 热情 | Passionate/Enthusiastic | 0.15 (3) | 0.93 | 0.94 |
| 和善 | Kind~and~gentle | 0.15 (3) | 0.82 | 0.85 |
| 猥琐 | Wretched | 0.15 (3) | 0.63 | 0.61 |

**Table S3.** Full correlation matrix, British participants tested in Study 2 (Caucasian faces below diagonal, Asian faces above diagonal).

| Trait | Friendly | Nice | Warm | Kind | Sweet | Quiet | Shy | Funny | Attract | Age | Mas. | Intel. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Friendly | - | .90** | .89** | .87** | .85** | -.78** | -.40** | .67** | .20** | -.01 | -.30** | -.06 |
| Nice | .91** | - | .88** | .86** | .87** | -.72** | -.35** | .58** | .30** | -.06 | -.36** | -.05 |
| Warm | .90** | .88** | - | .82** | .83** | -.72** | -.34** | .58** | .27** | -.07 | -.39** | -.07 |
| Kind | .90** | .89** | .87** | - | .82** | -.66** | -.32** | .58** | .20** | .02 | -.27** | -.04 |
| Sweet | .88** | .90** | .87** | .88** | - | -.64** | -.31** | .53** | .30** | -.07 | -.43** | -.09* |
| Quiet | -.79** | -.72** | -.74** | -.72** | -.68** | - | .54** | -.65** | -.22** | .09* | .24** | .12** |
| Shy | -.46** | -.47** | -.49** | -.41** | -.41** | .60** | - | -.41** | -.13** | -.03 | -.01 | -.06 |
| Funny | .66** | .55** | .57** | .59** | .50** | -.71** | -.34** | - | -.26** | .32** | .11* | -.19** |
| Attract | .17** | .28** | .23** | .18** | .30** | -.11* | -.34** | -.28** | - | -.79** | -.66** | .04 |
| Age | .10* | .05 | .08 | .12** | .04 | -.14** | -.01 | .37** | -.70** | - | .51** | .08 |
| Mascu. | -.19** | -.27** | -.25** | -.21** | -.39** | .10* | .04 | .17** | -.51** | .41** | - | .17** |
| Intel. | .28** | .31** | .28** | .31** | .28** | -.11* | -.11* | .17** | -.07 | .45** | .11* | - |

** $p < .01$, * $p < .05$, $n = 500$.

**Table S4.** Full correlation matrix, Chinese participants tested in Study 2 (Caucasian faces below diagonal, Asian faces above diagonal).

| | Passion. | Cheer. | Serious | Kind | Affable | Bene. | Age | Wretch. | Masc. | Att. | Cap. | Dipl. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Passionate/Enthus. | | .91** | -.88** | .82** | .71** | .50** | -.08 | -.17** | -.27** | .10* | -.07 | .34** |
| Cheerful/Outgoing | .93** | | -.89** | .77** | .64** | .43** | -.12** | -.09* | -.25** | .11* | -.05 | .35** |
| Serious | -.92** | -.91** | | -.74** | -.59** | -.31** | .23** | .04 | .31** | -.12** | .16** | -.25** |
| Kind~and~gentle | .87** | .86** | -.84** | | .80** | .57** | -.06 | -.32** | -.32** | .14** | -.06 | .29** |
| Affable | .82** | .81** | -.78** | .87** | | .71** | .16** | -.41** | -.27** | .10* | -.07 | .26** |
| Benevolent | .61** | .61** | -.54** | .68** | .73** | | .57** | -.32** | -.02 | -.17** | -.13** | .08 |
| Age | .02 | .04 | .02 | .07 | .16** | .57** | | .03 | .24** | -.57** | -.16** | -.16** |
| Wretched | -.36** | -.31** | .28** | -.48** | -.52** | -.36** | .07 | | .44** | -.41** | -.15** | -.19** |
| Masculinity | -.21** | -.14** | .18** | -.22** | -.27** | -.03 | .27** | .49** | | -.27** | .15** | -.06 |
| Attractive | .05 | .03 | -.03 | .11* | .13** | -.07 | -.43** | -.43** | -.29** | | .32** | .39** |
| Capable/Experience | -.02 | -.01 | .10* | .04 | .08 | .18** | .18** | -.26** | .07 | .26** | | .57** |
| Diplomatic | .50** | .52** | -.46** | .56** | .57** | .61** | .32** | -.36** | .00 | .07 | .31** | |

** $p < .01$, * $p < .05$, $n = 500$.

**Table S5.** Parallel analysis from Study 2 representing 95% percentile eigenvalues across 1,000 random datasets (500 cases for Asian/Caucasian face analyses and 1,000 cases for Pancultural face analyses, with 12 variables for the individual cultural models and 24 variables for the pancultural models).

| | | Eigenvalues Individual cultural models | Eigenvalues Pancultural, Asian/Caucasian face models | Eigenvalues Pancultural, All face models |
|---|---|---|---|---|
| Component number | 1 | 1.320 | 1.481 | 1.330 |
| | 2 | 1.236 | 1.402 | 1.278 |
| | 3 | 1.177 | 1.347 | 1.241 |
| | 4 | 1.127 | 1.296 | 1.207 |
| | 5 | 1.084 | 1.252 | 1.177 |
| | 6 | 1.040 | 1.213 | 1.151 |
| | 7 | 1.002 | 1.180 | 1.127 |
| | 8 | 0.966 | 1.144 | 1.105 |
| | 9 | 0.929 | 1.115 | 1.081 |
| | 10 | 0.893 | 1.084 | 1.061 |
| | 11 | 0.853 | 1.052 | 1.040 |
| | 12 | 0.807 | 1.023 | 1.020 |
| | 13 | - | 0.996 | 1.000 |
| | 14 | - | 0.969 | 0.981 |
| | 15 | - | 0.942 | 0.962 |
| | 16 | - | 0.913 | 0.942 |
| | 17 | - | 0.888 | 0.924 |
| | 18 | - | 0.860 | 0.904 |
| | 19 | - | 0.835 | 0.884 |
| | 20 | - | 0.808 | 0.866 |
| | 21 | - | 0.780 | 0.845 |
| | 22 | - | 0.750 | 0.825 |
| | 23 | - | 0.719 | 0.800 |
| | 24 | - | 0.684 | 0.771 |
| Cases (i.e. faces) | | 500 | 500 | 1,000 |
| Variables (i.e. traits) | | 12 | 24 | 24 |

**Table S6.** Dimensions of British impressions of Caucasian and Asian faces (principal component analyses, orthogonal varimax rotation). Loadings ≥ .30 are in bold.

| | Caucasian face dimensions | | | Asian face dimensions | |
|---|---|---|---|---|---|
| | Approach. | Youth-Attract | Capability | Approach. | Youth-Attract |
| Friendly | **0.93** | 0.07 | 0.20 | **0.95** | 0.10 |
| Nice | **0.89** | 0.19 | 0.27 | **0.92** | 0.20 |
| Warm | **0.90** | 0.15 | 0.23 | **0.90** | 0.20 |
| Kind | **0.88** | 0.11 | **0.30** | **0.89** | 0.10 |
| Sweet | **0.85** | 0.27 | **0.31** | **0.87** | 0.24 |
| Quiet | **-0.91** | 0.09 | 0.12 | **-0.84** | -0.10 |
| Funny | **0.75** | **-0.48** | -0.05 | **0.78** | **-0.40** |
| Shy | **-0.64** | -0.06 | 0.25 | **-0.50** | 0.07 |
| Age | 0.11 | **-0.78** | **0.44** | 0.09 | **-0.87** |
| Attractive | 0.17 | **0.87** | -0.12 | 0.13 | **0.92** |
| Masculine | -0.14 | **-0.77** | -0.13 | -0.24 | **-0.80** |
| Intelligent | 0.17 | -0.16 | **0.84** | -0.10 | -0.06 |
| Variance explained | 49% | 20% | 11% | 48% | 22% |

**Table S7.** Dimensions of Chinese impressions of Asian and Caucasian faces (principal component analyses, orthogonal varimax rotation). Loadings ≥ .30 are in bold.

| | | Asian face dimensions | | | | Caucasian face dimensions | | |
|---|---|---|---|---|---|---|---|---|
| Chinese ratings | Translations | Approach. | Youth | Attract. | Capability | Approach. | Youth-Attract | Capability |
| 热情 | Passion./enthusiastic | **0.95** | -0.05 | 0.08 | 0.07 | **0.95** | 0.11 | -0.01 |
| 开朗 | Cheerful/outgoing | **0.95** | 0.02 | 0.01 | 0.08 | **0.94** | 0.06 | 0.00 |
| 严肃 | Serious | **-0.94** | -0.16 | -0.01 | 0.06 | **-0.94** | -0.08 | 0.11 |
| 和善 | Kind~and~gentle | **0.86** | -0.16 | 0.28 | 0.06 | **0.92** | 0.15 | 0.13 |
| 和蔼 | Affable | **0.72** | **-0.41** | **0.37** | 0.08 | **0.89** | 0.16 | 0.24 |
| 慈祥 | Benevolent | **0.47** | **-0.79** | 0.21 | 0.01 | **0.71** | -0.21 | **0.51** |
| 年轻人/老年人 | Age | -0.13 | **-0.90** | -0.14 | -0.13 | 0.13 | **-0.67** | **0.56** |
| 猥琐 | Wretched | -0.05 | 0.18 | **-0.89** | -0.19 | **-0.35** | **-0.64** | **-0.42** |
| 女性化的/男性化的 | Masculine | -0.26 | -0.27 | **-0.70** | 0.25 | -0.19 | **-0.67** | 0.09 |
| 吸引力 | Attractive | 0.07 | **0.50** | **0.53** | **0.45** | -0.05 | **0.80** | 0.20 |
| 干练 | Capable/Experienced | -0.15 | 0.06 | 0.00 | **0.89** | -0.12 | 0.18 | **0.80** |
| 圆滑 | Diplomatic | **0.32** | 0.07 | 0.06 | **0.82** | **0.56** | -0.04 | **0.58** |
| Variance explained | | 37% | 17% | 15% | 15% | 44% | 17% | 16% |

**Table S8.** Inter-dimension correlations for British, Chinese and Pancultural models.

| Model | Dimensions | | Correlation |
|---|---|---|---|
| British Caucasian | Approachability | Youth-Attract | .128 |
| | Approachability | Capability | .354 |
| | Youth-Attract | Capability | -.038 |
| British Asian | Approachability | Youth-Attract | .133 |
| Chinese Caucasian | Approachability | Youth-Attract | .110 |
| | Approachability | Capability | .215 |
| | Youth-Attract | Capability | .028 |
| Chinese Asian | Approachability | Youth | -.080 |
| | Approachability | Attract | .324 |
| | Approachability | Capability | .055 |
| | Youth | Attract | .045 |
| | Youth | Capability | .109 |
| | Attract | Capability | .146 |
| Pancultural All faces | Approachability | Youth-Attract | .087 |
| | Approachability | Capability | .173 |
| | Youth-Attract | Capability | .055 |
| Pancultural Asian faces | Approachability | Youth-Attract | .123 |
| | Approachability | Capability | .007 |
| | Youth-Attract | Capability | .089 |
| Pancultural Caucasian faces | Approachability | Youth-Attract | .098 |
| | Approachability | Capability | .234 |
| | Youth-Attract | Capability | -.023 |

**Table S9.** Pancultural dimensions of impressions of 500 Asian, 500 Caucasian, and all 1,000 faces (principal axis factor analysis, structure matrices) in Study 2. These can be interpreted as akin to correlations between the factors and variables. Factor loadings ≥ .3 appear in bold. App. = Approachability, Youth-Attract = Youthful-attractiveness, Cap. = Capability.

| | | Pancultural faces | | | Asian faces | | | Caucasian faces | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | App. | Youth-Attract. | Cap. | App. | Youth-Attract. | Cap. | App. | Youth-Attract. | Cap. |
| Friendly | British | **.96** | .08 | .16 | **.96** | .13 | -.02 | **.96** | .11 | .24 |
| Passion-enthusiastic | Chinese | **.93** | .15 | .10 | **.93** | .17 | .03 | **.94** | .14 | .06 |
| Nice | British | **.93** | .17 | .24 | **.93** | .22 | .01 | **.92** | .24 | **.34** |
| Warm | British | **.92** | .16 | .20 | **.91** | .22 | -.04 | **.92** | .19 | **.31** |
| Cheerful-outgoing | Chinese | **.91** | .14 | .09 | **.90** | .20 | .04 | **.94** | .09 | .06 |
| Kind | British | **.90** | .08 | .24 | **.88** | .12 | 0 | **.91** | .15 | **.34** |
| Kind~and~gentle | Chinese | **.89** | .15 | .21 | **.87** | .17 | .04 | **.92** | .20 | .24 |
| Sweet | British | **.88** | .22 | .20 | **.88** | .25 | -.06 | **.89** | **.32** | **.35** |
| Serious | Chinese | **-.86** | -.22 | .07 | **-.85** | -.29 | .10 | **-.91** | -.12 | .04 |
| Shy | British | **-.46** | -.06 | -.21 | **.42** | 0 | -.18 | **-.50** | -.13 | -.14 |
| Quiet | British | **-.82** | -.05 | -.05 | **-.80** | -.14 | .01 | **-.84** | .02 | -.09 |
| Affable | Chinese | **.86** | .06 | .28 | **.79** | .01 | .06 | **.89** | .20 | **.36** |
| Benevolent | Chinese | **.68** | **-.36** | .28 | **.62** | **-.40** | 0 | **.74** | -.18 | **.59** |
| Funny | British | **.69** | **-.36** | -.12 | **.69** | **-.33** | -.23 | **.71** | **-.41** | .06 |
| Attractiveness | British | .19 | **.87** | .26 | .20 | **.92** | .23 | .16 | **.84** | -.06 |
| Attractiveness | Chinese | .08 | **.66** | **.40** | .09 | **.69** | **.37** | .04 | **.66** | .11 |
| Age | British | .08 | **-.87** | .16 | .02 | **-.88** | -.07 | .12 | **-.75** | **.59** |
| Age | Chinese | .10 | **-.81** | .13 | .04 | **-.82** | -.11 | .16 | **-.68** | **.61** |
| Masculine | British | -.27 | **-.69** | .00 | **-.31** | **-.74** | .17 | -.20 | **-.71** | -.02 |
| Masculine | Chinese | -.26 | **-.57** | .00 | **-.32** | **-.57** | .21 | -.19 | **-.64** | -.04 |
| Wretched | Chinese | **-.34** | **-.36** | **-.43** | -.25 | -.28 | -.16 | **-.41** | **-.59** | -.48 |
| Capable-experienced | Chinese | -.03 | .06 | **.71** | -.11 | .14 | **.83** | .04 | .09 | **.51** |
| Intelligent | British | .13 | -.17 | **.69** | -.05 | -.06 | **.73** | .25 | -.11 | **.68** |
| Diplomatic | Chinese | **.44** | .08 | **.63** | **.33** | .24 | **.72** | **.60** | -.01 | **.52** |
| Variance explained, varimax PCA | | 44% | 18% | 10% | 42% | 19% | 10% | 46% | 18% | 11% |

**Table S10.** Agreement across Chinese and British participant groups, as measured by Pearson's *r* correlations across ratings at the level of the faces.

| | All faces | Asian faces | Caucasian faces |
|---|---|---|---|
| Approachability | .927** | .943** | .921** |
| Youth | .976** | .985** | .980** |
| Attractiveness | .858** | .910** | .811** |
| Capability | .675** | .769** | .412 |

** $p < .001$, all faces $n = 42$, Asian/Caucasian faces $n = 18$.

**Table S11.** Mean differences in predicted trait ratings across pairs of high and low average faces on each dimension from each model. Dimensions were tested on the main predicted traits i.e. approachability, youth, attractiveness, and capability.

| British participants | Model | Mean high – low | SD high – low | *d* |
|---|---|---|---|---|
| British Caucasian | Factor 1: Approachability | 2.86** | 1.53 | 1.87 |
| | Factor 2: Youthful | 3.45** | 1.09 | 3.17 |
| | Factor 2: Attractiveness | 4.02** | 1.47 | 2.73 |
| | Factor 3: Capability | -0.52 | 1.93 | 0.27 |
| Chinese Caucasian | Factor 1: Approachability | 3.23** | 1.70 | 1.90 |
| | Factor 2: Youthful | 3.25** | 1.04 | 3.13 |
| | Factor 2: Attractiveness | 4.07** | 1.59 | 2.56 |
| | Factor 3: Capability | -0.36 | 1.83 | 0.20 |
| British Asian | Factor 1: Approachability | 3.20** | 1.69 | 1.89 |
| | Factor 2: Youthful | 5.30** | 1.15 | 4.61 |
| | Factor 2: Attractiveness | 4.07** | 1.56 | 2.61 |
| Chinese Asian | Factor 1: Approachability | 3.36** | 1.57 | 2.14 |
| | Factor 2: Youthful | 4.36** | 1.26 | 3.46 |
| | Factor 3: Attractiveness | 3.66** | 1.20 | 3.05 |
| | Factor 4: Capability | 2.00** | 1.68 | 1.19 |
| Chinese participants | | | | |
| British Caucasian | Factor 1: Approachability | 2.80** | 1.95 | 1.44 |
| | Factor 2: Youthful | 3.10** | 1.60 | 1.94 |
| | Factor 2: Attractiveness | 2.40** | 1.53 | 1.57 |
| | Factor 3: Capability | 0.80* | 1.84 | 0.43 |
| Chinese Caucasian | Factor 1: Approachability | 3.13** | 2.13 | 1.47 |
| | Factor 2: Youthful | 3.05** | 1.93 | 1.58 |
| | Factor 2: Attractiveness | 2.05** | 2.01 | 1.02 |
| | Factor 3: Capability | 0.65° | 2.03 | 0.32 |
| British Asian | Factor 1: Approachability | 2.68** | 2.30 | 1.17 |
| | Factor 2: Youthful | 5.20** | 1.22 | 4.26 |
| | Factor 2: Attractiveness | 2.45** | 1.81 | 1.35 |
| Chinese Asian | Factor 1: Approachability | 2.48** | 2.33 | 1.06 |
| | Factor 2: Youthful | 5.08** | 1.33 | 3.82 |
| | Factor 3: Attractiveness | 3.08** | 2.03 | 1.52 |
| | Factor 4: Capability | 1.90** | 2.36 | 0.81 |

** $p < .01$, * $p < .05$, ° $p = .05$

**SUTHERLAND ET AL. - INSTRUCTIONS**

## Study 1

### Overall instructions 1

In this study we are interested in what people think when they perceive the faces of others. Your task is to look at a set of 26 faces one by one, and write down everything that comes into your mind when you look at each face. This can be anything you like, no matter how silly or inappropriate. We just want to know exactly what you honestly think or feel. There is no right or wrong answer – just what you think. Please take your time to answer fully, but if you find your mind wandering, or that your thoughts are becoming less spontaneous, then move onto the next face. Please feel free to take a break at any point. The study will last no more than an hour. In the first part, you will just see faces, and in the second part, you will see faces along with a short description of where they are. When you have completed the study, we will give you a short demographics sheet to fill in.

在这项研究中，我们感兴趣的是当人们看到其他人的面孔时的想法。 你的任务为看一组逐个出现的面孔，并且写下当你看到每个面孔时脑海里浮现的所有想法。这可以是任何内容，不管有多么傻或者不恰当，因为你真实的想法正是我们所想知道的。答案没有对错之分，只是取决于你的想法。请利用时间充分作答，但如果你发现自己注意力不集中了或者想法变得没那么自然了，那就请看下一个面孔。请随时稍作休息。这个实验需时不超过一小时。在第一部分，你只会看到一些面孔；在第二部分，你会看到的是面孔以及一段对他们所在地的简短描述。当你完成这个实验的时候，我们需要你填写一份简短的人口统计调查表。

### Response instructions 1

Please write down your first impressions here.
Write whatever comes to mind.
Click "OK" when you're ready to move on.

请在这里写下你的第一印象。
写下所有在你脑海里浮现的。
当你准备好看下一个面孔的时候，点击“OK”

## Study 2

### Overall instructions 2

This study aims to investigate social impressions of faces. Your task is to look at a

set of 1000 faces and rate each one on a scale from 1 to 7 for how * they are. Please feel free to take a break at any point. When you have completed the study, we will give you a short demographics sheet to fill in. The experiment will last no more than an hour. There is no right or wrong answer - just what you think. Go with your gut instinct or first impression if you are unsure. Try and use the whole rating scale, and try and be consistent.

* was replaced with friendly, funny, intelligent, kind, nice, quiet, shy, sweet, warm, feminine or masculine, old, or attractive,.

这个实验目的在于探讨人们对面孔的印象。你的任务为看着一组面孔(1000 张)，并且按照一个从一到七的量表对每张面孔的*评分。你可以随时稍作休息。当你完成这个实验的时候，我们会给你填写一份简短的人口统计调查表。这个实验需时不会超过一小时。答案没有对错之分，只是取决于你的想法。如果你无法确定答案，那就请遵循你的直觉或第一印象。请尝试使用整个评分量表，并尽量保持一致。

* was replaced with 开朗(cheerful/outgoing), 严肃(serious), 慈祥(benevolent), 和蔼 (affable), 和善 (kind~and~gentle), 热情(passionate/enthusiastic), 干练 (capable/experienced), 圆滑 (diplomatic), 猥琐 (wretched), 吸引力 (attractive), 年龄 (age), or 女性化的/男性化的(feminine or masculine).

**Response instructions 2**

Response scales from: 1 (not very) to 7 (very): friendly, kind, intelligent, nice, warm, quiet, shy, funny, sweet, attractive, young to old, or feminine to masculine.

Response scale from: 1 (非常不, not very) to 7 (非常, very): 开朗(cheerful/outgoing), 严肃 (serious), 慈祥 (benevolent), 和蔼 (affable), 和善 (kind~and~gentle), 热情 (passionate/enthusiastic), 干练(capable/experienced), 圆滑 (diplomatic), 猥琐 (wretched), 吸引力(attractive), 年轻人 (young) to 老年人 (old), or 女性化的 (feminine) to 男性化的 (masculine).

Participants only saw one attribute; see main text for more details.

## Study 3

**Overall instructions 3**

This study aims to investigate social impressions of faces. Your task is to look at a set of faces and rate them for four different social traits, including how approachable, old, attractive and capable they look. Please feel free to take a break at any point. When you have completed the study, we will ask you some demographic questions. The experiment will last no more than 30 minutes.

这个实验目的在于探讨人们对面孔的印象。你的任务是看一组面孔，并且对这些面孔的四个不同的属性评分，这些属性包括：他们看起来有多平易近人、或老、或有吸引力、或干练。

你可以随时稍作休息。当你完成这个测试的时候，我们还会再问你一些人口统计调查问题。这个实验需时不会超过半个小时。

**Block instructions 3**

In this task, you will look at 44 faces and rate each one for how * they look. For each face, you will be asked to rate the person on a scale from 1 to 7 where 1 means not very *, and 7 means very *. There is no right or wrong answer – just what you think. Go with your gut instinct or first impression if you are unsure. Try and use the whole rating scale, and try and be consistent. This task will take around 5 minutes.

你的任务为看着一组面孔（44张），并且对每张面孔的*程度评分。

你将需要对每张面孔（这个人）的*程度按照一个从1到7的评分量表评分，1表示这个人是非常不*的，7表示这个人是非常*的。

如果你无法确定答案，那就请遵循你的直觉或第一印象。请尝试使用整个评分量表，并尽量保持评定标准的一致性。

这个测试将耗时大约 5 分钟。

* was replaced with approachable (平易近人), attractive (吸引力), capable (干练), or age (年龄).

**Response instructions 3**

Response scales from: 1 (not very) to 7 (very): approachable (平易近人), attractive ( 吸引力), capable (干练), or young (年轻人) to old (老年人).

Participants saw all attributes; see main text for more details.