Social perception of faces around the world: How well does the valence-dominance model generalize across world regions? (Registered Report Stage 1)

This is the first empirical study that has been selected to be run via the Psychological Science Accelerator, a new initiative for conducting large-scale psychological research (https://psysciacc.org/). The manuscript starts on page eight

Corresponding author: Benedict Jones (ben.jones@glasgow.ac.uk), Institute of Neuroscience & Psychology, University of Glasgow, Scotland, UK.

Benedict C Jones (Institute of Neuroscience & Psychology, University of Glasgow)
Lisa M DeBruine (Institute of Neuroscience and Psychology, University of Glasgow)
Jessica Kay Flake (Department of Psychology, McGill University)
Balazs Aczel (Institute of Psychology, ELTE, Eotvos Lorand University)
Matúš Adamkovič (Institute of Psychology, Faculty of Arts, University of Prešov)
Ravin Alaei (Psychology, University of Toronto)
Sinan Alper (Psychology, Baskent University)
Michael R Andreychik (Psychology, Fairfield University)
Daniel Ansari (Psychology, The University of Western Ontario)
Jack D Arnal (Psychology Department, McDaniel College)
Peter Babinčák (Institute of Psychology, Faculty of Arts, University of Prešov)
Gabriel Baník (Institute of Psychology, University of Presov)
Krystian Barzykowski (Institute of Psychology, Jagiellonian University)
Ernest Baskin (Food Marketing, Saint Joseph's University)
Carlota Batres (Department of Psychology, Franklin & Marshall College)
Khandis R Blake (Evolution and Ecology Research Centre, UNSW Sydney)
Martha Lucia Borras-Guevara (School of Psychology and Neuroscience, University of St Andrews)

Mark J Brandt (Department of Social Psychology, Tilburg University)
Debora I Burin (Instituto de Investigaciones, Facultad de Psicologia, Universidad de Buenos Aires - CONICET)
Sun Jun Cai (China, Qu Fu Normal University)
Dustin P Calvillo (Psychology, California State University San Marcos)
Priyanka Chandel (School of Studies in Life Science, Pt Ravishankar Shukla University, Raipur (Chhattisgarh))
Armand Chatard (Psychology, University of Poitiers & CNRS)
Sau-Chin Chen (Department of Human Development and Psyhology, Tzu-Chi Universitiy, Taiwan)
Coralie Chevallier (Département d'Etudes Cognitives, Paris Sciences et Lettres)
William J Chopik (Psychology, Michigan State University)
Cody D Christopherson (Psychology, Southern Oregon University)
Vinet Coetzee (Department of Biochemistry, Genetics and Microbiology, University of Pretoria)
Nicholas A Coles (Department of Psychology, University of Tennessee)
Melissa F Colloff (Centre for Applied Psychology, School of Psychology, University of Birmingham)
Corey L Cook (Department of Psychology, Pacific Lutheran University)
Matthew T Crawford (School of Psychology, Victoria University of Wellington)
Alexander F Danvers (Institute for the Study of Human Flourishing, University of Oklahoma)
Barnaby JW Dixson (The School of Psychology, The University of Queensland)
Vilius Dranseika (Institute of Philosophy, Vilnius University)
Yarrow Dunham (Psychology, Yale University)
Thomas Rhys Evans (School of Psychological, Social and Behavioural Science, Coventry University)
Ana Maria Fernandez (Laboratorio de Evolucion y Relaciones Interpersonales, Universidad de Santiago de Chile)
Heather D Flowe (Psychology, University of Birmingham)
Patrick S Forscher (Psychological Science, University of Arkansas)
Gwendolyn Gardiner (Psychology, University of California, Riverside)

Eva Gilboa-Schechtman (Department of Psychology and the Gonda Brain Science Center, Bar-Ilan university)
Michael Gilead (Psychology, Ben-Gurion University)
Tripat Gill (Lazaridis School of Business & Economics, Wilfrid Laurier University)
Isaac González-Santoyo (Psychology deparment, National Autonomous University of México)
Amanda C Hahn (Psychology, Humboldt State University)
Eric Hehman (Psychology, McGill University)
Chuan-Peng Hu (Neuroimaging Center, Johannes Gutenberg University Medical Center)
Hans IJzerman (LIP/PC2S, Université Grenoble Alpes)
Michael Inzlicht (Department of Psychology, University of Toronto)
Natalia Irrazabal (Fac Ciencias Sociales, Universidad de Palermo - CONICET)
Bastian Jaeger (Department of Social Psychology, Tilburg University)
Chaning Jang (Director, Busara Center for Behavioral Economics)
Steve M J Janssen (School of Psychology, University of Nottingham - Malaysia Campus)
Zhongqing Jiang (Psychology, Liaoning Normal University)
Pavol Kačmár (Department of Psychology, Faculty of Arts, Pavol Jozef Šafárik University in Košice)
Gwenael Kaminski (Cognition, Langues, Langage, Ergonomie, Toulouse University)
Aycan Kapucu (Psychology, Ege University)
Monica A Koehn (Department of Social Sciences and Psychology, Western Sydney University)
Vanja Kovic (Department of Psychology, Laboratory for Neurocognition and Applied Cognition, University of Belgrade)
Pratibha Kujur (SoS in Life Science, Pt Ravishankar Shukla University)
Chun-Chia Kung (Psychology, National Cheng Kung University)
Ai-Suan Lee (Department of Psychology, Universiti Tunku Abdul Rahman)
Nicole Legate (Psychology, Illinois Institute of Technology)
Juan David Leongómez (Facultad de Psicología, Universidad El Bosque)

Carmel A Levitan (Cognitive Science, Occidental College)
Hause Lin (Psychology, University of Toronto)
Samuel Lins (Psychology, University of Porto)
Qinglan Liu (Department of Psychology, Hubei University)
Marco Tullio Liuzza (Department of Surgical and Medical Sciences, Magna Graecia University of Catanzaro)
Johannes Lutz (Department of Psychology, University of Potsdam)
Harry Manley (Faculty of Psychology, Chulalongkorn University (Bangkok, Thailand))
Tara C Marshall (Department of Life Sciences, Brunel University London)
Randy J McCarthy (Center for the Study of Family Violence and Sexual Assault, Northern Illinois University)
Nicholas M Michalak (Psychology, University of Michigan)
Jeremy K Miller (Psychology, Willamette University)
Arash Monajem (Faculty of Psychology and Education, University of Tehran)
JA Muñoz-reyes (Laboratorio de Comportamiento Animal y Humano, Centro de Estudios Avanzados, Universidad de Playa Ancha)
Erica D Musser (Department of Psychology, Center for Children and Families, Florida Internationl University)
Lison Neyroud (LIP/PC2S, Université Grenoble Alpes)
Tonje Kvande Nielsen (Department of Psychology, University of Oslo)
Ceylan Okan, (Department of Social Sciences and Psychology, Western Sydney University)
Jerome Olsen (Department of Applied Psychology: Work, Education, and Economy, University of Vienna)
Asil Ali Özdoğru (Department of Psychology, Üsküdar University)
Babita Pande (SoS in Life Science, Pt Ravishankar Shukla University)
Arti Parganiha (SoS in Life Science, Pt Ravishankar Shukla University)
Noorshama Parveen (SoS in Life Science, Pt Ravishankar Shukla University)
Gerit Pfuhl (Department of Psychology, UiT The Arctic University of Norway)
Michael C Philipp (Psychology, Massey University)
Isabel R Pinto (Social Psychology Lab, University of Porto)
Pablo Polo (Laboratorio de Comportamiento Animal y Humano, Centro de Estudios Avanzados, Universidad de Playa Ancha)

Sraddha Pradhan (SoS in Life Science, Pt Ravishankar Shukla University)
John Protzko (Psychological and Brain Sciences, University of California, Santa Barbara)
Yue Qi (CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences)
Dongning Ren (Department of Social Psychology, Tilburg University)
Ivan Ropovik (Faculty of Education, University of Presov)
Nicholas O Rule (Psychology, University of Toronto)
Oscar R Sánchez (Facultad de Psicología, Universidad El Bosque)
S Adil Saribay (Psychology, Boğaziçi University)
Blair Saunders (Psychology, School of Social Sciences, University of Dundee)
Vidar Schei (Department of Strategy and Management, NHH Norwegian School of Economics)
Kathleen Schmidt (Psychology, Southern Illinois University Carbondale)
Martin Seehuus (Psychology, Middlebury College)
MohammadHasan Sharifian (Faculty of Psychology and Education, University of Tehran)
Victor Kenji M Shiramizu (Brain Institute, UFRN)
Almog Simchon (Psychology, Ben-Gurion University of the Negev)
Margaret Messiah Singh (SoS in Life Science, Pandit Ravishankar Shukla University)
Miroslav Sirota (Deaprtment of Psychology, University of Essex)
Guyan Sloane (Department of Psychology, University of Essex)
Sara Álvarez Solas (Biociencias, Universidad Regional Amazónica Ikiam)
Tiago Jessé Souza de Lima (Department of Psychology, University of Fortaleza)
Ian D Stephen (Department of Psychology, Macquarie University)
Stefan Stieger (Department of Psychology, Karl Landsteiner University of Health Sciences)
Daniel Storage (Psychology, University of Illinois)
Therese E Sverdrup (Department of Strategy and Management, NHH)
Peter Szecsi (Department of Affective Psychology, Eötvös Loránd University)
Christian K Tamnes (Department of Psychology, University of Oslo)

Chrystalle B Y Tan (Department of Community and Family Medicine, Universiti Malaysia Sabah)
Martin Thirkettle (Department of Psychology, Sociology & Politics, Sheffield Hallam University)
Dong Tiantian (China, QuFu Normal University)
Enrique Turiegano (Biology, Universidad autónoma de Madrid)
Kim Uittenhove (Department of developmental psychology, University of Geneva)
Heather L Urry (Psychology, Tufts University)
Eugenio Valderrama (Facultad de Psicología, Universidad El Bosque)
Jaroslava Varella Valentova (Department of Experimental Psychology, Institute of Psychology, University of Sao Paulo)
Nicolas Van der Linden (Center for Social and Cultural Psychology, Université Libre de Bruxelles (ULB))
Wolf Vanpaemel (Faculty of Psychology and Educational Sciences, University of Leuven)
Varella, M A C (Dept of Experimental Psychology, Institute of Psychology, University of São Paulo)
Milena Vásquez-Amézquita (Facultad de Psicología, Universidad El Bosque)
Leigh Ann Vaughn (Psychology, Ithaca College)
Evie Vergauwe (Department of Psychology and Educational Sciences, University of Geneva)
Michelangelo Vianello (Department of Philosophy, Sociology, Education and Applied Psychology, University of Padova)
Tan Kok Wei (School of Psychology and Clinical Language Sciences, University of Reading Malaysia)
David White (School of Psychology, UNSW Sydney)
John Paul Wilson (Psychology, Montclair State University)
Anna Wlodarczyk (Escuela de Psicología, Universidad Católica del Norte)
Qi Wu (Psychology, Liaoning Normal University)
Wen-Jing Yan (Institute of Psychology and Behavior Sciences, Wenzhou University)
Xin Yang (Psychology, Yale University)

Ilya Zakharov (Developmental Behavioral Genetics Lab, Psychological Institute of Russian Academy of Education)
Janis H Zickfeld (Department of Psychology, University of Oslo)
Christopher R Chartier (Department of Psychology, Ashland University)

Benedict Jones, Lisa DeBruine and Jessica Flake are joint first authors. Christopher Chartier (last author) is the Director of the Psychological Science Accelerator. All other authors are listed in alphabetical order.

Author contributions

Benedict Jones, Lisa DeBruine and Jessica Flake proposed and designed the project, designed the analysis plan, drafted and revised the Stage 1 submission, will carry out data collection. Christopher Chartier is the Director of the Psychological Science Accelerator, will carry out data collection, drafted and revised Stage 1 submission. All other authors had input into design of project and analysis plan, revised the Stage 1 submission, will carry out data collection

Funding

Hans IJzerman is supported by French National Research Agency "Investissements d'avenir" program (ANR-15-IDEX-02)
Coralie Chevallier is supported by ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL; Yue Qi is suported by Beijing Natural Science Foundation (5184035) and the Scientific Foundation of the Institute of Psychology, Chinese Academy of Sciences (Y5CX122005); Lisa M. DeBruine is supported by ERC KINSHIP; Muñoz-reyes is supported by Fondecyt regular 1170513; Erica D. Musser is suported by National Institutes of Mental Health R03-MH110812-02; Wen-Jing Yan is supported by National Natural Science Foundation of China (31500875); Ravin Alaei is supported by Social Sciences and Humanities Research Council of Canada; González-Santoyo is supported by PAPIIT UNAM IA209416 and Project CONACYT Ciencia Básica 241744; Pablo Polo is supported by Partially supported by grant FONDECYT regular 1170513; Tripat Gill is supported by Social Science and Humanities Research Council of Canada; Nicholas O. Rule is supported by Social Sciences and

Humanities Research Council of Canada; Eric Hehman is supported by SSHRC Insight Development Grant (430-2016-00094); David White is supported by Supported by an Australian Research Council Linkage Project grant (LP160101523); Evie Vergauwe is supported by Swiss National Science Foundation PZ00P1_154911; Nicholas A. Coles is supported by National Science Foundation Graduate Research Fellowship #R010138018; Michael Inzlicht is supported by This research was supported by grant RGPIN-2014-03744 from the Natural Sciences and Engineering Research Council of Canada; Lison Neyroud is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR15IDEX02); Krystian Barzykowski is supported by National Science Centre, Poland (2015/19/D/HS6/00641).

Social perception of faces around the world: How well does the valence-dominance model generalize across world regions?

Abstract

Over the last ten years, Oosterhof and Todorov's (2008) valence-dominance model of social judgments of faces has emerged as the most prominent account of how we evaluate faces on social dimensions. In this model, two dimensions (valence and dominance) underpin social judgments of faces. How well this model generalizes across world regions is a critical, yet unanswered, question. We will address this question by replicating Oosterhof and Todorov's (2008) methodology across all world regions (Africa, Asia, Central America and Mexico, Eastern Europe, Middle East, USA and Canada, Australia and New Zealand, Scandinavia, South America, UK, Western Europe, total N ≥ 9525) and using a diverse set of face stimuli. If we uncover systematic regional differences in social judgments, this will fundamentally change how social perception research is done and interpreted. If we find consistency across regions, this will ground future theory in an appropriately powered empirical test of an underlying assumption.

Introduction

People quickly and involuntarily form impressions of others based on their facial appearance (Olivola & Todorov, 2010; Ritchie et al., 2017; Willis & Todorov, 2006). These impressions then influence important social outcomes (Olivola et al., 2014; Todorov et al., 2015). For example, people are more likely to cooperate in socioeconomic interactions with individuals whose faces are evaluated as more trustworthy (Van 't Wout & Sanfey, 2008), vote for individuals whose faces are evaluated as more competent (Todorov et al., 2005), and seek romantic relationships with individuals whose faces are evaluated as more attractive (Langlois et al., 2000). Facial appearance can even influence life-or-death outcomes. For example, untrustworthy-looking defendants are more likely to receive death sentences (Wilson & Rule, 2015). Given evaluations of faces influence social outcomes, understanding how people in society evaluate others' faces can provide insight into a potentially important route through which social stereotypes impact behavior (Jack & Schyns, 2017; Todorov et al., 2008).

Over the last decade, the valence-dominance model (Oosterhof & Todorov, 2008) has emerged as the most prominent account of how we evaluate faces on social dimensions (840 citations in Google Scholar at May 10th 2018). Oosterhof and Todorov (2008) identified 13 different traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, unhappiness, intelligence, meanness, responsibility, sociability, trustworthiness, and weirdness) that perceivers spontaneously evaluate faces on when forming trait impressions. From these traits they derived a two-dimensional model of perception: *valence* and *dominance*. *Valence,* best characterized by rated trustworthiness, was defined as the extent to which the target was perceived as having the *intention* to harm the viewer (Oosterhof & Todorov, 2008). *Dominance,* best characterized by rated dominance, was defined as the extent to which the target was perceived as having the *ability* to inflict harm on the viewer (Oosterhof & Todorov, 2008). Crucially, the model proposes that these two dimensions are sufficient to drive social evaluations of faces. As a consequence, the majority of research on the effects of social evaluations of faces has focused on one or both of these dimensions (see Olivola et al., 2014 and Todorov et al., 2015 for reviews).

The valence-dominance model is widely employed in research investigating person perception, with little challenge to its assumed universality (Sutherland et al., 2018; Wang et al., 2018). Successful replications of this model have only been conducted in Western samples (Morrison et al., 2017; Wang et al., 2016). This focus on Western samples is consistent with research on human behavior more broadly, which typically draws general assumptions from analyses of Western participants' responses (Henrich et al., 2010). Kline et al. (2018) recently termed this problematic practice the *Western centrality assumption* and argued that regional variation, rather than universality, is likely the default for human behavior. Indeed, two recent studies of social evaluation of faces by Chinese participants (Sutherland et al., 2018; Wang et al., 2018) found that Chinese participants' social evaluations of faces were underpinned by a valence dimension similar to that reported for Western participants by Oosterhoff and Todorov (2008), but not by a corresponding dominance dimension. Instead, both studies reported a second dimension, referred to as *capability*, that was best characterized by rated intelligence. These results demonstrate that the Western centrality assumption is an important barrier to understanding how people evaluate faces on social dimensions. Crucially, these studies also suggest that the valence-dominance model is not a universal account of social evaluations of faces. While these studies demonstrate that the valence-dominance model is not perfectly universal, the extent of its global generality is an open, but important, question.

To establish the generalizability of the valence-dominance model across world regions, we will replicate Oosterhof and Todorov's (2008) methodology in a wide range of world regions (Africa, Asia, Middle East, Central America and Mexico, USA and Canada, Eastern Europe, Western Europe, Australia and New Zealand, Scandinavia, South America, UK).

Our study will be the most comprehensive test to date of social evaluations of faces. Table 1 details the world regions we will examine, the countries from those regions where we will carry out testing, the researchers responsible for carrying out that testing, and the number of raters each research group will collect data from (planned total number of raters = 9425). Participating research groups were recruited via the Psychological Science Accelerator project (Chartier et al., 2018; Chawla, 2017). Previous studies compared two cultures to demonstrate regional differences (Sutherland et al., 2018; Wang et al., 2018). By contrast, the scale and scope of our study will allow us to generate the most comprehensive picture of how social evaluations of faces differ across the world. Our accepted registered protocol will be posted on the Open Science Framework.

World region	**Countries and regions**	**Researchers**	**Number of raters**
Africa	Kenya, South Africa	Chaning Jang, Vinet Coetzee	250
Asia	China, India, Malaysia, Taiwan, Thailand	Dong Tiantian, Sun Juncai, Wen-Jing Yan, Chuan-Peng Hu, Yue Qi, Qinglan Liu, Zhongqing Jiang, Qi Wu, Arti Parganiha, Steve Janssen, Ai-Suan Lee, Tan Kok Wei, Chun-Chia Kung, Sau-Chin Chen, Harry Manley, Pratibha Kujur, Sraddha Pradhan Noorshama Parveen, Chrystalle Tan, Margaret Messiah Singh, Priyanka Chandel, Babita Pande	1050
Middle East	Iran, Israel, Turkey	Mohamma Hasan Sharifian, Eva Gilboa-Schechtman, Michael Gilead, Almog Simchon, Sinan Alper, Asil Özdoğru, Adil Saribay, Aycan Kapucu	600
Central America and Mexico	Ecuador, El Salvador, Mexico	Sara Álvarez Solas, Carlota Batres, Isaac González-Santoyo	250
USA and Canada	USA, Canada	Daniel Ansari, Hause Lin, Michael Inzlicht, Nick Rule, Ravin Alaei, Eric Hehman, Sally Xie, Tripat Gill, Daniel Storage, Cody Christopherson, Kathleen Schmidt, Nikki Legate, Randy McCarthy, Jeremy Miller, Gwen Gardiner, Chris Chartier, Dustin Calvillo, Nicholas Coles, Nicholas Michalak, Amanda Hahn, Martin Seehuus, Carmel Levitan, Michael Andreychik, Erica Musser, Yarrow Dunham, Xin Yang, Heather Urry, Ernest Baskin, William Chopik, Jack Arnal, Alexander Danvers, Corey Cook, John Paul Wilson, Patrick Forscher, Leigh Ann Vaughn, John Protzko	2525
Eastern Europe	Hungary,	Balazs Aczel, Vilius Dranseika, Krystian	775

	Lithuania, Poland, Russia, Serbia, Slovakia	Barzykowski, Ilya Zakharov, Vanja Kovic, Pavol Kačmár, Gabriel Baník, Ivan Ropovik, Matúš Adamkovič, Peter Babinčák, Peter Szecsi	
Western Europe	Austria, Belgium, France, Switzerland, Germany, Italy, Netherlands, Portugal, Spain, Switzerland	Stefan Stieger, Jerome Olsen, Wolf Vanpaemel, Nicolas Van der Linden, Armand Chatard, Coralie Chevallier, Kaminski Gwenaël, Hans IJzerman, Lison Neyroud, Johannes Lutz, Michelangelo Vianello, Marco Tullio Liuzza, Dongning Ren, Mark Brandt, Bastian Jaeger, Samuel Lins, Enrique Turiégano, Evie Vergauwe, Kim Uittenhove	1325
Australia and New Zealand	Australia, New Zealand	Khandis Blake, Ian Stephen, David White, Barnaby Dixson, Monica Koehn, Ceylan Okan, Michael Philipp, Matt Crawford	725
Scandinavia	Norway	Christian Tamnes, Tonje Kvande Nielsen, Janis Zickfeld, Vidar Schei, Therese Sverdrup, Gerit Pfuhl	325
South America	Argentina, Brazil, Chile, Colombia	Debora Burin, Natalia Irrazabal, Victor Shiramizu, Tiago Jessé Souza de Lima, Jaroslava Varella Valentova, Marco Antonio Correa Varella, José Antonio Muñoz Reyes, Pablo Polo Rodrigo, Anna Wlodarczyk, Ana María Fernández, Juan David Leongómez, Oscar Sánchez, Milena Vasquez-Amézquita, Eugenio Valderrama, Martha Lucia Borras Guevara	950
UK	England, Scotland, Wales	Melissa Colloff, Heather Flowe, Blair Saunders, Benedict Jones, Lisa DeBruine, Miroslav Sirota, Guyan Sloane, Martin Thirkettle, Tara Marshall, Thomas Rhys Evans	650

Table 1. The world regions we will examine, the countries from those regions where we will collect data, the researchers carrying out that testing, and the total number of raters data will be collected from in each region. Researchers will each collect data from between 50 and 200 raters.

Methods

Procedure

Oosterhof and Todorov (2008) derived their valence-dominance model from a principal component analysis of ratings (by US raters) of 62 faces for 13 different traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, unhappiness, intelligence, meanness, responsibility, sociability, trustworthiness, and weirdness). Using the criteria of the number of components with an Eigenvalue >1, this analysis produced two principal components. The first component explained 63% of the variance in trait ratings, was strongly correlated with rated trustworthiness, and weakly correlated with rated dominance. The second component explained 18% of the variance in trait ratings, was strongly correlated with rated dominance, and weakly correlated with rated trustworthiness. We will replicate Oosterhof and Todorov's method in each world region we examine.

Stimuli in our study will be an open-access, full-color, face image set (49 women, 53 men, diverse ethnicity), taken under standardized photographic conditions (DeBruine & Jones, 2017). Like Oosterhof and Todorov (2008), the individuals photographed are posed looking directly at the camera with a neutral expression. Like Oosterhof and Todorov (2008), background, lighting and clothing (here, a white t-shirt) are constant across images.

In our study, adult raters will be randomly allocated to rate all 102 faces for one of the 13 adjectives tested by Oosterhof and Todorov (aggressive, attractive, caring, confident, dominant, emotionally stable, unhappy, intelligent, mean, responsible, sociable, trustworthy, weird). Following Oosterhof and Todorov (2008), ratings will be made using 1 (not at all) to 7 (very) scales, the order in which faces will be presented (i.e., trial order) will be fully randomized, and the rating task will be self-paced. See a demo of the English-language version here: http://faceresearch.org/project?PSAeng&auto and an example trial in Figure 1. Because all researchers will collect data through an identical interface (excepting differences in instruction language), data collection protocols will be highly standardized across labs.

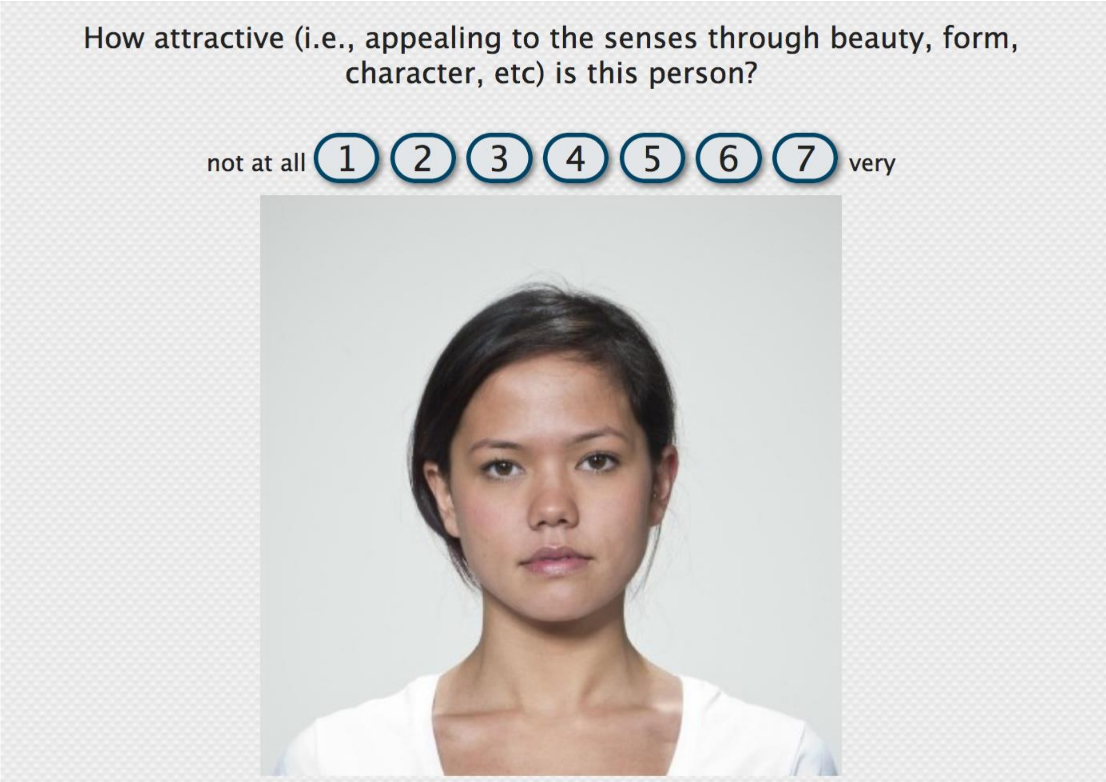

Figure 1. An example of a rating-task trial from a block where faces would be rated for attractiveness.

After completing the rating task, raters will complete a short questionnaire requesting demographic information (sex, age, ethnicity). These variables were not considered in Oosterhof and Todorov's analyses but will be collected in our study so that other researchers can use them in secondary analyses of the published data. The data from this study will be by far the largest and most comprehensive open access set of face ratings from around the world with open stimuli, proving an invaluable resource for further research addressing the Western centrality assumption in person perception research.

Raters will complete the task in a language appropriate for their country (see Translations guidelines section, below, for details of our procedure for translating instructions). To mitigate potential problems with translating single-word labels, dictionary definitions for each of the 13 traits will be provided. Twelve of these dictionary definitions have previously been used to test for effects of social impressions on the memorability of face photographs (Bainbridge et al., 2013). Dominance (not rated in Bainbridge et al., 2013) will be defined as "strong; important". All definitions (and other instructions) in all languages used will be made publicly available on the Open Science Framework.

Raters

We plan to test a total of 9425 raters (see Table 1 for break down by world region). In each world region, at least 15 different raters will rate each of the 13 traits. This minimum number of raters per trait in each world region was chosen following simulations we ran (see https://osf.io/x7fus/ for code and data) that sampled from a population of 2513 raters, each of whom had rated the attractiveness of the 102 faces that will be used in our study. These simulations showed that >99% of 1000 random samples of 15 raters produced Cronbach's alphas >.8. This indicates that the dependent variable in our analysis (averages of ratings from 15 or more raters) will be highly reliable. Each research group has approval from their local Ethics Committee or IRB to conduct the study, has explicitly indicated that their institution does not require approval for the researchers to conduct this type of face-rating task, or has explicitly indicated that the current study is covered by a preexisting approval. Data collection will be completed by May 1st 2019.

Analysis plan

For each world region studied, our analyses will directly replicate the principal component analysis reported by Oosterhof and Todorov (2008). This will test the theoretical model proposed by Oosterhof and Todorov (2008) in each world region. Ratings from each world region will be analyzed separately. Raw data (anonymous) will be published on the Open Science Framework. First, we will calculate the average rating for each face separately for each of the 13 traits. Like Oosterhof and Todorov (2008), we will then subject these mean ratings to principal component analysis with orthogonal components and no rotation. Using the criteria reported in Oosterhof and Todorov's (2008) paper, we will retain and interpret the components with an Eigenvalue > 1. The code that will be used for these analyses is publicly available at the Open Science Framework (https://osf.io/87rbg/) and included in the supplemental materials of this manuscript.

Criteria for replicating Oosterhof and Todorov's model

Oosterhof and Todorov's valence-dominance model will be judged to have been replicated in a given world region if the first two components both have Eigenvalues > 1, the first component (i.e., the one explaining more of the variance in ratings) is correlated strongly (loading > .5) with trustworthiness and weakly (loading < .5) with dominance, and the second component (i.e., the one explaining less of the variance in ratings) is correlated strongly (loading > .5) with dominance and weakly (loading < .5) with trustworthiness. All three criteria need to be met to conclude that the model was replicated in a given world region.

Exclusions

Following Oosterhof and Todorov (2008), data from raters who fail to complete all 102 ratings will be excluded from analyses. Data from raters who provide invariant responses (i.e., give the same rating for 75% or more of the faces) will also be excluded from analyses. There will be no other rater exclusions.

Data-quality check

Following previous research testing the valence-dominance model (Morrison et al., 2017; Oostehof & Todorov, 2008; Wang et al., 2018), data quality will be checked by calculating the inter-rater agreement (indicated by Cronbach's alpha) for each trait separately for each world region. A trait will only be included in the analysis for that world region if this alpha is greater than .70. Where alpha is <.70, this low inter-rater agreement will be reported and discussed.

Power analysis

Simulations show we have >95% power to detect the key effect of interest (two components meeting the criteria for replicating Oosterhof & Todorov that are described above). We used the open data from Morrison et al's (2017) replication of Oosterhof & Todorov (2008) to generate a variance-covariance matrix representative of typical interrelationships among the 13 traits that will be tested in our study. We then generated 1000 samples of 102 faces from these distributions and ran our planned principal component analysis (which is identical to that reported by Oosterhof & Todorov, 2008) on each sample (see https://osf.io/x7fus/ for code and data). Results of 100% of these analyses matched our criteria for replicating Oosterhof & Todorov (2008). This demonstrates that 102 faces will give us >95% power to replicate Oosterhof and Todorov's results.

Robustness analyses

Oostehorf and Todorov (2008) extracted and interpreted components with an Eigenvalue > 1 using an unrotated principal components analysis. As described above, we will directly replicate their methodology in our main analyses. However, we acknowledge that this type of analysis has been criticized. First, it has been argued that exploratory factor analysis with rotation, rather than an unrotated principal components analysis, is more appropriate when one intends to measure correlated latent factors, as is the case in the current study (e.g., Fabrigar et al., 1999; Park et al., 2002). Second, the extraction rule of Eigenvalues >1 has been criticized for not indicating the optimal number of components, as well as producing unreliable components (e.g., Cliff, 1988; Zwick & Velicer, 1986). To address these methodological limitations, we will repeat our main analyses, this time using exploratory factor analysis with an oblimin rotation as the model and a parallel analysis (Horn, 1965) as the extraction method. We will use parallel analysis as the extraction method because it has been described as yielding the optimal number of components (or factors) across the largest array of scenarios (Fabrigar et al., 1999; O'Connor, 2000; Schmitt, 2011). The purpose of these additional analyses is twofold. First, to address potential methodological limitations in the original study and, second, to ensure that the results of our replication of Oostehorf and Todorov's (2008) study are robust to the implementation of those more rigorous analytical techniques. The code that will be used for these robustness analyses is publicly available at the Open Science Framework (https://osf.io/87rbg/) and included in the supplemental materials of this manuscript. The same critieria for replicating Oosterhof and Todorov's model that are described above (see "Criteria for replicating Oosterhof and Todorov's model") will be applied to this analysis

Translation guidelines

This section describes the procedure we will use to translate instructions, trait labels, and trait definitions from English to the languages to be used for testing in each country. This process reflects and extends best practice in translating for cross-cultural research, as described in Brislin (1970).

Translation Personnel

Language Coordinator: Will coordinate translation process and discuss final version with translators.
"A" Translators: Will translate from English to target language and discuss final version with coordinator and B Translators (N=2, both bilingual).
"B" Translators: Will translate from target language to English and discuss final version with coordinator and A Translators (N=2, both bilingual).
External Readers: Will read materials for final clarity check (N=2, both non-academics).
Individual researchers (or research groups) carrying out data collection: Will provide final checks and suggest any necessary cultural adjustments.

Translation Process

Step 1 (Translation). Original document is translated from English to target language by A Translators resulting in document Version A.
Step 2 (Back-translation). Version A is translated back from target language to English by B Translators independently resulting in Version B.
Step 3 (Discussion). Version A and B are discussed among translators and the language coordinator, discrepancies in Version A and B are detected and solutions are discussed. Version C is created.
Step 4 (External readings). Version C is tested on two non-academics fluent in the target language. Members of the fluent group are asked how they perceive and understand the translation. Possible misunderstandings are noted and again discussed as in Step 3.
Step 5 (Possible cultural adjustments). Data collection labs read materials and identify any adjustments for their local participant sample. Adjustments are

discussed with the Language Coordinator, who makes any necessary changes, resulting in the final version for each site.

This process will produce the Final Translated Document, containing the instructions that will be used in the study.

Conclusions

If our project uncovers systematic regional differences in how people make social judgments from physical appearance, this will fundamentally change the way social perception research is done and interpreted. If we find consistency across world regions, this will ground future theory in an appropriately powered empirical test of an underlying assumption. As a result, the outcome of this project will have a large and lasting impact on social perception research.

References

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. Journal of Experimental Psychology: General, 142, 1323-1334.

Brislin, R. W. (1970). Back-translation for cross-cultural research. Journal of Cross-Cultural Psychology, 1, 185-216.

Chartier, C, McCarthy, R., & Urry, H. (2018). The Psychological Science Accelerator, APS Observer, 31, 30.

Chawla, D. S. (2017). A new ‘accelerator’ aims to bring big science to psychology. Science, doi:10.1126/science.aar4464

Cliff, N. (1988). The Eigenvalues-greater-than-one rule and the reliability of components. Psychological Bulletin, 103, 276–279.

DeBruine, L., & Jones, B. (2017). Face Research Lab London Set (Version 3). Figshare. https://doi.org/10.6084/m9.figshare.5047666.v3

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, 272–299.

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33, 61-83.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179–185.

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. Annual Review of Psychology, 68, 269-297.

Kline, M. A., Shamsudheen, R., & Broesch, T. (2018). Variation is the universal: Making cultural evolution work in developmental psychology. Philosophical Transactions of the. Royal . Society. B, 373, 20170059.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. Psychological Bulletin, 126, 390-423.

Morrison, D., Wang, H., Hahn, A. C., Jones, B. C., & DeBruine, L. M. (2017). Predicting the reward value of faces and bodies from social perception. PLoS ONE, 12, e0185093.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. Behavior Research Methods, Instruments, & Computers, 32, 396–402.

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. Journal of Nonverbal Behavior, 34, 83-110.

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. Trends in Cognitive Sciences, 18, 566-570.

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

Park, H. S., Dailey, R., & Lemus, D. (2002). The use of exploratory factor analysis and principal components analysis in communication research. Human Communication Research, 28, 562-577.

Ritchie, K. L., Palermo, R., & Rhodes, G. (2017). Forming impressions of facial attractiveness is mandatory. Scientific Reports, 7, 469.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. Journal of Psychoeducational Assessment, 29, 304–321.

Sutherland, C. A. M., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial first impressions across culture: Data-driven modeling of Chinese and British perceivers' unconstrained facial impressions. Personality and Social Psychology Bulletin, 44, 521-537.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. Science, 308, 1623-1626.

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, 519-545.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. Trends in Cognitive Sciences, 12, 455-460.

Van 't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. Cognition, 108, 796-803.

Wang, H., Hahn, A. C., DeBruine, L. M., & Jones, B. C. (2016). The motivational salience of faces is related to both their valence and dominance. PLoS ONE, 11, e0161114.

Wang, H., Han, C., Hahn, A., Fasolt, V., Morrison, D.,... Jones, B. C. (2018). A data-driven study of Chinese participants' social judgments of Chinese faces. PsyArXiv.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. Psychological Science, 17, 592-598.

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. Psychological Science, 26, 1325-1331.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. Psychological Bulletin, 99, 432–442.