To Which World Regions Does the Valence-Dominance Model of Social Perception Apply?

(PSA001; Registered Report Stage 2)

This manuscript is under review at *Nature Human Behaviour.*

(BCJ, LMD, JKF joint first authors)

Benedict C. Jones[1], Lisa M. DeBruine[1], Jessica K. Flake[2], Marco Tullio Liuzza[3], Jan Antfolk[4], Nwadiogo C. Arinze[5], Izuchukwu L. G. Ndukaihe[5], Nicholas G. Bloxsom[6], Savannah C. Lewis[6], Francesco Foroni[7], Megan L. Willis[7], Carmelo P. Cubillas[8], Miguel A. Vadillo[8], Michael Gilead[9], Almog Simchon[9], S. Adil Saribay[10], Nicholas C. Owsley[11], Dustin P. Calvillo[12], Anna Wlodarczyk[13], Yue Qi[14], Kris Ariyabuddhiphongs[15], Somboon Jarukasemthawee[15], Harry Manley[15], Panita Suavansri[15], Nattasuda Taephant[15], Ryan M. Stolier[16], Thomas R. Evans[17], Judson Bonick[18], Jan W. Lindemans[18], , Logan F. Ashworth[19], Coralie Chevallier[20], Aycan Kapucu[21], Aslan Karaaslan[21], Juan David Leongómez[22], Oscar R. Sánchez[22], Eugenio Valderrama[22], Milena Vásquez-Amézquita[22], Balazs Aczel[24], Nandor Hajdu[23, 24], Peter Szecsi[24], Michael Andreychik[25], Erica D. Musser[26], Carlota Batres[27], Chuan-Peng Hu[28], Qing-Lan Liu[29], Nicole Legate[30], Leigh Ann Vaughn[31], Krystian Barzykowski[32], Karolina Golik[32], Irina Schmid[33], Stefan Stieger[33], Richard Artner[34], Chiel Mues[34], Wolf Vanpaemel[35], Zhongqing Jiang[36], Qi Wu[36], Gabriela M. Marcu[37], Ian D. Stephen[38], Jackson G. Lu[39], Michael C. Philipp[40], Jack D. Arnal[41], Eric Hehman[2], Sally Y. Xie[2], William J. Chopik[42], Martin Seehuus[43], Soufian Azouaghe[44, 79], Abdelkarim Belhaj[44], Jamal Elouafa[44], John P. Wilson[45], Elliott Kruse[46], Marietta Papadatou-Pastou[47], Alan E. Barba-Sãnchez[48], Anabel De La Rosa-Gómez[48], Isaac González-Santoyo[49], Tsuyueh Hsu[50], Chun-Chia Kung[50], Hsiao-Hsin Wang[50], Jonathan B. Freeman[51], DongWon Oh[52], Vidar Schei[53], Therese E. Sverdrup[53], Carmel A. Levitan[54], Corey L. Cook [55], Priyanka Chandel[56], Pratibha Kujur[56], Arti Parganiha[56], Noorshama Parveen[56], Atanu Kumar Pati[56], Sraddha Pradhan[56], Margaret M. Singh[56], Babita Pande[57], Jozef Bavolar[58], Pavol Kačmár[58], Ilya Zakharov[59], Sara Álvarez-Solas[60], Ernest Baskin[61], Martin Thirkettle[62], Kathleen Schmidt[63], Cody D. Christopherson[64], Jordan W. Suchow[65], Jonas K. Olofsson[66], Ai-Suan Lee[67], Jennifer L. Beaudry[68], Taylor D. Gogan[68], Julian A. Oldmeadow[68], Barnaby J. W. Dixson[120], Laura M. Stevens[70], Gianni Ribeiro[120], Mark J. Brandt[72], Karlijn Hoyer[72], Bastian Jaeger[72], Dongning Ren[72], Willem W. A. Sleegers[72], Joeri Wissink[72], Gwenaël Kaminski[73], Victoria A. Floerke[74], Heather L. Urry[74], Sau-Chin Chen[75], Gerit Pfuhl[76], Zahir Vally[77], Dana M. Basnight-Brown[78], Hans IJzerman[79], Elisa Sarda[79], Touhami Badidi[81], Nicolas Van der Linden[82], Chrystalle B. Y. Tan[83], Vanja Kovic[84], Melissa F. Colloff[70], Heather D. Flowe[70]Debora I. Burin[87], Gwendolyn Gardiner[88], John Protzko[89], Christoph Schild[90], Karolina A. Scigala[90], Ingo Zettler[90], Erin M. O'Mara Kunz[91], Daniel Storage[92], Fieke M. A. Wagemans[93], Blair Saunders[94], Miroslav Sirota[95], Guyan V. Sloane[95], Tiago J. S. Lima[96], Kim Uittenhove[97], Evie Vergauwe[97], Katarzyna Jaworska[1], Lilian Carvalho, Karl Ask[99], Casper J. J. van Zyl[100], Anita Körner[101], Sophia C. Weissgerber[102], Jordane Boudesseul[103], Fernando Ruiz-Dodobara[103], Kay L. Ritchie[104], Nicholas M. Michalak[105], Khandis R. Blake[106, 142], David White[106], Alasdair R. Gordon-Finlayson[107], Michele Anne[108], Steve M. J. Janssen[108], Kean Mun Lee[108], Tonje K. Nielsen[109], Christian K. Tamnes[109], Janis H. Zickfeld[109], Anna Dalla Rosa[110], Ferenc Kocsor[111], Luca Kozma[111], Ádám Putz[111], Patrizio Tressoldi[112], Michelangelo Vianello[110], Natalia Irrazabal[113], Armand Chatard[114], Samuel Lins[115], Isabel R. Pinto[115], Johannes Lutz[116], Matus Adamkovic[117], Peter Babincak[117], Gabriel Baník[117], Ivan Ropovik[118,143], Vinet Coetzee[119], Kim Peters[120], Niklas K. Steffens[120,] Kok Wei Tan[121], Christopher A. Thorstenson[122], Ana Maria Fernandez[123], Rafael M. C. S. Hsu[124], Jaroslava V. Valentova[124], Marco A. C. Varella[124], Nadia S. Corral-Frías[125], Martha Frias-Armenta[125], Javad Hatami[126], Arash Monajem[126], MohammadHasan Sharifian[126], Brooke Frohlich[127], Hause Lin[128], Michael Inzlicht[128], Claus Lamm[129], Ekaterina Pronizius[129], Martin Voracek[129], Jerome Olsen[130], Erik Mac Giolla[131], Aysegul Akgoz[132], Asil A. Özdoğru[132], Matthew T. Crawford[133], Brooke Bennett-Day[134], Monica A. Koehn[135], Ceylan Okan[136], Daniel Ansari[137], Tripat Gill[138], Jeremy K. Miller[139], Yarrow Dunham[140], Xin Yang[140], Sinan Alper[141], Ravin Alaei[128], Martha Lucia Borras-Guevara[144], Sun Jun Cai[145], Alexander F. Danvers[146], Vilius Dranseika[147], Eva Gilboa-Schechtman[148], Amanda C. Hahn[19], Chaning Jang[11], Tara C. Marshall[149], Randy J. McCarthy[150], Jose Antonio Muñoz-reyes[151], Lison Neyroud[80], Pablo Polo[151], Nicholas Rule[128], Victor K.M. Shiramazu[152], Dong Tiantian[145], Enrique Turiegano[8], Wen-Jing Yan[153], Benjamin Balas[69], Paulo Ferreira[154], Julia Jünger[98], Georgina Mburu[11], Walido

Sampaio[155], Diana Santos[154], Sami Gülgöz[71], Julia Stern[98], Patrick S. Forscher[80], Christopher R. Chartier[6], Nicholas A. Coles[127]

[1]University of Glasgow, Institute of Neuroscience and Psychology, UK. [2]McGill University, Department of Psychology, Canada. [3]"Magna Graecia" University of Catanzaro, Department of Medical and Surgical Sciences, Italy. [4]Åbo Akademi University, Faculty of Arts, Psychology and Theology, Finland. [5]Alex Ekwueme Federal University Ndufu Alike, Department of Psychology, Nigeria. [6]Ashland University, Department of Psychology, USA. [7]Australian Catholic University, School of Behavioural and Health Sciences, Australia. [8]Autonomous University of Madrid, Department of Basic Psychology, Spain. [9]Ben-Gurion University of the Negev, Department of Psychology, Israel. [10]Boğaziçi University, Department of Psychology, Turkey. [11]Busara Center for Behavioral Economics, Kenya. [12]California State University San Marcos, Psychology Department, USA. [13]Catholic University of the North, School of Psychology, Chile. [14]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. [15]Chulalongkorn University, Faculty of Psychology, Thailand. [16]Columbia University, Department of Psychology, USA. [17]Coventry University, School of Psychological, Social and Behavioural Sciences, UK. [18]Duke University, Center for Advanced Hindsight, USA. [19]Humboldt State University, Department of Psychology, USA. [20]Laboratoire de Neurosciences Cognitives et Computationnelles, Département d'Études Cognitives, INSERM U960, École Normale Supérieure, France. [21]Ege University, Psychology Department, Turkey. [22]Universidad El Bosque, Faculty of Psychology, Colombia. [23]ELTE Eötvös Loránd University, Doctoral School of Psychology, Hungary. [24]ELTE Eötvös Loránd University, Institute of Psychology, Hungary. [25]Fairfield University, Department of Psychology, USA. [26]Florida International University, Department of Psychology, USA. [27]Franklin and Marshall College, Department of Psychology, USA. [28]German Resilience Center, Germany. [29]Hubei University, Department of Psychology, China. [30]Illinois Institute of Technology, Department of Psychology, USA. [31]Ithaca College, Department of Psychology, USA. [32]Jagiellonian University, Institute of Psychology, Poland. [33]Karl Landsteiner University of Health Sciences, Department of Psychology and Psychodynamics, Austria. [34]Katholieke Universiteit Leuven, Research Group of Quantitative Psychology and Individual Differences, Belgium. [35]Katholieke Universiteit Leuven, Faculty of Psychology and Educational Sciences, Belgium. [36]Liaoning Normal University, Department of Psychology, China. [37]Lucian Blaga University of Sibiu, Department of Psychology, Romania. [38]Macquarie University, Department of Psychology, Australia. [39]Massachusetts Institute of Technology, Sloan School of Management, USA. [40]Massey University, School of Psychology, New Zealand. [41]McDaniel College, Psychology Department, USA. [42]Michigan State University, Department of Psychology, USA. [43]Middlebury College, Department of Psychology, USA. [44]Mohammed V University, Department of Psychology, Morocco. [45]Montclair State University, Psychology Department, USA. [46]Monterrey Institute of Technology and Higher Education, EGADE Business School, Mexico. [47]National and Kapodistrian University of Athens, School of Education, Greece. [48]National Autonomous University of Mexico, School of Higher Studies Iztacala, Mexico. [49]National Autonomous University of Mexico, Department of Psychology, [50]National Cheng Kung University, Department of Psychology, Taiwan. [51]New York University, Department of Psychology and Center for Neural Science, USA. [52]New York University, Department of Psychology, USA. [53]NHH Norwegian School of Economics, Department of Strategy and Management, Norway. [54]Occidental College, Department of Cognitive Science, USA. [55]Pacific Lutheran University, Department of Psychology, USA. [56]Pandit Ravishankar Shukla University, School of Studies in Life Science, India. [57]Pandit Ravishankar Shukla University, Center for Basic Sciences, India. [58]Pavol Jozef Šafárik University in Košice, Department of Psychology, Slovakia. [59]Psychological Institute of Russian Academy of Education, Developmental Behavioral Genetics Lab, Russia. [60]Universidad Regional Amazónica Ikiam, Department of Biosciences, Ecuador. [61]Saint Joseph's University, Department of Food Marketing, USA. [62]Sheffield Hallam University, Department of Psychology, Sociology and Politics, UK. [63]Southern Illinois University,

Department of Psychology, USA. [64]Southern Oregon University, Psychology Department, USA. [65]Stevens Institute of Technology, School of Business, USA. [66]Stockholm University, Department of Psychology, Sweden. [67]Sunway University, Department of Psychology, Malaysia. [68]Swinburne University of Technology, Department of Psychological Sciences, Australia.[69]North Dakota State University, Department of Psychology, USA. [70]University of Birmingham , School of Psychology, UK. [71]Koç University, Turkey. [72]Tilburg University, Department of Social Psychology, Netherlands. [73]Toulouse University, CLLE, France. [74]Tufts University, Department of Psychology, USA. [75]Tzu-Chi University, Department of Human Development and Psychology, Taiwan. [76]UiT The Arctic University of Norway, Department of Psychology, Norway. [77]United Arab Emirates University, Department of Psychology & Counseling, United Arab Emirates. [78]United States International University - Africa, Kenya. [79]Université Grenoble Alpes, LIPPC2S, France. [80]Université Grenoble Alpes, Department of Psychology, France. [81]Université Ibn Tofail Kénitra, Department of Psychology, Morocco. [82]Université Libre de Bruxelles, Center for Social and Cultural Psychology, Belgium. [83]Universiti Malaysia Sabah, Department of Community and Family Medicine, Malaysia. [84]University of Belgrade, Department of Psychology, Faculty of Philosophy, Serbia. [85]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. [87]Universidad de Buenos Aires, Instituto de Investigaciones, Facultad de Psicologia, Argentina. [88]University of California, Riverside, Department of Psychology, USA. [89]University of California, Santa Barbara, Department of Psychological & Brain Sciences, USA. [90]University of Copenhagen, Department of Psychology, Denmark. [91]University of Dayton, Department of Psychology, USA. [92]University of Denver, Department of Psychology, USA. [93]University of Duisburg-Essen, Institute for Socio-Economics, Germany. [94]University of Dundee, School of Social Sciences, UK. [95]University of Essex, Department of Psychology, UK. [96]University of Fortaleza, Department of Psychology, Brazil. [97]University of Geneva, Faculty of Psychology and Educational Sciences, Switzerland. [98]University of Goettingen, Department of Psychology, Germany. [99]University of Gothenburg, Department of Psychology, Sweden. [100]University of Johannesburg, Department of Psychology, South Africa. [101]University of Kassel, Department of Psychology, Germany. [102]University of Kassel, Department of Psychology, Germany. [103]University of Lima, Institute of Scientific Research, Peru. [104]University of Lincoln, School of Psychology, UK [105]University of Michigan, Department of Psychology, USA. [106]University of New South Wales Sydney, Evolution and Ecology Research Centre, Australia. [107]University of Northampton, Faculty of Health, Education & Society, UK. [108]University of Nottingham Malaysia, School of Psychology, Malaysia. [109]University of Oslo, Department of Psychology, Norway. [110]University of Padova, Department of Philosophy, Sociology, Education and Applied Psychology, Italy. [111]University of Pécs, Institute of Psychology, Hungary. [112]University of Padova, Department of General Psychology, Italy. [113]University of Palermo, Faculty of Social Sciences, Argentina. [114]University of Poitiers, Psychology Department, France. [115]University of Porto, Department of Psychology, Portugal. [116]University of Potsdam, Department of Psychology, Germany. [117]University of Presov, Institute of Psychology, Faculty of Arts, Slovakia. [118]University of Presov, Faculty of Education, Slovakia. [119]University of Pretoria, Department of Biochemistry, Genetics and Microbiology, South Africa. [120]University of Queensland, School of Psychology, Australia. [121]University of Reading Malaysia, School of Psychology and Clinical Language Sciences, Malaysia. [122]University of Rochester, Department of Clinical and Social Sciences in Psychology, USA. [123]University of Santiago, Chile, School of Psychology, Chile. [124]University of São Paulo, Department of Experimental Psychology, Institute of Psychology, Brazil. [125]University of Sonora, Department of Psychology, Mexico. [126]University of Tehran, Department of Psychology, Iran. [127]University of Tennessee, Knoxville, Department of Psychology, USA. [128]University of Toronto, Department of Psychology, Canada. [129]University of Vienna, Department of Basic Psychological Research and Research Methods, Austria. [130]University of Vienna, Department of Applied Psychology: Work, Education, and Economy, Austria. [131]University West, Department of Behavioral Sciences, Sweden. [132]Üsküdar University, Department of Psychology, Turkey. [133]Victoria University of

Wellington, School of Psychology, New Zealand. [134]Wesleyan College, Department of Psychology, USA. [135]Western Sydney University, School of Psychology, Australia. [136]Western Sydney University, School of Social Science and Psychology, Australia. [137]Western University, Department of Psychology, Canada. [138]Wilfrid Laurier University, Lazaridis School of Business & Economics, Canada. [139]Willamette University, Department of Psychology, USA. [140]Yale University, Department of Psychology, USA, [141]Yasar University, Department of Psychology, Turkey. [142]University of Melbourne, Melbourne School of Psychological Sciences, Australia. [143]Charles University, Faculty of Education, Institute for Research and Development of Education, Czechia, [144]University of St. Andrews, United Kingdom, [145]Qufu Normal University, China, [146]University of Oklahoma, USA, [147]Vilnius University, Lithuania, [148]Bar-Ilan University, Israel, [149]Brunel University London, UK, [150]Northern Illinois University, [150]Playa Ancha University of Educational Sciences, Chile, [152]Federal University of Rio Grande do Norte, Brazil, [153]Wenzhou University, China, [154]Universidade Federal da Grande Dourados, Brazil, [155]Universidade Federal de São Carlos, Brazil.

Funding: Claus Lamm was supported by the Viennese Science and Technology Fund (WWTF VRG13-007, to Claus Lamm); Lisa DeBruine was supported by ERC #647910 (KINSHIP); Luca Kozma, Ferenc Kocsor and Ádám Putz were supported by the European Social Fund (EFOP-3.6.1.-16-2016-00004 – "Comprehensive Development for Implementing Smart Specialization Strategies at the University of Pécs"). Kim Uittenhove and Evie Vergauwe were supported by a grant from the Swiss National Science Foundation (PZ00P1_154911 to Evie Vergauwe). Tripat Gill is supported by the Social Sciences and Humanities Research Council of Canada (SSHRC). Miguel Vadillo was supported by grants 2016-T1/SOC-1395 (Comunidad de Madrid) and PSI2017-85159-P (AEI/FEDER UE). Rystian Barzykowski was supported by a grant from the National Science Centre, Poland [No.: 2015/19/D/HS6/00641]. Judson Bonick and Jan W. Lindemans were supported by the Joep Lange Institute. Gabriel Baník was supported by Slovak Research and Development Agency [APVV-17-0418]. Hans IJzerman and Elisa Sarda were supported by a French National Research Agency "Investissements d'avenir" program grant (ANR-15-IDEX-02). The Raipur Group is thankful to (1) the University Grants Commission, New Delhi, India for the research grants received through its DRS-SAP (Phase-III) scheme sanctioned to the School of Studies in Life Science and (2) logistic support received from the Center for Translational Chronobiology at the School of Studies in Life Science, PRSU, Raipur, India. Karl Ask was supported by a small grant from the Department of Psychology, University of Gothenburg. Yue Qi was supported by grants from the Beijing Natural Science Foundation (5184035) and CAS Key Laboratory of Behavioral Science, Institute of Psychology. Nicholas Coles was supported by the National Science Foundation Graduate Research Fellowship R010138018.

Acknowledgments: We would like to acknowledge the following research assistants: Alan E. Barba-Sãnchez (National Autonomous University of Mexico); Jordan Muriithi and Joyce Ngugi (United States International University - Africa); Elisa Adamo, Viviana Ciambrone, Francesca Dolce, and Domenico Cafaro ("Magna Graecia" University of Catanzaro); Eugenia De Stefano (University of Padova); Samile A. Escobar Abadia (University of Lincoln); Lene Elisabeth Grimstad (NHH Norwegian School of Economics); Teodor Jernsäther (Stockholm University); Luciana Chavarria Zamora (Franklin and Marshall College); Ryan E Liang and Ruth Christy Lo (Universiti Tunku Abdul Rahman); Ashley Short and Liam Allen (Massey University, New Zealand)' Arda Ateş, Ezgi Güneş, and Salih Can Özdemir (Boğaziçi University); Ida Pedersen and Tove Roos (Åbo Akademi University); Nicole Paetz (Escuela de Comunicación Mónica Herrera); Johan Green (University of Gothenburg); Morris Krainz, BSc (University of Vienna, Austria); Boryana Todorova (University of Vienna, Austria).

The project contributions of each author according to author self-nominations using the CRediT taxonomy (https://www.casrai.org/credit.html). Light blue indicates a supporting role, dark blue a leading role; bold names indicate the leadership team, italics the administrative team. Figure format is adapted from Steinmetz (https://twitter.com/SteinmetzNeuro/status/1147241128858570752). At the time of submission we had not received CRediT responses from 36 co-authors, and they are omitted from the current version of the table. Figure detail and code at https://osf.io/6dbxg/

Abstract

Over the last ten years, Oosterhof and Todorov's valence-dominance model has emerged as the most prominent account of how people evaluate faces on social dimensions. In this model, two dimensions (valence and dominance) underpin social judgments of faces. Because this model has primarily been developed and tested in Western regions, it is unclear whether these findings apply to other regions. We addressed this question by replicating Oosterhof and Todorov's methodology across 11 world regions, 41 countries, and 11,481 participants. When we used Oosterhof's and Todorov's original analysis strategy, the valence-dominance model generalized across regions. When we used an alternative strategy that allowed for a more optimal number of correlated latent factors, we observed much less generalization. These results underscore how each analysis strategy embeds substantive assumptions that can strongly influence theoretical conclusions.

To Which World Regions Does the Valence-Dominance Model of Social Perception Apply?

People quickly and involuntarily form impressions of others based on their facial appearance[1-3]. These impressions then influence important social outcomes[4,5]. For example, people are more likely to cooperate in socioeconomic interactions with individuals whose faces are evaluated as more trustworthy[6], vote for individuals whose faces are evaluated as more competent[7], and seek romantic relationships with individuals whose faces are evaluated as more attractive[8]. Facial appearance can even influence life-or-death outcomes. For example, untrustworthy-looking defendants are more likely to receive death sentences[9]. Given that such evaluations influence profound outcomes, understanding how people evaluate others' faces can provide insight into a potentially important route through which social stereotypes impact behavior[10,11].

Over the last decade, Oosterhof and Todorov's valence-dominance model[12] has emerged as the most prominent account of how we evaluate faces on social dimensions[5]. Oosterhof and Todorov identified 13 different traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, unhappiness, intelligence, meanness, responsibility, sociability, trustworthiness, and weirdness) that perceivers spontaneously use to evaluate faces when forming trait impressions[12]. 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[12]. *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[12]. 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[4,5].

Successful replications of the valence-dominance model have only been conducted in Western samples[13,14]. This focus on the West is consistent with research on human behavior more broadly, which typically draws general assumptions from analyses of Western participants' responses[15]. Kline et al. recently termed this problematic practice the *Western centrality assumption* and argued that regional variation, rather than universality, is likely the default for human behavior[16].

Consistent with Kline et al's notion that human behavior is best characterized by regional variation, two recent studies of social evaluation of faces by Chinese participants indicate different factors underlie their impressions[17,18]. Both studies reported that Chinese participants' social evaluations of faces were underpinned by a valence dimension similar to that reported by Oosterhof and Todorov for Western participants, but not by a corresponding dominance dimension. Instead, both studies reported a second dimension, referred to as *capability*, which was best characterized by rated intelligence. Furthermore, the ethnicity of the faces rated only subtly affected perceptions[17]. Research into potential cultural differences in the effects of experimentally manipulated facial characteristics on social perceptions has also found little evidence that cultural differences in social perceptions of faces depend on the ethnicity of the faces presented[19-21]. Collectively, these results suggest that the Western centrality assumption may be an important barrier to understanding how people evaluate faces on social dimensions. Crucially, these studies also suggest that the valence-dominance model is not necessarily a universal account of social evaluations of faces and warrants further investigation in the broadest set of samples possible.

Although the studies described above demonstrate that the valence-dominance model is not perfectly universal, to which specific world regions it does and does not apply are open and important questions. Demonstrating differences between British and Chinese raters is evidence against the universality of the valence-dominance model, but it does not adequately address these questions. Social perception in China may be unique in not fitting the valence-dominance model because of the atypically high general importance placed on status-related traits, such as capability, during social interactions in China[22,23]. Indeed, Tan et al. demonstrated face-processing differences between Chinese participants living in mainland China and Chinese participants living in nearby countries, such as Malaysia[24]. Insights regarding the unique formation of social perceptions in other cultures and world regions are lacking. Only a large-scale study investigating social perceptions in many different world regions can provide such insights.

To establish the world regions to which the valence-dominance model applies, we will replicate Oosterhof and Todorov's methodology[12] in a wide range of world regions (Africa, Asia, Australia and New Zealand, Central America and Mexico, Eastern Europe, the Middle East, the USA and Canada, Scandinavia, South America, the UK, and Western Europe; see Table 1). Our study will be the most comprehensive test of social evaluations of faces to date, including more than 9,000 participants. Participating research groups were recruited via the Psychological Science Accelerator project[25-27]. Previous studies compared two cultures to demonstrate regional differences[17,18]. By contrast, the scale and scope of our study will allow us to generate the most comprehensive picture of the world regions to which the valence-dominance model does and does not apply.

We will test two specific competing predictions.

Prediction 1. The valence-dominance model will apply to all world regions.

Prediction 2. The valence-dominance model will apply in Western-world regions, but not other world regions.

Table 1

World Regions, Countries, and Localities of Planned Data Collection

World region	Countries and Localities
Africa	Kenya, Nigeria, South Africa
Asia	China, India, Malaysia, Taiwan, Thailand
Australia and New Zealand	Australia, New Zealand
Central America and Mexico	Ecuador, El Salvador, Mexico
Eastern Europe	Hungary, Lithuania, Poland, Russia, Serbia, Slovakia
The Middle East	Iran, Israel, Turkey
The USA and Canada	Canada, the USA
Scandinavia	Denmark, Finland, Norway, Sweden
South America	Argentina, Brazil, Chile, Colombia
The UK	England, Scotland, Wales
Western Europe	Austria, Belgium, France, Germany, Greece, Italy, the Netherlands, Portugal, Spain, Switzerland

Note. We collected data from a minimum of 350 raters per world region based on the simulations described in the Methods section below.

Methods

Ethics

Each research group had approval from their local Ethics Committee or IRB to conduct the study, had explicitly indicated that their institution did not require approval for the researchers to conduct this type of face-rating task, or had explicitly indicated that the current study was covered by a preexisting approval. Although the specifics of the consent procedure differed across research groups, all participants provided informed consent. All data was stored centrally on University of Glasgow servers.

Procedure

Oosterhof and Todorov derived their valence-dominance model from a principal components analysis of ratings (by US raters) of 66 faces for 13 different traits (aggressiveness, attractiveness, caringness, confidence, dominance, emotional stability, intelligence, meanness, responsibility, sociability, trustworthiness, unhappiness, and weirdness)[12]. Using the criteria of the number of components with eigenvalues greater than 1.0, this analysis produced two principal components. The first component explained 63% of the variance in trait ratings, strongly correlated with rated trustworthiness (r = .94), and weakly correlated with rated dominance (r = -.24). The second component explained 18% of the variance in trait ratings, strongly correlated with rated dominance (r = .93), and weakly correlated with rated trustworthiness (*r* = -.06). We replicated Oosterhof and Todorov's method[12] and primary analysis in each world region we examined.

Stimuli in our study came from an open-access, full-color, face image set[28] consisting of 60 men and 60 women taken under standardized photographic conditions (M_{age} = 26.4 years, *SD* = 3.6 years, Range = 18 to 35 years). These 120 images consisted of 30 Black (15 male, 15 female), 30 White (15 male, 15 female), 30 Asian (15 male, 15 female), and 30 Latin faces (15 male, 15 female). As in Oosterhof and Todorov's study[12], the individuals photographed posed looking directly at the camera with a neutral expression, and all of background, lighting, and clothing (here, a grey t-shirt) were constant across images.

In our study, adult raters were randomly assigned to rate the 13 adjectives tested by Oosterhof and Todorov using scales ranging from 1 (*Not at all*) to 9 (*Very*) for all 120 faces in a fully randomized order at their own pace. Because all researchers collected data through an identical interface (except for differences in instruction language), data collection protocols were highly standardized across labs. Each participant completed the block of 120 face-rating trials twice so that we could report test-retest reliabilities of ratings; ratings from the first and second blocks were averaged for all analyses (see CODE 1.5.5 in the Supplemental Materials).

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

Raters completed the task in a language appropriate for their country (see below). To mitigate potential problems with translating single-word labels, dictionary definitions for each of the 13 traits were provided. Twelve of these dictionary definitions had previously been used to test for effects of social impressions on the memorability of face photographs[19]. Dominance (not included in that study) was defined as "strong, important."

Participants

Simulations determined that we should obtain at least 25 different raters for each of the 13 traits in every region (see https://osf.io/x7fus/ for code and data). We focused on ratings of attractiveness and intelligence for the simulations because they showed the highest and lowest agreement among the traits analyzed by Oosterhof and Todorov, respectively. First, we sampled from a population of 2,513 raters, each of whom had rated the attractiveness of 102 faces; these simulations showed that more than 99% of 1,000 random samples of 25 raters produced good or excellent interrater reliability coefficients (Cronbach's αs >.80). We then repeated these simulations sampling from a population of 37 raters, each of whom rated the intelligence of 100 faces, showing that 93% of 1,000 random samples of 25 raters produced good or excellent interrater reliability coefficients (Cronbach's αs >.80). Thus, averages of ratings from 25 or more raters will produce reliable dependent variables in our analyses; we plan to test at least 9,000 raters in total.

In addition to rating the faces for the 13 traits examined by Oosterhof and Todorov, 25 participants in each region were randomly assigned to rate the targets' age in light of Sutherland et al.'s results showing that a youth/attractiveness dimension emerged from analyses of a sample of faces with a very diverse age range[30]. Age ratings were not included in analyses relating to replications of Oosterhof and Todorov's valence-dominance model.

Analysis Plan

The code used for our analyses is included in the Supplemental Materials and publicly available from the Open Science Framework (https://osf.io/87rbg/).

Ratings from each world region were analyzed separately and anonymous raw data is published on the Open Science Framework. Our main analyses directly replicated the principal component analysis reported by Oosterhof and Todorov to test their theoretical model in each region sampled (CODE 2.1). First, we calculated the average rating for each face separately for each of the 13 traits (CODE 2.1.2). We then subjected these mean ratings to principal component analysis with orthogonal components and no rotation, as Oosterhof and Todorov did (CODE 2.1.3). Using the criteria they reported, we retained and interpreted components with eigenvalues greater than 1.0 (CODE 2.1.3.1).

Criteria for replicating Oosterhof and Todorov's valence-dominance model. We used multiple sources of evidence to judge whether Oosterhof and Todorov's valence-dominance model replicated in a given world region. First, we examined the solution from the principal components analysis conducted in each region and determined if Oosterhof and Todorov's primary pattern replicated according to three criteria: (i) the first two components had eigenvalues greater than 1.0, (ii) the first component (i.e., the one explaining more of the variance in ratings) correlated strongly with trustworthiness ($\lambda > .7$) and weakly with dominance ($\lambda < .5$), and (iii) the second component (i.e., the one explaining less of the variance in ratings) correlated strongly with dominance ($\lambda > .7$) and weakly with trustworthiness ($\lambda < .5$). If the solution in a world region met all three of these criteria, we concluded that the primary pattern of the model replicated in that region (CODE 2.1.3.3).

In addition to reporting whether the primary pattern was replicated in each region, we also reported Tucker's coefficient of congruence[31,32]. The congruence coefficient, ϕ, ranges from -1 to 1 and quantifies the similarity between two vectors of loadings[33]. It is:

$$\phi(x, y) = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}}$$

where x_i and y_i are the loadings of variable i (i = 1, …, n number of indicators in the analysis) onto factors x and y. For the purposes of the current research, we compared the vector of loadings from the first component from Oosterhof and Todorov to the vector of loadings from the first component estimated from each world region. We repeated this analysis for the second component. This produced a standardized measure of component similarity for each component in each world region that was not sensitive to the mean size of the loadings[34]. Further, this coefficient was fitting for the current study because it does not require an a priori specification of a factor structure for each group, as would be needed if we were to compare the factor structures in a multiple-group confirmatory factor analysis. Following previous guidelines[34], we concluded that the components in Oosterhof and Todorov were not similar to those estimated in a given world region if the coefficient was less than .85, were fairly similar if it was between .85 - .94, and were equal if it was greater than .95. (CODE 2.1.4).

Thus, we reported whether the solution had the same primary pattern that Oosterhof and Todorov found and quantified the degree of similarity between each component and the corresponding component from Oosterhof and Todorov's work. This connects to our competing predictions:

Prediction 1 (The valence-dominance model applies to all world regions) was supported if the solution from the principal components analysis conducted in each region satisfied *all* of the criteria described above. Specifically, the primary pattern was replicated and the components had at least a fair degree of similarity as quantified by a ϕ of .85 or greater.

Prediction 2 (The valence-dominance model will applies in Western-world regions, but not other world regions) was supported if the solutions from the principal components analysis conducted in Australia and New Zealand, The USA and Canada, Scandinavia, The UK, and Western Europe, but not Africa, Asia, Central America and Mexico, Eastern Europe, The Middle East, or South America, satisfied the criteria described above.

Exclusions. Data from raters who failed to complete all 120 ratings in the first block of trials or who provided the same rating for 75% or more of the faces was excluded from analyses (CODES 1.5.1,1.5.3, and 1.5.5).

Data-quality checks. Following previous research testing the valence-dominance model[12-14], data quality was checked by separately calculating the interrater agreement (indicated by Cronbach's α and test-retest reliability) for each trait in every world region (CODE 2.1.1). A trait was only included in the analysis for that region if the coefficient exceeded .70. Test-retest reliability of traits was not used to exclude traits from analysis.

Power analysis. Simulations showed we had more than 95% power to detect the key effect of interest (i.e., two components meeting the criteria for replicating Oosterhof and Todorov's work, as described above). We used the open data from Morrison et al.'s replication[13] of Oosterhof and Todorov's research to generate a variance-covariance matrix representative of typical interrelationships among the 13 traits tested in our study. We then generated 1,000 samples of 120 faces from these distributions and ran our planned principal components analysis (which is identical to that reported by Oosterhof & Todorov) on each sample (see https://osf.io/87rbg/ for code and data). Results of >99% of these analyses matched our criteria for replicating Oosterhof and Todorov's findings. Thus, 120 faces gave us more than 95% power to replicate Oosterhof and Todorov's results.

Robustness analyses. Oosterhof and Todorov extracted and interpreted components with an eigenvalue greater than 1.0 using an unrotated principal components analysis. As described above, we directly replicated their method in our main analyses but 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 was the case in the current study[35,36]. Second, the extraction rule of eigenvalues greater than 1.0

has been criticized for not indicating the optimal number of components, as well as for producing unreliable components[37,38].

To address these limitations, we repeated our main analyses using exploratory factor analysis with an oblimin rotation as the model and a parallel analysis to determine the number of factors to extract. We also recalculated the congruence coefficient described above for these exploratory factor analysis results (CODE 2.2.1).

We used parallel analysis to determine the number of factors to extract because it has been described as yielding the optimal number of components (or factors) across the largest array of scenarios[35,39,40] (CODE 2.2.1). In a parallel analysis, random data matrices are generated such that they have the same number of cases and variables as the real data. The mean eigenvalue from the components of the random data is compared to the eigenvalue for each component from the real data. Components are then retained if their eigenvalues exceed those from the randomly generated data[41].

The purpose of these additional analyses was twofold. First, to address potential methodological limitations in the original study and, second, to ensure that the results of our replication of Oosterhof and Todorov's study are robust to the implementation of those more rigorous analytic techniques. The same criteria for replicating Oosterhof and Todorov's model described above was applied to this analysis (CODE 2.2.1.3).

Results and Discussion

Analyzed data set. Following the planned data exclusions (see supplemental materials for a break down of these exclusions, CODE 1.5), the analyzed data set is summarized in Table 2.

Table 2

Number of participants per region and Cronbach's alphas following data quality checks and exclusions

	aggressive	attractive	caring	confident	dominant	emostable	intelligent	mean	responsible	sociable	trustworthy	unhappy	weird
Africa	α = 0.81 n = 45	α = 0.87 n = 38	α = 0.86 n = 44	α = 0.81 n = 31	α = 0.79 n = 38	α = 0.78 n = 38	α = 0.76 n = 37	α = 0.89 n = 51	α = 0.81 n = 36	α = 0.82 n = 34	α = 0.87 n = 49	α = 0.80 n = 43	α = 0.89 n = 37
Asia	α = 0.93 n = 59	α = 0.96 n = 52	α = 0.95 n = 73	α = 0.96 n = 72	α = 0.92 n = 55	α = 0.91 n = 55	α = 0.93 n = 64	α = 0.91 n = 51	α = 0.93 n = 63	α = 0.95 n = 65	α = 0.93 n = 61	α = 0.94 n = 61	α = 0.94 n = 49
Australia & New Zealand	α = 0.96 n = 77	α = 0.98 n = 88	α = 0.96 n = 90	α = 0.97 n = 93	α = 0.94 n = 66	α = 0.96 n = 88	α = 0.95 n = 81	α = 0.95 n = 71	α = 0.94 n = 68	α = 0.97 n = 95	α = 0.95 n = 72	α = 0.95 n = 85	α = 0.96 n = 70
Central America & Mexico	α = 0.86 n = 35	α = 0.95 n = 34	α = 0.84 n = 35	α = 0.90 n = 36	α = 0.88 n = 44	α = 0.85 n = 27	α = 0.83 n = 41	α = 0.86 n = 32	α = 0.84 n = 31	α = 0.90 n = 37	α = 0.87 n = 34	α = 0.83 n = 34	α = 0.83 n = 23
Eastern Europe	α = 0.94 n = 59	α = 0.97 n = 58	α = 0.93 n = 56	α = 0.95 n = 60	α = 0.95 n = 74	α = 0.92 n = 56	α = 0.94 n = 64	α = 0.94 n = 68	α = 0.95 n = 65	α = 0.95 n = 68	α = 0.94 n = 54	α = 0.96 n = 74	α = 0.96 n = 53
Middle East	α = 0.89 n = 27	α = 0.93 n = 26	α = 0.92 n = 34	α = 0.94 n = 34	α = 0.89 n = 30	α = 0.91 n = 28	α = 0.87 n = 38	α = 0.87 n = 27	α = 0.83 n = 28	α = 0.93 n = 32	α = 0.88 n = 37	α = 0.93 n = 49	α = 0.88 n = 24
Scandinavia	α = 0.95 n = 48	α = 0.97 n = 44	α = 0.95 n = 46	α = 0.96 n = 56	α = 0.94 n = 49	α = 0.95 n = 67	α = 0.96 n = 54	α = 0.91 n = 36	α = 0.92 n = 37	α = 0.97 n = 64	α = 0.95 n = 58	α = 0.95 n = 55	α = 0.95 n = 39
South America	α = 0.95 n = 88	α = 0.98 n = 99	α = 0.94 n = 101	α = 0.97 n = 104	α = 0.95 n = 110	α = 0.95 n = 96	α = 0.94 n = 102	α = 0.94 n = 86	α = 0.93 n = 108	α = 0.97 n = 101	α = 0.95 n = 100	α = 0.96 n = 80	α = 0.97 n = 108
UK	α = 0.88 n = 16	α = 0.95 n = 22	α = 0.94 n = 34	α = 0.93 n = 30	α = 0.89 n = 34	α = 0.90 n = 30	α = 0.91 n = 34	α = 0.87 n = 27	α = 0.89 n = 37	α = 0.93 n = 28	α = 0.92 n = 27	α = 0.94 n = 24	α = 0.90 n = 18
USA & Canada	α = 0.98 n = 248	α = 0.99 n = 224	α = 0.99 n = 257	α = 0.99 n = 303	α = 0.98 n = 246	α = 0.99 n = 270	α = 0.98 n = 239	α = 0.98 n = 270	α = 0.98 n = 269	α = 0.99 n = 246	α = 0.98 n = 263	α = 0.98 n = 252	α = 0.99 n = 226
Western Europe	α = 0.98 n = 152	α = 0.99 n = 147	α = 0.98 n = 136	α = 0.98 n = 156	α = 0.97 n = 150	α = 0.98 n = 141	α = 0.97 n = 141	α = 0.97 n = 120	α = 0.98 n = 138	α = 0.99 n = 188	α = 0.98 n = 141	α = 0.98 n = 140	α = 0.98 n = 113

Main analysis (principal components analysis, PCA, CODE 2.1). Oosterhof and Todorov reported the results of a PCA with orthogonal components, no rotation, and retaining components with eigenvalues > 1. Using an identical analysis, we extracted the same number of components in two world regions: Africa and South America. In the other world regions we extracted three components, following the eigenvalues > 1 rule. In the world regions where a third component was extracted the trait ratings of "unhappy" and "weird' tended to have the highest loadings on that component. We are cautious to interpret or describe this component with any authority because it varied across world regions and explained only a small proportion of additional variance.

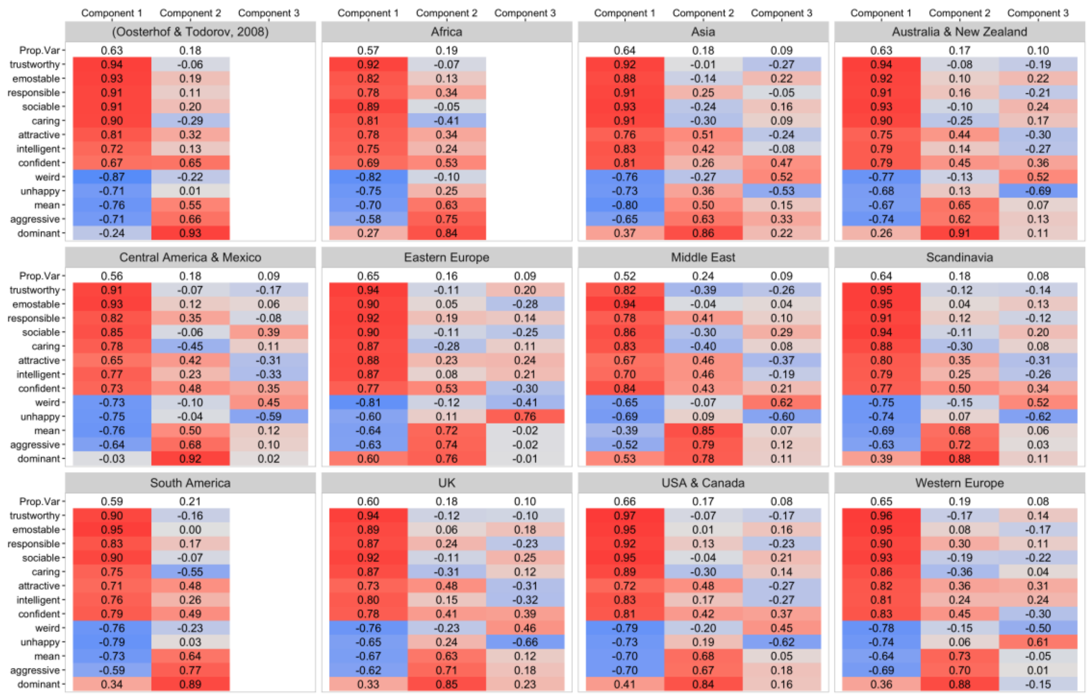

Figure 1. Principal component analysis (PCA) loading matrices for each region. Positive loadings are shaded red and negative loadings shaded blue; darker colors correspond to stronger loadings. The proportion of variance explained by each component is included at the top of each table.

The primary pattern Oosterhof and Todorov reported (a first component that was highly correlated with rated trustworthiness, but not rated dominance, and a second component that was highly correlated with rated dominance, but not rated trustworthiness) was present in all world regions except for Eastern Europe and the Middle East. In those latter two regions, both dominance and trustworthiness ratings were too highly correlated with the first factor. Figure 1 shows the full loading matrices for each region and Table 3 shows how these relate to our replication criteria.

Table 3

Replication criteria for the principal component analysis (PCA) for each region

	Component 1		**Component 2**		
Region	**Dominant**	**Trustworthy**	**Dominant**	**Trustworthy**	**Replicated**
(Oosterhof & Todorov, 2008)	-0.244	0.941	0.929	-0.060	Yes
Africa	0.271	0.924	0.843	-0.065	Yes
Asia	0.370	0.922	0.863	-0.006	Yes
Australia & New Zealand	0.257	0.943	0.907	-0.076	Yes
Central America & Mexico	-0.030	0.913	0.923	-0.066	Yes
Eastern Europe	0.599	0.938	0.755	-0.113	No
Middle East	0.528	0.816	0.778	-0.394	No
Scandinavia	0.392	0.953	0.881	-0.121	Yes
South America	0.343	0.899	0.894	-0.155	Yes
UK	0.331	0.944	0.851	-0.121	Yes
USA & Canada	0.406	0.966	0.841	-0.073	Yes
Western Europe	0.357	0.957	0.875	-0.166	Yes

Note: Oosterhof and Todorov's valence-dominance model was judged to have been replicated in a given world region if the first component had a loading < .5 with dominance and > .7 with trustworthiness, and the second component had a loading > .7 with dominance and < .5 with trustworthiness.

Tucker's coefficient of congruence, ϕ, indicated that the first component was congruent with the first component in Oosterhof and Todorov's original study for all world regions (i.e., ϕ > .95). The second component was also congruent with the second component reported by Oosterhof and Todorov in all of the world regions (i.e., all ϕ > .85), except Asia (ϕ = .848). Table 4 summarizes these results.

Table 4

Factor congruence for each region's principal component analysis (PCA)

	Component 1		Component 2	
Region	**Loading**	**Congruence**	**Loading**	**Congruence**
Africa	0.980	equal	0.947	fairly similar
Asia	0.974	equal	0.843	not similar
Australia & New Zealand	0.982	equal	0.959	equal
Central America & Mexico	0.993	equal	0.953	equal
Eastern Europe	0.953	equal	0.948	fairly similar
Middle East	0.944	fairly similar	0.853	fairly similar
Scandinavia	0.973	equal	0.960	equal
South America	0.973	equal	0.948	fairly similar
UK	0.976	equal	0.938	fairly similar
USA & Canada	0.972	equal	0.952	equal
Western Europe	0.975	equal	0.936	fairly similar

Together, these results suggest the valence-dominance model generalizes across world regions when using an identical analysis to Oosterhof and Todorov's original study. Thus, the results of our PCA support prediction one (that the valence-dominance model will apply to all world regions), but not prediction two (that the valence-dominance model will apply in Western world regions, but not other world regions). However, we note here that generalization to Eastern Europe and Middle East was poorer than for the other regions.

Robustness analyses[*] (Exploratory Factor Analysis, CODE 2.2). Following our analysis plan, we conducted additional robustness analyses that directly addressed criticisms of the type of statistical analyses used by

[*] Note for reviewers: During the robustness analysis, using our registered code, an error ("The estimated weights for the factor scores are probably incorrect. Try a different factor extraction method.") appeared in the output for one region (USA and Canada). Inspection of the weights did not show any aberrant or out of range values. By using a parallel analysis extraction method, we did not even use an extraction technique as executed by the factor analysis features of the psych package in R. It is possible this error is not even relevant. To fully understand if this error is meaningful and impactful to the trustworthiness of the results, we will need to conduct some exploratory analyses and consider if our results are sensitive across shifting around the extraction and estimation options and compare them to results from other software packages. Given we did not register such analyses, we would like guidance from the reviewers and editor on how to report on these and strategies for moving forward.

Oosterhof and Todorov (see[42] for a discussion of these criticisms). These analyses employed EFA with an oblimin rotation as the model and used parallel analysis to identify the number of factors to extract. We conducted this analysis on Oosterhof and Todorov's original data and found a similar result to their PCA solution. With the EFA, all other regions showed more than two factors. Full EFA loading matrices for each region and Oosterhof and Todorov's original data are shown in Figure 2.

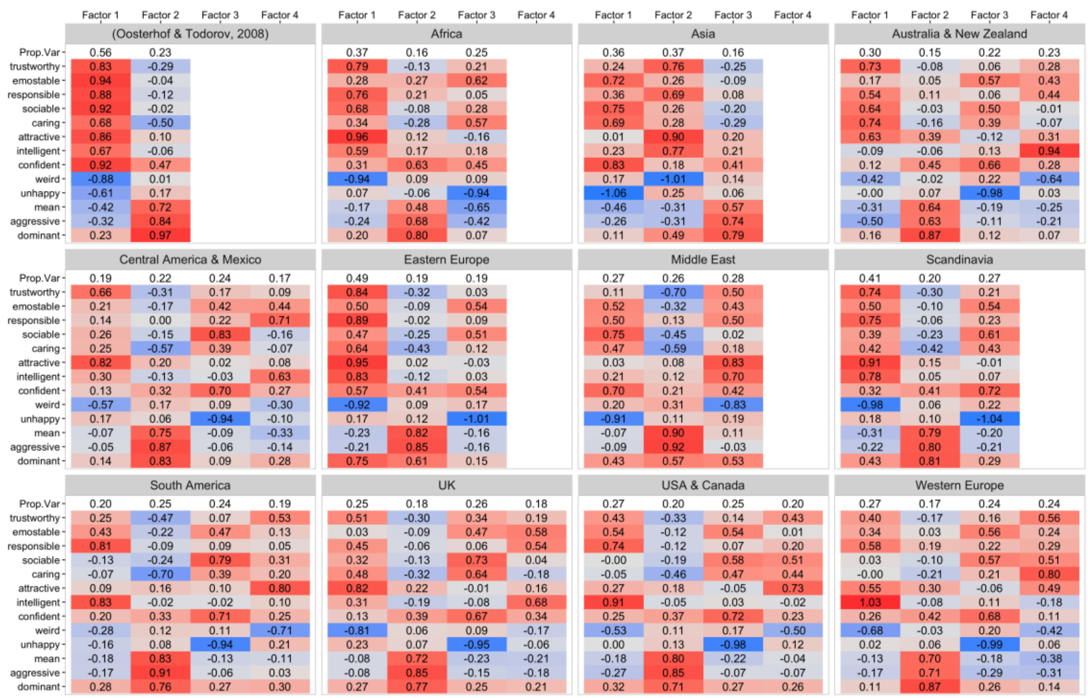

Figure 2

Exploratory factor analysis (EFA) loading matrices for each region. Positive loadings are shaded red and negative loadings shaded blue; darker colours correspond to stronger loadings. The proportion of variance explained by each factor is included at the top of each table.

In contrast to our PCA, the results of our robustness analyses showed little evidence that the valence-dominance model generalizes across world regions. A summary of the results for our replication criteria is given in Table 5. These results showed that our replication criteria were met for Australia and New Zealand, Africa, and Scandinavia, but for none of the other world regions.

Table 5

Replication criteria for the exploratory factor analysis (EFA) for each region

	Factor 1		**Factor 2**		
Region	**Dominant**	**Trustworthy**	**Dominant**	**Trustworthy**	**Replicated**
(Oosterhof & Todorov, 2008)	0.228	0.826	0.970	-0.288	Yes
Africa	0.200	0.786	0.796	-0.133	Yes
Asia	0.110	0.236	0.487	0.761	No
Australia & New Zealand	0.157	0.730	0.873	-0.078	Yes
Central America & Mexico	0.142	0.662	0.831	-0.311	No
Eastern Europe	0.750	0.843	0.609	-0.322	No
Middle East	0.427	0.112	0.566	-0.699	No
Scandinavia	0.428	0.744	0.806	-0.304	Yes
South America	0.278	0.255	0.757	-0.472	No
UK	0.265	0.510	0.766	-0.299	No
USA & Canada	0.320	0.426	0.711	-0.335	No
Western Europe	0.111	0.398	0.869	-0.172	No

Note: Oosterhof and Todorov's valence-dominance model was judged to have been replicated in a given world region if the first factor had a loading < .5 with dominance and > .7 with trustworthiness, and the second factor had a loading > .7 with dominance and < .5 with trustworthiness.

Tucker's coefficient of congruence, ϕ, indicated that the first factor was congruent with the first factor in Oosterhof and Todorov's original study (i.e., ϕ

> .85) for Africa, Eastern Europe, and Scandinavia. The second factor was congruent with the second factor reported by Oosterhof and Todorov in all of world regions (i.e., all $\phi > .85$), except Asia ($\phi = -.090$). Table 6 summarizes these results.

Table 6

Factor congruence for each region's exploratory factor analysis (EFA)

	Factor 1		Factor 2	
Region	**Loading**	**Congruence**	**Loading**	**Congruence**
Africa	0.894	fairly similar	0.900	fairly similar
Asia	0.765	not similar	-0.090	not similar
Australia & New Zealand	0.810	not similar	0.933	fairly similar
Central America & Mexico	0.777	not similar	0.972	equal
Eastern Europe	0.891	fairly similar	0.957	equal
Middle East	0.736	not similar	0.855	fairly similar
Scandinavia	0.884	fairly similar	0.980	equal
South America	0.682	not similar	0.956	equal
UK	0.774	not similar	0.977	equal
USA & Canada	0.772	not similar	0.975	equal
Western Europe	0.774	not similar	0.949	fairly similar

Thus, the results of our EFA support neither Prediction one (that the valence-dominance model will apply to all world regions) nor Prediction two (that the valence-dominance model will apply to Western-world regions, but not other world regions). There was, however, some evidence that a second factor that was highly correlated with dominance was present in all world regions except Asia.

Conclusions

Our primary analyses, PCAs identical to those reported by Oosterhof and Todorov, suggested that the valence-dominance model of social perception of faces generalizes relatively well across world regions. However, most world regions showed a third component not discussed in the original work. The presence of this third component suggests that a simple valence-dominance model does not fully capture the richness of social perception in many world regions. Further work is needed to interpret this component.

In contrast to the results of our PCAs, an alternative analysis that addressed common criticisms of the type of analysis Oosterhof and Todorov employed showed little evidence that the valence-dominance model is useful for summarizing social perception of faces in different world regions. For example, according to our primary replication criteria, the valence-dominance model replicated in only four world regions. Although some previous research on the generalizability of social perception of faces has focused on investigating generalization across different world regions, our study extends this work by emphasizing the additional importance of generalization across analysis models. We show that conclusions about the extent to which models of social perception generalize across world regions can depend, at least to some extent, on the specific model employed to analyze the data.

A necessary next step for moving forward in person perception research is addressing which analysis model (PCA or EFA) best aligns with theory, so that those models and theories can be revised and expanded appropriately in future research. Crucially, the two models make different assumptions about trait ratings of faces. The PCA model does not assume that a latent factor causes the trait ratings of the faces. The component simply captures an aggregate, maximized to explain variance. Furthermore, in the original valence-dominance model, those components were assumed to be unrelated. By contrast, the theory underlying the EFA model is that a latent factor causes the trait ratings, and any unexplained variance in that rating is measurement error. Additionally, our EFA models allowed for the factors to be correlated.

Theory can guide which model we use to analyze person perception data. A person perception theory that aligns with a PCA model would state that there are no underlying, latent factors that cause a person to rate a face in a particular way. There are, instead, perceptual processes that vary across contexts, those doing the rating, and those being rated, and the differential processes give rise to components that can be used to reduce the data. Because the ratings come from context specific processes (and not a causal, latent factor) the estimated components can vary across contexts, raters, and those being rated. This theory of person perception would move forward with identifying the shared processes across contexts, those rating, and those being rated, to see if there are predictable patterns in how the data are reduced. However, a person perception theory that aligns with an EFA model would state that latent factors (e.g., valence or dominance) cause the trait ratings and, once we account for the correct latent factors, any variability left in the ratings is measurement error. This theory would move forward with identifying and defining those latent factors, confirming their existence, reliability, and generalizability.

Our study is one of several recent studies that have begun to address these key questions[21,43,44] by exploring how the structure of trait ratings vary systematically. This growing body of work catalogues variations in trait ratings by target demographic[21,43, 45], target status[46], target age[47], perceiver knowledge[48], and cultural factors[17,18]. Further, from this growing body of work dynamic theories of person perception and more flexible statistical models for capturing them have been proposed[21,43,44,49].

Our results are consistent with this recent work in that they do not provide strong evidence that there are a few generalizable latent factors that cause the trait ratings across world regions. They do however, suggest a dynamic process of person perception and elucidate the differential patterns of ratings across world regions. We can use these data, representing impressions formed on a global scale, to expand or refine our theories and guide the selection of statistical models to represent those theories.

References

1. Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. Journal of Nonverbal Behavior, 34, 83-110.
2. Ritchie, K. L., Palermo, R., & Rhodes, G. (2017). Forming impressions of facial attractiveness is mandatory. Scientific Reports, 7, 469.
3. Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. Psychological Science, 17, 592-598.
4. Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. Trends in Cognitive Sciences, 18, 566-570.
5. 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.

6. 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.

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

8. 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.

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

10. 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.

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

12. 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.

13. 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.

14. 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.

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

16. 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.

17. 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.

18. Wang, H., Han, C., Hahn, A., Fasolt, V., Morrison, D.,... Jones, B. C. (2019). A data-driven study of Chinese participants' social judgments of Chinese faces. PloS one, 14(1), e0210315.

19. Han, C., Wang, H., Hahn, A. C., Fisher, C. I., Kandrik, M., Fasolt, V., ... & Jones, B. C. (2018). Cultural differences in preferences for facial coloration. Evolution and Human Behavior, 39, 154-159.

20. Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., ... & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. Nature, 394, 884.

21. Xie, S. Y., Flake, J. K., & Hehman, E. (2019). Perceiver and target characteristics contribute to impression formation differently across race and gender. Journal of Personality and Social Psychology, 117, 364.

22. Li, N. P., Valentine, K. A., & Patel, L. (2011). Mate preferences in the US and Singapore: A cross-cultural test of the mate preference priority model. Personality and Individual Differences, 50, 291-294.

23. Ting-Toomey, S. (1994). Face and facework: An introduction. In S. Ting-Toomey (Ed.),The Challenge of Facework: Cross-Cultural and Interpersonal Issues (pp. 1-14). Albany, NY: State University of New York Press.

24. Tan, C. B. Y., Stephen, I. D., Whitehead, R., Sheppard, E. (2012). You look familiar: How Malaysian Chinese recognize faces. PLOS ONE 7, e29714.

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

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

27. Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., ... & Castille, C. M. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. Advances in Methods and Practices in Psychological Science, 1(4), 501-515.

28. Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The chicago face database: A free stimulus set of faces and norming data. Behavior Research Methods, 47, 1122-1135.

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

30. Sutherland, C. A., 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, 105-118.

31. Burt, C. (1948). The factorial study of temperament traits. British Journal of Psychology, Statistical Section, 1, 178-203.

32. Tucker, L. R. (1951). A method for synthesis of factor analysis studies (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.

33. Davenport, E. C., Jr., (1990). Significance testing of congruence coefficients: A good idea? Educational and Psychological Measurement, 50, 289-296.

34. Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. Methodology, 2, 57-64.

35. 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.

36. 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.

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

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

39. 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.

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

41. Courtney, M. G. R. (2013). Determining the number of factors to retain in EFA: Using the SPSS R-Menu v2.0 to make more judicious estimations. Practical Assessment, Research & Evaluation, 18, 1-14.

42. Widaman, K. F. (2018). On common factor and principal component representations of data: Implications for theory and for confirmatory replications. Structural Equation Modeling, 25, 829-847.

43. Hehman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. Journal of Personality and Social Psychology, 113, 513.

44. Sutherland, C. A., Rhodes, G., Burton, N. S., & Young, A. W. (in press). Do facial first impressions reflect a shared social reality? British Journal of Psychology.

45. Oh, D., Dotsch, R., Porter, J., & Todorov, A. (in press). Gender biases in impressions from faces: Empirical studies and computational models. *Journal of Experimental Psychology: General*.

46. Oh, D., Shafir, E., & Todorov, A. (preprint). Economic Status Cues from Clothes Affect Perceived Competence from Faces. PsyArxiv.

47. Collova, J. R., Sutherland, C. A., & Rhodes, G. (2019). Testing the functional basis of first impressions: Dimensions for children's faces are not the same as for adults' faces. *Journal of Personality and Social Psychology*.

48. Stolier, R. M., Hehman, E., Keller, M. D., Walker, M., & Freeman, J. B. (2018). The conceptual structure of face impressions. *Proceedings of the National Academy of Sciences*, *115*(37), 9210-9215.

49. Stolier, R. M., Hehman, E., & Freeman, J. B. (2018). A dynamic structure of social trait space. *Trends in Cognitive Sciences*, *22*(3), 197-200.