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

Benedict C. Jones[1]‡* (psysciacc.001@gmail.com), 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], Enrique Turiegano[8], Michael Gilead[9], Almog Simchon[9], S. Adil Saribay[10], Nicholas C. Owsley[11], Chaning Jang[11], Georgina Mburu[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], Amanda C. Hahn[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], Nandor Hajdu[23, 24],Balazs Aczel[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],, 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], Trinity Leonis[64], Jordan W. Suchow[65], Jonas K. Olofsson[66], Teodor Jernsäther[66], Ai-Suan Lee[67], Jennifer L. Beaudry[68], Taylor D. Gogan[68], Julian A. Oldmeadow[68], Benjamin Balas[69], Laura M. Stevens[70], Melissa F. Colloff[70], Heather D. Flowe[70],Sami Gülgöz[71],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], Lison Neyroud[80], Touhami Badidi[81], Nicolas Van der Linden[82], Chrystalle B. Y. Tan[83], Vanja Kovic[84], Waldir Sampaio[85], Paulo Ferreira[86], Diana Santos[86], Debora I. Burin[87], Gwendolyn Gardiner[88], John Protzko[89], Christoph Schild[90], Karolina A. Ścigała[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], Julia Stern[98], 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], Michelangelo Vianello[110], Ferenc Kocsor[111], Luca Kozma[111], Ádám Putz[111], Patrizio Tressoldi[112], 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], , Barnaby J. W. Dixson[120], Gianni Ribeiro[120], 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 Frías-Armenta[125], Javad Hatami[126], Arash Monajem[126], MohammadHasan Sharifian[126], Brooke Frohlich[127], Hause Lin[128], Michael Inzlicht[128], Ravin Alaei[128], Nicholas O. Rule[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], Martha Lucia Borras-Guevara[144], Sun Jun Cai[145], Dong Tiantian[145], Alexander F. Danvers[146], David R. Feinberg[147], Marie M Armstrong[147], Eva Gilboa-Schechtman[148], Randy J. McCarthy[149], Jose Antonio Muñoz-reyes[150], Pablo Polo[150], Victor K.M. Shiramazu[151], Wen-Jing Yan[152], Patrick S. Forscher[80], Christopher R. Chartier[6], Nicholas A. Coles[153]

[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]Leibniz Institute for Resilience Research, 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 in Rabat, 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, Facultad de Ciencias de la Vida, Ecuador. [61]Saint Joseph's University, Department of Food Marketing, USA. [62]Sheffield Hallam University, Centre for Behavioural Science and Applied Psychology, UK. [63]Southern Illinois University, School of Psychological and Behavioral Sciences, 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, LIP/PC2S, 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]Universidade Federal de São Carlos, Brazil, [86]Universidade Federal da Grande Dourados, Brazil, [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 Brasília, Department of Social and Work 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, Faculty of Psychology, Department of Cognition, Emotion, and Methods in Psychology, Austria. [130]University of Vienna, Faculty of Psychology, 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]McMaster University, Department of Psychology, Neuroscience, & Behaviour, Canada. [148]Bar-Ilan University, Israel. [149]Northern Illinois University. [150]Playa Ancha University of Educational Sciences, Chile. [151]Federal University of Rio Grande do Norte, Brazil. [152]Wenzhou University, China. [153]Harvard University, Harvard Kennedy School, USA.

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 11,570 participants. When we used Oosterhof and Todorov's original analysis strategy, the valence-dominance model generalized across regions. When we used an alternative methodology to allow for correlated dimensions we observed much less generalization. Collectively, these results suggest that, while the valence-dominance model generalizes very well across regions when dimensions are forced to be orthogonal, regional differences are revealed when we use different extraction methods, correlate and rotate the dimension reduction solution.

Introduction

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 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 is the most comprehensive test of social evaluations of faces to date, including more than 11,000 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 allows us to generate the most comprehensive picture of the world regions to which the valence-dominance model does and does not apply.

We test two specific competing predictions:

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

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

Results

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

Main analysis (principal component 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. We conducted an identical analysis and report (1) the number of components extracted per the registered criteria, (2) if the 1st and 2nd components had the same primary pattern as Oosterhof and Todorov, and (3) the similarity of the 1st and 2nd factors as quantified with a congruence coefficient.

We extracted the same number of components (2) as Oosterhof and Todorov in two world regions, Africa and South America, and a different number of components (3) in the other world regions (see Figure 1). 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, but those ratings also crossloaded on the first component. We hesitate to interpret or describe this component with any authority because it varied across world regions, consisted of crossloaded traits, and explained only a small proportion of additional variance.

The primary pattern reported by Oosterhof and Todorov (a first component that strongly correlated with rated trustworthiness, but not with rated dominance, and a second component that strongly correlated with rated dominance, but not with rated trustworthiness) was present in all world regions except Eastern Europe, where dominance was correlated with the first component more strongly than our registered criterion (that dominance would correlate weakly with the first component < .5). Figure 1 shows the full loading matrices for each region and Table 3 shows how these relate to our registered criteria.

We report Tucker's coefficient of congruence, ϕ, which quantifies the loading similarity of Oosterhof and Todorov's reported component to the corresponding component we extracted. However, it is important to interpret ϕ with caution when the number of components differs across the solutions being compared. When comparing loadings across solutions, an assumption is that the configuration of the traits to components is the same (i.e., configural invariance). To the extent that the structure of the loading matrices differs across solutions, the comparability of the loadings is compromised (i.e., loadings estimated from different dimensional spaces are not on the same scale). For world regions that did not have the same configuration of traits to components (different number of components extracted, different primary pattern observed), ϕ is uninterpretable. This is because the differences in configuration across the two solutions are conflated with the loading differences.

Our analyses indicated that the first component was equal to (ϕ > .95) the first component in Oosterhof and Todorov's original study for all world regions. The second component was equal to (ϕ > .95) or fairly similar to (ϕ > .85) the second component reported by Oosterhof and Todorov in all of the world regions except Asia (ϕ = .848). Table 4 summarizes these results.

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 1 (that the valence-dominance model will apply to all world regions), but not Prediction 2 (that the valence-dominance model will apply in Western world regions, but not other world regions). However, we note here that in most world regions, we extracted a 3rd component not extracted in the original study, that Eastern Europe did not demonstrate the same primary pattern, and that ϕ should be interpreted with caution for all world regions except Africa and South America.

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 Oosterhof and Todorov (see[42] for a discussion of these criticisms). These robustness analyses employed EFA with an oblimin rotation as the model and used parallel analysis to identify the number of factors to extract. The goal of an EFA with an oblimin rotation is to simplify the loading matrix and yield interpretable factors.

We conducted this analysis on Oosterhof and Todorov's original data and found a similar result to their PCA solution: two factors extracted, with Factor 1 characterized by a high loading for trustworthiness and Factor 2 characterized by a high loading for dominance. However, for all other world regions, we extracted more than two factors using parallel analysis. Full EFA loading matrices for each region and Oosterhof and Todorov's original data are shown in Figure 2. The four-factor solution for the USA and Canada did not converge. We did not register a contingency for nonconvergence, but because parallel analysis can lead to over extraction, we reran the EFA with one less than the number of suggested factors. The model converged when estimating three factors.

In contrast to the PCA, the results of our robustness analyses showed less evidence that the valence-dominance model generalizes across world regions. For example, we extracted a different number of factors than the original solution for all world regions. A summary of the results for our replication criteria is given in Table 5.

Because the number of factors differed from the original solution in all world regions and the loading matrices were differentially rotated from the original solution, it is not valid to compare the differences in the loadings from the original solution to those observed in the world regions reported here, as we had initially planned. Loadings quantify the relationship of traits to a factor. To compare loadings across samples, we must first determine whether we extracted the same factor in each sample (i.e., satisfied the assumption of configural invariance). Our registered analyses included the calculation of Tucker's coefficient of congruence, ϕ, in order to compare the 1st factor from the original study to the 1st factor we extracted in a given world region, and to compare the 2nd factor from the original study to the 2nd factor extracted in a given world region. However, because we extracted a different number of factors from the original solution in all world regions, it is not valid to compare the loadings across these different factors or quantify their differences using ϕ.

The congruence coefficient is only appropriate to report when we can ensure the factors are comparable across samples. That the number of factors extracted did not replicate the original pattern and that the EFAs were rotated differently across world regions negates the comparability of the loadings. Consistent with our registered analysis code, we reported ϕ for the 1st factor from Oosterhof and Todorov to the factor with the most explained variance in a world region, and ϕ for the 2nd factor from Oosterhof and Todorov to the factor with the 2nd most explained variance in a world region only in the supplemental materials. However, we stress that these coefficients are quantifying loadings that link to different factors from different dimensional spaces and are not necessarily comparable.

In summary, the results of our EFA support neither Prediction 1 (that the valence-dominance model will apply to all world regions) nor Prediction 2 (that the valence-dominance model will apply to Western-world regions, but not other world regions).

Discussion

Our primary analyses, PCAs identical to those reported by Oosterhof and Todorov (2008), suggested that the valence-dominance model of social perception of faces generalizes well across world regions. Although most world regions showed a third component not discussed in the original work, this third component is actually similar to the third component in Oosterhof and Todorov's original study. In Oosterhof and Todorov's original study, they did not interpret the third component because its eigenvalue was below 1, whereas in our analyses the eigenvalues of the third components in most of the regions are just above 1. Nonetheless, the third component in each region has a factor congruence between 0.77 and 0.90 with the third component for Oosterhof and Todorov's data. However, we emphasize here that many of these dimensions accounted for a relatively small proportion of the variance explained and, thus, may be of limited theoretical importance.

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 much less generalization across world regions. We used modern extraction techniques and EFAs with correlated factor rotations. The correlated rotation methods aim to simplify the loading matrix with the goal of estimating interpretable factors, and in our data, revealed more regional variation. These results suggest that, if the dimensions of face perception are indeed correlated, using analytic techniques that force these dimensions to be uncorrelated may be obscuring important regional differences in the structure of face perceptions.

A necessary next step for moving forward in person-perception research is to address 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 captures linear combinations of the original variables, maximized to explain variance. Furthermore, in the original valence-dominance model, those components were assumed to be orthogonal. 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. This theory of person perception would move forward with identifying the shared processes across contexts, raters, and ratees to see if there are predictable patterns in how the data are reduced.

A person-perception theory that aligns with an EFA model makes different assumptions about the processes that give rise to face ratings. This theory 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. We suggest that more careful and explicit consideration of how theory connects to these approaches, and of which approach is best-suited to different research questions, will benefit the field.

Our study is one of several recent studies that have begun to utilize different statistical models and explore more dynamic theories of trait ratings[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, this growing body of work proposes dynamic theories of person perception and more flexible statistical models for capturing them[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. Given the accumulating evidence for variation in trait ratings, it is important that the connection between the statistical models used to represent theories of person perception are explicit and can accommodate the complexities of the impression formation process.

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 the 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 planned 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. These age-ratings were collected to allow for planned exploratory analyses including rated age, but we did not perform these analyses.

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/). The specific sections of code are cited below as (CODE x.x.x).

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 (⌊ > .7) and weakly with dominance (⌊ < .5), and (iii) the second component (i.e., the one explaining less of the variance in ratings) correlated strongly with dominance (⌊ > .7) and weakly with trustworthiness (⌊ < .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 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 were 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. Cases in which the coefficient does not exceed .70 will be reported and discussed. There were no cases in which the coefficient did not exceed .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 2.2.2).

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 were to be applied to this analysis (CODE 2.2.1.3).

Protocol registration

The Stage 1 protocol for this Registered Report was accepted in principle on 05 November 2018. The protocol, as accepted by the journal, can be found at https://dx.doi.org/10.6084/m9.figshare.7611443.v1.

Data availability

Full data are publicly available at https://osf.io/87rbg/

Code availability

Full analysis code is publicly available at https://osf.io/87rbg/

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. (2020). Economic Status Cues from Clothes Affect Perceived Competence from Faces. Nat Hum Behav 4, 287–293

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, 117(5), 900–924.

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.

Acknowledgements

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); Debora Burin and Natalia Irrazabal received funding from CONICET, Argentina; 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). Krystian 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). Taylor Gogan was supported by an Australian Government Research Training Program (RTP) Scholarship. 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. 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, Domenico Cafaro, Viviana Ciambrone, Francesca Dolce, and Eva Tolomeo (“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); 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 funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

B.C.J., L.M.D., J.K.F., J.P.W., J.B.F., S.Á., H.I., S.M.J.J., and H.L. contributed to Conceptualization.

B.C.J., L.M.D., N.C.A., N.G.B., Y.Q., J.W.L., K.G., G.M.M., J.G.L., J.B.F., P.C., A.P., N.P., S.P., M.M.S., B.P., M.J.B., V.K., J.P., D.S., S.C.W., J.V.V., P.S.F., C.R.C., and N.A.C. contributed to Data Curation.

B.C.J., L.M.D., J.K.F., Y.Q., and J.B.F. contributed to Formal Analysis.

N.C.O., Y.Q., J.W.L., C.C., J.L., O.R.S., E.V., M.V., J.G.L., M.C.P., J.B.F., J.K.O., G.K., H.I., H.D.F., T.J.S.L., E.V., K.A., K.T., M.I., C.L., P.S.F., and C.R.C. contributed to Funding Acquisition.

B.C.J., L.M.D., M.L., J.A., I.L.G.N., N.G.B., S.C.L., F.F., M.L.W., C.P.C., M.A.V., S.A.S., N.C.O., D.P.C., A.W., Y.Q., H.M., P.S., T.R.E., J.B., J.W.L., C.C., A.K., A.K., J.L., O.R.S., E.V., M.V., B.A., P.S., M.A., E.D.M., C.B., C.H., Q.L., L.V., K.B., K.G., I.S., S.S., R.A., C.M., W.V., Z.J., Q.W., G.M.M., I.D.S., J.G.L., M.C.P., J.D.A., E.H., S.Y.X., W.J.C., M.S., J.P.W., E.K., M.P., A.E.B., A.d., I.G., H.W., J.B.F., D.O., V.S., T.E.S., C.A.L., C.L.C., A.K.P., J.B., P.K., I.Z., S.Á., E.B., M.T., K.S., C.D.C., J.W.S., J.K.O., A.L., T.D.G., J.A.O., B.J.W.D., L.M.S., G.R., M.J.B., B.J., D.R., G.K., V.A.F., H.L.U., S.C., G.P., Z.V., D.M.B., H.I., N.V., C.B.Y.T., V.K., M.F.C., H.D.F., D.I.B., G.G., J.P., C.S., K.A.S., E.M.O., D.S., B.S., M.S., G.V.S., T.J.S.L., K.U., E.V., J.S., K.A., C.J.J.v., A.K., S.C.W., J.B., F.R., K.L.R., N.M.M., K.R.B., D.W., A.R.G., M.A., S.M.J.J., K.L.,

T.K.N., C.K.T., J.H.Z., A.D., L.K., M.V., N.I., A.C., S.L., J.L., M.A., P.B., G.B., I.R., V.C., K.P., N.K.S., K.T., C.A.T., A.F., R.M.C.S.H., J.V.V., N.S.C., M.F., J.H., A.M., M.S., B.F., H.L., M.I., C.L., E.P., M.V., J.O., E.M., A.A., A.A.Ö., M.T.C., B.B., M.A.K., C.O., T.G., J.K.M., Y.D., X.Y., S.A., P.S.F., C.R.C., and N.A.C. contributed to Investigation.

B.C.J., L.M.D., J.K.F., S.C.L., L.V., M.S., S.A., A.B., J.E., J.P.W., J.B.F., C.A.L., C.D.C., K.H., B.J., J.W., G.K., H.I., T.B., N.V., H.D.F., J.P., F.M.A.W., S.M.J.J., and H.L. contributed to Methodology.

B.C.J., L.M.D., N.G.B., S.C.L., M.L.W., M.G., A.S., N.C.O., A.W., Y.Q., H.M., R.M.S., J.B., J.W.L., C.C., A.K., A.K., J.L., O.R.S., E.V., M.V., B.A., C.B., C.H., L.V., K.B., K.G., I.S., S.S., I.D.S., M.C.P., S.Y.X., W.J.C., M.S., A.d., I.G., C.K., J.B.F., D.O., C.A.L., J.B., P.K., I.Z., K.S., C.D.C., J.W.S., J.L.B., J.A.O., B.J.W.D., M.J.B., B.J., D.R., G.P., Z.V., E.S., N.V., V.K., M.F.C., H.D.F., J.P., C.S., K.A.S., E.M.O., B.S., M.S., T.J.S.L., K.U., E.V., K.A., C.J.J.v., S.C.W., J.B., F.R., K.L.R., D.W., S.M.J.J., C.K.T., J.H.Z., L.K., S.L., V.C., N.K.S., K.T., R.M.C.S.H., J.V.V., A.M., M.S., B.F., H.L., C.L., E.P., M.V., A.A., A.A.Ö., M.A.K., T.G., X.Y., S.A., P.S.F., C.R.C., and N.A.C. contributed to Project Administration.

B.C.J., L.M.D., M.L., S.C.L., C.P.C., M.A.V., S.A.S., A.W., Y.Q., K.A., S.J., H.M., P.S., N.T., R.M.S., C.C., A.K., J.L., M.V., N.H., C.B., L.V., K.B., K.G., Z.J., G.M.M., I.D.S., J.G.L., S.Y.X., W.J.C., M.S., S.A., A.B., J.E., A.d., C.K., J.B.F., C.A.L., A.K.P., P.K., I.Z., E.B., K.S., C.D.C., J.K.O., J.L.B., B.J.W.D., D.R., W.W.A.S., S.C., G.P., D.M.B., T.B., C.B.Y.T., V.K., H.D.F., G.G., C.S., K.A.S., E.M.O., B.S., M.S., G.V.S., T.J.S.L., K.U., E.V., K.J., K.A., J.B., F.R., N.M.M., S.M.J.J., C.K.T., A.D., F.K., Á.P., P.T., M.V., A.C.,

S.L., J.L., M.A., P.B., V.C., A.F., R.M.C.S.H., J.V.V., N.S.C., M.F., A.M., M.S., H.L., C.L., M.V., E.M., A.A.Ö., M.A.K., C.O., X.Y., S.A., P.S.F., and C.R.C. contributed to Resources.

B.C.J., L.M.D., J.K.F., G.M.M., I.D.S., N.P., B.P., C.D.C., H.D.F., C.S., K.R.B., R.M.C.S.H., C.R.C., and N.A.C. contributed to Software.

B.C.J., L.M.D., J.K.F., M.L., S.C.L., M.L.W., N.C.O., A.W., H.M., J.W.L., C.C., A.K., J.L., O.R.S., E.V., M.V., M.A., E.D.M., C.B., L.V., K.B., I.D.S., M.C.P., E.H., W.J.C., M.S., C.K., J.B.F., C.A.L., P.K., I.Z., K.S., C.D.C., J.W.S., J.K.O., A.L., J.L.B., J.A.O., B.J.W.D., M.J.B., H.I., V.K., M.F.C., H.D.F., J.P., C.S., E.M.O., D.S., B.S., M.S., T.J.S.L., K.U., E.V., K.A., C.J.J.v., D.W., S.M.J.J., A.C., S.L., K.P., N.K.S., K.T., A.F., J.V.V., M.S., M.I., C.L., M.V., A.A., A.A.Ö., M.A.K., S.A., P.S.F., C.R.C., and N.A.C. contributed to Supervision.

B.C.J., L.M.D., J.K.F., C.C., Q.W., S.Y.X., M.S., C.L.C., A.K.P., I.Z., J.W.S., E.S., V.K., H.D.F., J.P., M.S., E.V., C.J.J.v., P.T., J.H., M.V., and M.A.K. contributed to Validation.

B.C.J., L.M.D., J.K.F., H.D.F., M.A.K., and P.S.F. contributed to Visualization.

B.C.J., L.M.D., J.K.F., F.F., Y.Q., C.B., I.G., J.B.F., K.S., B.J.W.D., G.K., H.L.U., H.I., H.D.F., D.I.B., J.P., C.S., D.S., K.L.R., S.M.J.J., A.D., N.K.S., J.O., A.A.Ö., M.A.K., P.S.F., and N.A.C contributed to Writing (original draft).

B.C.J., L.M.D., J.K.F., M.L., J.A., I.L.G.N., S.C.L., F.F., M.L.W., M.A.V., A.S., D.P.C., A.W., Y.Q., K.A., H.M., T.R.E., J.B., J.W.L., C.C., J.L., B.A., N.H., P.S., M.A., E.D.M., C.B., N.L., L.V., K.B., I.S., S.S., Z.J., I.D.S., M.C.P., J.D.A., E.H., S.Y.X., W.J.C., M.S., S.A., A.B., J.E., J.P.W., E.K., M.P., A.E.B., A.d., J.B.F., V.S., T.E.S., C.A.L., C.L.C., P.C., P.K., A.P., N.P., A.K.P., S.P., M.M.S., B.P., P.K., I.Z., S.Á., E.B., M.T., K.S., C.D.C., J.W.S., J.K.O., A.L., J.L.B., T.D.G., J.A.O., B.J.W.D., G.R., M.J.B., K.H., B.J., G.K., V.A.F., H.L.U., G.P., Z.V., H.I., T.B., N.V., C.B.Y.T., V.K., M.F.C., H.D.F., D.I.B., G.G., C.S., E.M.O., D.S., B.S., M.S., T.J.S.L., K.U., E.V., J.S., K.A., C.J.J.v., A.K., K.L.R., K.R.B., D.W., A.R.G., S.M.J.J., T.K.N., C.K.T., J.H.Z., M.V., N.I., M.A., G.B., I.R., V.C., K.P., N.K.S., K.T., C.A.T., A.F., R.M.C.S.H., J.V.V., B.F., H.L., C.L., E.P., M.V., J.O., E.M., A.A., A.A.Ö., B.B., M.A.K., T.G., J.K.M., Y.D., P.S.F., C.R.C., and N.A.C. contributed to Writing (review & editing).

Note: some collaborators did not list their specific contributions.

Competing interests

The authors declare no competing interests

Figure legends

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.

Figure 2. Exploratory factor analysis (EFA) 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 factor is included at the top of each table.

Tables (see following pages)

Table 1. World Regions, Countries, and Localities of 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	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, Ecuador
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. Countries in parentheses were added to the list after acceptance in principle of the Stage 1 protocol. Ecuador was incorrectly classified as Central America and Mexico in our Stage 1 submission, but has been classified as South America for analyses and our Stage 2 submission.

Table 2. Number of participants per region and Cronbach's alphas following data quality checks and exclusions

Region	aggressive	attractive	caring	confident	dominant	emotionally stable	intelligent	mean	responsible	sociable	trustworthy	unhappy	weird
Western Europe	α = 0.978 n = 152	α = 0.991 n = 147	α = 0.976 n = 136	α = 0.985 n = 156	α = 0.973 n = 150	α = 0.981 n = 141	α = 0.975 n = 141	α = 0.969 n = 120	α = 0.978 n = 138	α = 0.988 n = 188	α = 0.978 n = 141	α = 0.983 n = 140	α = 0.982 n = 113
USA & Canada	α = 0.983 n = 248	α = 0.991 n = 224	α = 0.986 n = 257	α = 0.989 n = 303	α = 0.977 n = 246	α = 0.986 n = 270	α = 0.979 n = 239	α = 0.984 n = 270	α = 0.984 n = 269	α = 0.988 n = 246	α = 0.984 n = 263	α = 0.985 n = 252	α = 0.987 n = 226
UK	α = 0.879 n = 16	α = 0.949 n = 22	α = 0.936 n = 34	α = 0.93 n = 30	α = 0.886 n = 34	α = 0.9 n = 30	α = 0.911 n = 34	α = 0.87 n = 27	α = 0.892 n = 37	α = 0.932 n = 28	α = 0.92 n = 27	α = 0.937 n = 24	α = 0.899 n = 18
South America	α = 0.948 n = 97	α = 0.982 n = 108	α = 0.944 n = 112	α = 0.968 n = 108	α = 0.957 n = 121	α = 0.949 n = 100	α = 0.938 n = 110	α = 0.949 n = 95	α = 0.937 n = 117	α = 0.974 n = 110	α = 0.952 n = 107	α = 0.961 n = 87	α = 0.973 n = 116
Scandinavia	α = 0.95 n = 48	α = 0.969 n = 44	α = 0.949 n = 46	α = 0.96 n = 56	α = 0.941 n = 49	α = 0.955 n = 67	α = 0.958 n = 54	α = 0.912 n = 36	α = 0.915 n = 37	α = 0.969 n = 64	α = 0.949 n = 58	α = 0.952 n = 55	α = 0.952 n = 39
Middle East	α = 0.912 n = 32	α = 0.949 n = 32	α = 0.934 n = 42	α = 0.943 n = 39	α = 0.9 n = 35	α = 0.903 n = 33	α = 0.896 n = 48	α = 0.901 n = 36	α = 0.87 n = 34	α = 0.944 n = 41	α = 0.895 n = 42	α = 0.943 n = 57	α = 0.896 n = 32
Eastern Europe	α = 0.941 n = 59	α = 0.971 n = 58	α = 0.926 n = 56	α = 0.946 n = 60	α = 0.952 n = 74	α = 0.923 n = 56	α = 0.939 n = 64	α = 0.937 n = 68	α = 0.953 n = 65	α = 0.955 n = 68	α = 0.937 n = 54	α = 0.964 n = 74	α = 0.956 n = 53
Central America & Mexico	α = 0.845 n = 26	α = 0.93 n = 25	α = 0.788 n = 24	α = 0.89 n = 32	α = 0.859 n = 33	α = 0.835 n = 23	α = 0.832 n = 33	α = 0.817 n = 23	α = 0.824 n = 22	α = 0.882 n = 28	α = 0.851 n = 27	α = 0.771 n = 27	α = 0.842 n = 15
Australia & New Zealand	α = 0.956 n = 77	α = 0.98 n = 88	α = 0.964 n = 90	α = 0.972 n = 93	α = 0.936 n = 66	α = 0.957 n = 88	α = 0.951 n = 81	α = 0.947 n = 71	α = 0.937 n = 68	α = 0.972 n = 95	α = 0.953 n = 72	α = 0.948 n = 85	α = 0.962 n = 70
Asia	α = 0.932 n = 59	α = 0.957 n = 52	α = 0.948 n = 73	α = 0.959 n = 72	α = 0.917 n = 55	α = 0.908 n = 55	α = 0.927 n = 64	α = 0.909 n = 51	α = 0.931 n = 63	α = 0.952 n = 65	α = 0.93 n = 61	α = 0.937 n = 61	α = 0.942 n = 49
Africa	α = 0.808 n = 45	α = 0.873 n = 38	α = 0.865 n = 44	α = 0.805 n = 31	α = 0.79 n = 38	α = 0.779 n = 38	α = 0.756 n = 37	α = 0.889 n = 51	α = 0.811 n = 36	α = 0.819 n = 34	α = 0.867 n = 49	α = 0.795 n = 43	α = 0.889 n = 37

Table 3. Replication criteria for the principal component analysis (PCA) for each region

Region	Component 1		Component 2		Replicated
	Trustworthy	Dominant	Dominant	Trustworthy	
(Oosterhof & Todorov, 2008)	0.941	-0.244	0.929	-0.060	Yes
Africa	0.924	0.271	0.843	-0.065	Yes
Asia	0.922	0.370	0.863	-0.006	Yes
Australia & New Zealand	0.943	0.257	0.907	-0.076	Yes
Central America & Mexico	0.918	0.007	0.915	-0.050	Yes
Eastern Europe	0.938	0.599	0.755	-0.113	No
Middle East	0.831	0.490	0.810	-0.382	Yes
Scandinavia	0.953	0.392	0.881	-0.121	Yes
South America	0.898	0.309	0.905	-0.151	Yes
UK	0.944	0.331	0.851	-0.121	Yes
USA & Canada	0.966	0.406	0.841	-0.073	Yes
Western Europe	0.957	0.357	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 > .7 with trustworthiness and < .5 with dominance, and the second component had a loading > .7 with dominance and < .5 with trustworthiness.

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	0.992	equal	0.935	fairly similar
Eastern Europe	0.953	equal	0.948	fairly similar
Middle East	0.952	equal	0.859	fairly similar
Scandinavia	0.973	equal	0.960	equal
South America	0.976	equal	0.953	equal
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

Table 5. Replication criteria for the exploratory factor analysis (EFA) for each region

	Factor 1		Factor 2		
Region	Trustworthy	Dominant	Dominant	Trustworthy	Replicated
(Oosterhof & Todorov, 2008)	0.826	0.228	0.970	-0.288	Yes
Africa	0.786	0.200	0.069	0.214	No
Asia	0.761	0.487	0.110	0.236	No
Australia & New Zealand	0.730	0.157	0.071	0.281	No
Central America & Mexico	0.268	0.108	0.241	0.591	No
Eastern Europe	0.843	0.750	0.609	-0.322	No
Middle East	0.177	0.502	0.600	-0.686	No
Scandinavia	0.744	0.428	0.293	0.211	No
South America	-0.458	0.778	0.261	0.058	No
UK	0.338	0.249	0.265	0.510	No
USA & Canada	0.768	0.491	0.264	0.189	No
Western Europe	0.398	0.111	0.256	0.164	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 > .7 with trustworthiness and < .5 with dominance, and the second factor had a loading > .7 with dominance and < .5 with trustworthiness.

Figure 1

(Oosterhof & Todorov, 2008)

	Component 1	Component 2	Component 3
Prop.Var	0.63	0.18	
trustworthy	0.94	-0.06	
emostable	0.93	0.19	
responsible	0.91	0.11	
sociable	0.91	0.20	
caring	0.90	-0.29	
attractive	0.81	0.32	
intelligent	0.72	0.13	
confident	0.67	0.65	
weird	-0.87	-0.22	
unhappy	-0.71	0.01	
mean	-0.76	0.55	
aggressive	-0.71	0.66	
dominant	-0.24	0.93	

Africa

	Component 1	Component 2	Component 3
Prop.Var	0.57	0.19	
trustworthy	0.92	-0.07	
emostable	0.82	0.13	
responsible	0.78	0.34	
sociable	0.89	-0.05	
caring	0.81	-0.41	
attractive	0.78	0.34	
intelligent	0.75	0.24	
confident	0.69	0.53	
weird	-0.82	-0.10	
unhappy	-0.75	0.25	
mean	-0.70	0.63	
aggressive	-0.58	0.75	
dominant	0.27	0.84	

Asia

	Component 1	Component 2	Component 3
Prop.Var	0.64	0.18	0.09
trustworthy	0.92	-0.01	-0.27
emostable	0.88	-0.14	0.22
responsible	0.91	0.25	-0.05
sociable	0.93	-0.24	0.16
caring	0.91	-0.30	0.09
attractive	0.76	0.51	-0.24
intelligent	0.83	0.42	-0.08
confident	0.81	0.26	0.47
weird	-0.76	-0.27	0.52
unhappy	-0.73	0.36	-0.53
mean	-0.80	0.50	0.15
aggressive	-0.65	0.63	0.33
dominant	0.37	0.86	0.22

Australia & New Zealand

	Component 1	Component 2	Component 3
Prop.Var	0.63	0.17	0.10
trustworthy	0.94	-0.08	-0.19
emostable	0.92	0.10	0.22
responsible	0.91	0.16	-0.21
sociable	0.93	-0.10	0.24
caring	0.90	-0.25	0.17
attractive	0.75	0.44	-0.30
intelligent	0.79	0.14	-0.27
confident	0.79	0.45	0.36
weird	-0.77	-0.13	0.52
unhappy	-0.68	0.13	-0.69
mean	-0.67	0.65	0.07
aggressive	-0.74	0.62	0.13
dominant	0.26	0.91	0.11

Central America & Mexico

	Component 1	Component 2	Component 3
Prop.Var	0.56	0.18	0.08
trustworthy	0.92	-0.05	-0.14
emostable	0.92	0.06	0.10
responsible	0.85	0.36	-0.03
sociable	0.84	-0.14	0.33
caring	0.78	-0.41	0.07
attractive	0.67	0.46	-0.30
intelligent	0.76	0.27	-0.28
confident	0.74	0.45	0.36
weird	-0.71	-0.15	0.50
unhappy	-0.74	0.07	-0.58
mean	-0.75	0.48	0.14
aggressive	-0.62	0.69	0.16
dominant	0.01	0.92	0.07

Eastern Europe

	Component 1	Component 2	Component 3
Prop.Var	0.65	0.16	0.09
trustworthy	0.94	-0.11	0.20
emostable	0.90	0.05	-0.28
responsible	0.92	0.19	0.14
sociable	0.90	-0.11	-0.25
caring	0.87	-0.28	0.11
attractive	0.88	0.23	0.24
intelligent	0.87	0.08	0.21
confident	0.77	0.53	-0.30
weird	-0.81	-0.12	-0.41
unhappy	-0.60	0.11	0.76
mean	-0.64	0.72	-0.02
aggressive	-0.63	0.74	-0.02
dominant	0.60	0.76	-0.01

Middle East

	Component 1	Component 2	Component 3
Prop.Var	0.54	0.25	0.09
trustworthy	0.83	-0.38	-0.21
emostable	0.94	-0.06	0.04
responsible	0.77	0.39	0.11
sociable	0.87	-0.28	0.28
caring	0.83	-0.41	0.06
attractive	0.66	0.48	-0.38
intelligent	0.72	0.52	-0.19
confident	0.81	0.45	0.26
weird	-0.70	-0.10	0.61
unhappy	-0.72	0.10	-0.58
mean	-0.44	0.84	0.09
aggressive	-0.55	0.77	0.13
dominant	0.49	0.81	0.11

Scandinavia

	Component 1	Component 2	Component 3
Prop.Var	0.64	0.18	0.08
trustworthy	0.95	-0.12	-0.14
emostable	0.95	0.04	0.13
responsible	0.91	0.12	-0.12
sociable	0.94	-0.11	0.20
caring	0.88	-0.30	0.08
attractive	0.80	0.35	-0.31
intelligent	0.79	0.25	-0.26
confident	0.77	0.50	0.34
weird	-0.75	-0.15	0.52
unhappy	-0.74	0.07	-0.62
mean	-0.69	0.68	0.06
aggressive	-0.63	0.72	0.03
dominant	0.39	0.88	0.11

South America

	Component 1	Component 2	Component 3
Prop.Var	0.59	0.21	0.08
trustworthy	0.90	-0.15	-0.24
emostable	0.95	0.02	0.09
responsible	0.82	0.18	-0.21
sociable	0.90	-0.05	0.37
caring	0.76	-0.55	0.12
attractive	0.71	0.48	-0.21
intelligent	0.77	0.24	-0.32
confident	0.79	0.50	0.31
weird	-0.77	-0.22	0.42
unhappy	-0.79	0.01	-0.57
mean	-0.73	0.64	0.09
aggressive	-0.60	0.76	0.09
dominant	0.31	0.91	0.08

UK

	Component 1	Component 2	Component 3
Prop.Var	0.60	0.18	0.10
trustworthy	0.94	-0.12	-0.10
emostable	0.89	0.06	0.18
responsible	0.87	0.24	-0.23
sociable	0.92	-0.11	0.25
caring	0.87	-0.31	0.12
attractive	0.73	0.48	-0.31
intelligent	0.80	0.15	-0.32
confident	0.78	0.41	0.39
weird	-0.76	-0.23	0.46
unhappy	-0.65	0.24	-0.66
mean	-0.67	0.63	0.12
aggressive	-0.62	0.71	0.18
dominant	0.33	0.85	0.23

USA & Canada

	Component 1	Component 2	Component 3
Prop.Var	0.66	0.17	0.08
trustworthy	0.97	-0.07	-0.17
emostable	0.95	0.01	0.16
responsible	0.92	0.13	-0.23
sociable	0.95	-0.04	0.21
caring	0.89	-0.30	0.14
attractive	0.72	0.48	-0.27
intelligent	0.83	0.17	-0.27
confident	0.81	0.42	0.37
weird	-0.79	-0.20	0.45
unhappy	-0.73	0.19	-0.62
mean	-0.70	0.68	0.05
aggressive	-0.70	0.67	0.18
dominant	0.41	0.84	0.16

Western Europe

	Component 1	Component 2	Component 3
Prop.Var	0.65	0.19	0.08
trustworthy	0.96	-0.17	0.14
emostable	0.95	0.08	-0.17
responsible	0.90	0.30	0.11
sociable	0.93	-0.19	-0.22
caring	0.86	-0.36	0.04
attractive	0.82	0.36	0.31
intelligent	0.81	0.24	0.24
confident	0.83	0.45	-0.30
weird	-0.78	-0.15	-0.50
unhappy	-0.74	0.06	0.61
mean	-0.64	0.73	-0.05
aggressive	-0.69	0.70	0.01
dominant	0.36	0.88	-0.15

Figure 2

(Oosterhof & Todorov, 2008)

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.56	0.23		
trustworthy	0.83	-0.29		
emostable	0.94	-0.04		
responsible	0.88	-0.12		
sociable	0.92	-0.02		
caring	0.68	-0.50		
attractive	0.86	0.10		
intelligent	0.67	-0.06		
confident	0.92	0.47		
weird	-0.88	0.01		
unhappy	-0.61	0.17		
mean	-0.42	0.72		
aggressive	-0.32	0.84		
dominant	0.23	0.97		

Africa

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.37	0.25	0.16	
trustworthy	0.79	0.21	-0.13	
emostable	0.28	0.62	0.27	
responsible	0.76	0.05	0.21	
sociable	0.68	0.28	-0.08	
caring	0.34	0.57	-0.28	
attractive	0.96	-0.16	0.12	
intelligent	0.59	0.18	0.17	
confident	0.31	0.45	0.63	
weird	-0.94	0.09	0.09	
unhappy	0.07	-0.94	-0.06	
mean	-0.17	-0.65	0.48	
aggressive	-0.24	-0.42	0.68	
dominant	0.20	0.07	0.80	

Asia

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.37	0.36	0.16	
trustworthy	0.76	0.24	-0.25	
emostable	0.26	0.72	-0.09	
responsible	0.69	0.36	0.08	
sociable	0.26	0.75	-0.20	
caring	0.28	0.69	-0.29	
attractive	0.90	0.01	0.20	
intelligent	0.77	0.23	0.21	
confident	0.18	0.83	0.41	
weird	-1.01	0.17	0.14	
unhappy	0.25	-1.06	0.06	
mean	-0.31	-0.46	0.57	
aggressive	-0.31	-0.26	0.74	
dominant	0.49	0.11	0.79	

Australia & New Zealand

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.30	0.23	0.22	0.15
trustworthy	0.73	0.28	0.06	-0.08
emostable	0.17	0.43	0.57	0.05
responsible	0.54	0.44	0.06	0.11
sociable	0.64	-0.01	0.50	-0.03
caring	0.74	-0.07	0.39	-0.16
attractive	0.63	0.31	-0.12	0.39
intelligent	-0.09	0.94	0.13	-0.06
confident	0.12	0.28	0.66	0.45
weird	-0.42	-0.64	0.22	-0.02
unhappy	-0.00	0.03	-0.98	0.07
mean	-0.31	-0.25	-0.19	0.64
aggressive	-0.50	-0.21	-0.11	0.63
dominant	0.16	0.07	0.12	0.87

Central America & Mexico

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.26	0.23	0.21	0.11
trustworthy	0.27	0.59	-0.31	0.06
emostable	0.51	0.19	-0.23	0.37
responsible	0.34	0.38	0.02	0.46
sociable	0.81	0.25	-0.14	-0.25
caring	0.43	0.33	-0.46	-0.22
attractive	0.04	0.83	0.19	0.05
intelligent	0.04	0.40	-0.16	0.55
confident	0.75	0.15	0.31	0.24
weird	0.11	-0.76	0.17	-0.12
unhappy	-0.93	0.22	0.10	-0.10
mean	-0.14	-0.12	0.74	-0.23
aggressive	-0.05	-0.07	0.90	-0.09
dominant	0.11	0.24	0.79	0.23

Eastern Europe

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.49	0.19	0.19	
trustworthy	0.84	-0.32	0.03	
emostable	0.50	-0.09	0.54	
responsible	0.89	-0.02	0.09	
sociable	0.47	-0.25	0.51	
caring	0.64	-0.43	0.12	
attractive	0.95	0.02	-0.03	
intelligent	0.83	-0.12	0.03	
confident	0.57	0.41	0.54	
weird	-0.92	0.09	0.17	
unhappy	0.17	0.12	-1.01	
mean	-0.23	0.82	-0.16	
aggressive	-0.21	0.85	-0.16	
dominant	0.75	0.61	0.15	

Middle East

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.29	0.28	0.26	
trustworthy	0.18	-0.69	0.44	
emostable	0.54	-0.36	0.38	
responsible	0.56	0.10	0.41	
sociable	0.74	-0.47	-0.01	
caring	0.45	-0.63	0.16	
attractive	0.07	0.10	0.82	
intelligent	0.27	0.16	0.71	
confident	0.80	0.21	0.29	
weird	0.18	0.30	-0.89	
unhappy	-0.89	0.18	0.21	
mean	-0.03	0.93	0.06	
aggressive	-0.06	0.92	-0.06	
dominant	0.50	0.60	0.45	

Scandinavia

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.41	0.27	0.20	
trustworthy	0.74	0.21	-0.30	
emostable	0.50	0.54	-0.10	
responsible	0.75	0.23	-0.06	
sociable	0.39	0.61	-0.23	
caring	0.42	0.43	-0.42	
attractive	0.91	-0.01	0.15	
intelligent	0.78	0.07	0.05	
confident	0.32	0.72	0.41	
weird	-0.98	0.22	0.06	
unhappy	0.18	-1.04	0.10	
mean	-0.31	-0.20	0.79	
aggressive	-0.22	-0.21	0.80	
dominant	0.43	0.29	0.81	

South America

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.25	0.24	0.20	0.19
trustworthy	-0.46	0.06	0.25	0.54
emostable	-0.21	0.47	0.41	0.16
responsible	-0.07	0.07	0.86	0.01
sociable	-0.23	0.80	-0.13	0.30
caring	-0.70	0.39	-0.05	0.19
attractive	0.17	0.10	0.06	0.82
intelligent	-0.02	-0.02	0.80	0.12
confident	0.34	0.71	0.21	0.23
weird	0.12	0.10	-0.27	-0.72
unhappy	0.08	-0.96	-0.14	0.20
mean	0.83	-0.13	-0.18	-0.12
aggressive	0.91	-0.07	-0.17	0.03
dominant	0.78	0.26	0.28	0.28

UK

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.26	0.25	0.18	0.18
trustworthy	0.34	0.51	-0.30	0.19
emostable	0.47	0.03	-0.09	0.58
responsible	0.06	0.45	-0.06	0.54
sociable	0.73	0.32	-0.13	0.04
caring	0.64	0.48	-0.32	-0.18
attractive	-0.01	0.82	0.22	0.16
intelligent	-0.08	0.31	-0.19	0.68
confident	0.67	0.13	0.39	0.34
weird	0.09	-0.81	0.06	-0.17
unhappy	-0.95	0.23	0.07	-0.06
mean	-0.23	-0.08	0.72	-0.21
aggressive	-0.15	-0.08	0.85	-0.18
dominant	0.25	0.27	0.77	0.21

USA & Canada

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.41	0.28	0.19	
trustworthy	0.77	0.19	-0.30	
emostable	0.47	0.57	-0.12	
responsible	0.87	0.11	-0.11	
sociable	0.39	0.64	-0.14	
caring	0.31	0.54	-0.41	
attractive	0.86	-0.00	0.22	
intelligent	0.80	0.07	-0.07	
confident	0.35	0.76	0.38	
weird	-0.97	0.15	0.08	
unhappy	0.22	-1.05	0.13	
mean	-0.19	-0.27	0.79	
aggressive	-0.32	-0.12	0.84	
dominant	0.49	0.26	0.72	

Western Europe

	Factor 1	Factor 2	Factor 3	Factor 4
Prop.Var	0.27	0.24	0.24	0.17
trustworthy	0.40	0.16	0.56	-0.17
emostable	0.34	0.56	0.24	0.03
responsible	0.58	0.22	0.29	0.19
sociable	0.03	0.57	0.51	-0.10
caring	-0.00	0.21	0.80	-0.21
attractive	0.55	-0.06	0.49	0.30
intelligent	1.03	0.11	-0.18	-0.08
confident	0.26	0.68	0.11	0.42
weird	-0.68	0.20	-0.42	-0.03
unhappy	0.02	-0.99	0.06	0.06
mean	-0.13	-0.18	-0.38	0.70
aggressive	-0.17	-0.29	-0.31	0.71
dominant	0.11	0.26	0.14	0.87

Code

PSA001: Stage 2 RR Analyses

2020-11-05

1 Load Data

Code

```
## Loading required package: MASS
```

Code

```
## ── Attaching packages ─────────────────────────────────────────── tidyverse 1.3.0 ──
```

```
## ✓ ggplot2 3.3.0     ✓ purrr   0.3.4
## ✓ tibble  3.0.1     ✓ dplyr   0.8.5
## ✓ tidyr   1.0.2     ✓ stringr 1.4.0
## ✓ readr   1.3.1     ✓ forcats 0.5.0
```

```
## ── Conflicts ────────────────────────────────────────────── tidyverse_conflicts() ──
## x ggplot2::%+%()      masks psych::%+%()
## x ggplot2::alpha()    masks psych::alpha()
## x dplyr::filter()     masks stats::filter()
## x dplyr::group_rows() masks kableExtra::group_rows()
## x dplyr::lag()        masks stats::lag()
## x dplyr::select()     masks MASS::select()
```

Code

```
## [1] "R version 4.0.0 (2020-04-24)"
```

Code

1.1 Simulate Study Data (for Stage 1 RR)

See https://osf.io/87rbg/ (https://osf.io/87rbg/) for Stage 1 RR code. The code below is modified from the original to account for a different raw data structure and to add additional tables and graphs. All analysis code is identical.

1.2 Load Study Data (for Stage 2 RR)

Load study data and demographic questionnaires from the data folder.

Code

Join experiment and questionnaire data

Code

1.3 Load Auxillary Data

Data on regions and stimuli.

1.3.1 Load Region Data

Code

1.3.2 Load Stimulus Info

Code

ethnicity	gender	n	mean_age	sd_age
asian	female	15	26.15	3.33
asian	male	15	26.40	3.21
black	female	15	27.00	3.51
black	male	15	28.07	4.27
latinx	female	15	25.27	2.42
latinx	male	15	26.31	4.00
white	female	15	25.77	3.03
white	male	15	26.06	4.46

Code

Stimuli in our study will be an open-access, full-color, face image set consisting of 60 men and 60 women (mean age=26.38 years, SD=3.57 years, range=18.7307692 to 34.9310345 years), taken under standardized photographic conditions (Ma et al., 2015).

1.3.3 Load O&T 2008 Data

Code

1.4 Data Processing

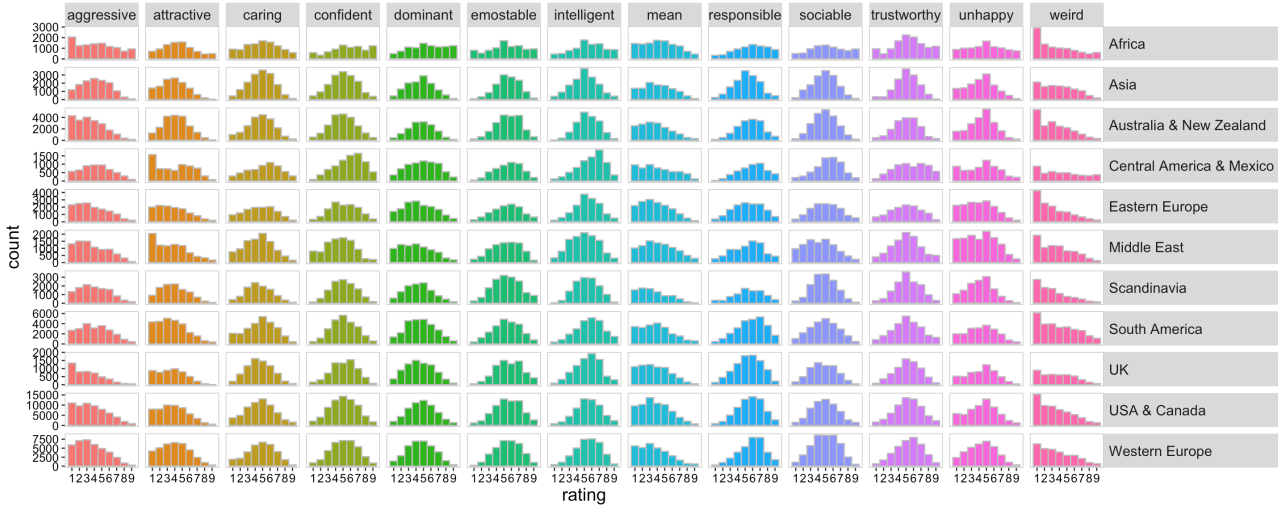
1.4.1 Join Data

Code

1.4.2 Graph distributions for trait by region

Code

Code



1.5 Data checks

Code

1.5.1 How many participants completed at least one rating for each of 120 stimuli

Participants rated the 120 stimuli in two blocks. Some participants quit the study before completing both. Our pre-registered exclusion criteria require that we have at least one rating for each of the 120 stimuli.

Code

region	rated < 120	rated all 120
Africa	12	553
Asia	49	822
Australia & New Zealand	21	1105
Central America & Mexico	14	350
Eastern Europe	106	847
Middle East	14	525
Scandinavia	48	689
South America	49	1464
UK	4	382
USA & Canada	66	3528
Western Europe	39	1973

1.5.2 Participants who did not complete exactly 240 trials

In rare cases, participants completed more than 240 trials because they restarted the study before completion (e.g., when a wifi outage required a page refresh).

Code

region	rated < 120	rated > 240	rated 120-239	rated 240
Africa	10	4	103	448

region	rated < 120	rated > 240	rated 120-239	rated 240
Asia	47	9	209	606
Australia & New Zealand	18	5	222	881
Central America & Mexico	11	0	98	255
Eastern Europe	104	0	224	625
Middle East	13	3	71	452
Scandinavia	47	2	150	538
South America	48	7	304	1154
UK	4	2	47	333
USA & Canada	65	2	382	3145
Western Europe	37	4	170	1801

1.5.3 Participants with low-variance responses in block 1

Code

trait	Africa	Asia	Australia & New Zealand	Central America & Mexico	Eastern Europe	Middle East	Scandinavia	South America	UK	USA & Canada	Western Europe	TOTAL
aggressive	4	5	16	0	2	3	2	4	4	21	14	75
attractive	2	4	3	3	1	4	3	13	0	16	9	58
caring	1	0	2	1	1	1	0	5	0	16	3	30
confident	4	1	0	1	1	1	1	4	0	4	2	19
dominant	4	4	0	1	2	1	1	2	0	10	4	29
emostable	3	3	5	1	1	2	4	3	4	14	10	50
intelligent	2	5	4	0	5	1	4	4	2	15	7	49
mean	2	4	6	2	5	2	6	6	3	27	17	80
responsible	0	3	5	0	3	0	2	5	3	19	8	48
sociable	1	1	0	0	0	1	1	5	0	12	3	24
trustworthy	3	6	2	1	5	2	6	6	3	15	15	64
unhappy	1	3	6	0	2	1	3	4	0	12	5	37
weird	6	3	12	3	10	3	5	15	1	36	13	107
TOTAL	33	42	61	13	38	22	38	76	20	217	110	670

1.5.4 Participants with no region

Code

```
## # A tibble: 0 x 3
## # … with 3 variables: user_id <dbl>, country <chr>, lab <chr>
```

1.5.5 Remove excluded data and average ratings

N.B. Oosterhof & Todorov (2008) additionally standardised the means of each rater before computing the mean for each face. The resulting scores are correlated >.99, so we retained our pre-registered protocol.

Code

1.6 Participant Demographics

Code

language	n
EL	99
ENG	5860
ES-PE	115
FAS	48
FR-BE	86
FR-CH	109
FRE	211
GER	452
HU	174
ITA	153
NL	235
NOR	326
POL	36
PT	76
PT-BR	201
RO	27
RU	240
SLO	263
SPA	1680
SRP	69
SV	198
THA	81
TUR	364
ZH-CN	101
ZH-S	366

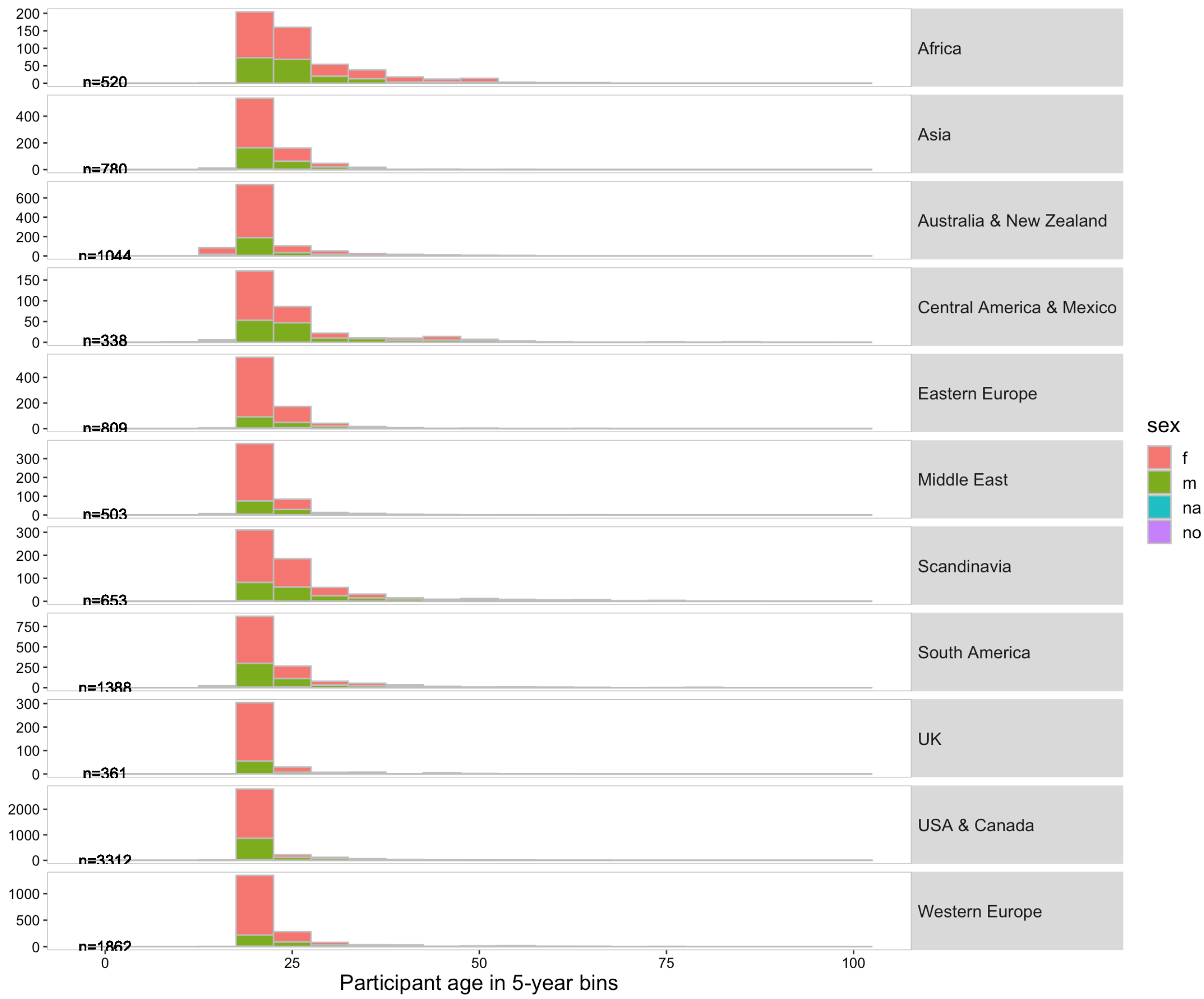
Code

region	n
Africa	520
Asia	780

region	n
Australia & New Zealand	1044
Central America & Mexico	338
Eastern Europe	809
Middle East	503
Scandinavia	653
South America	1388
UK	361
USA & Canada	3312
Western Europe	1862
TOTAL	11570

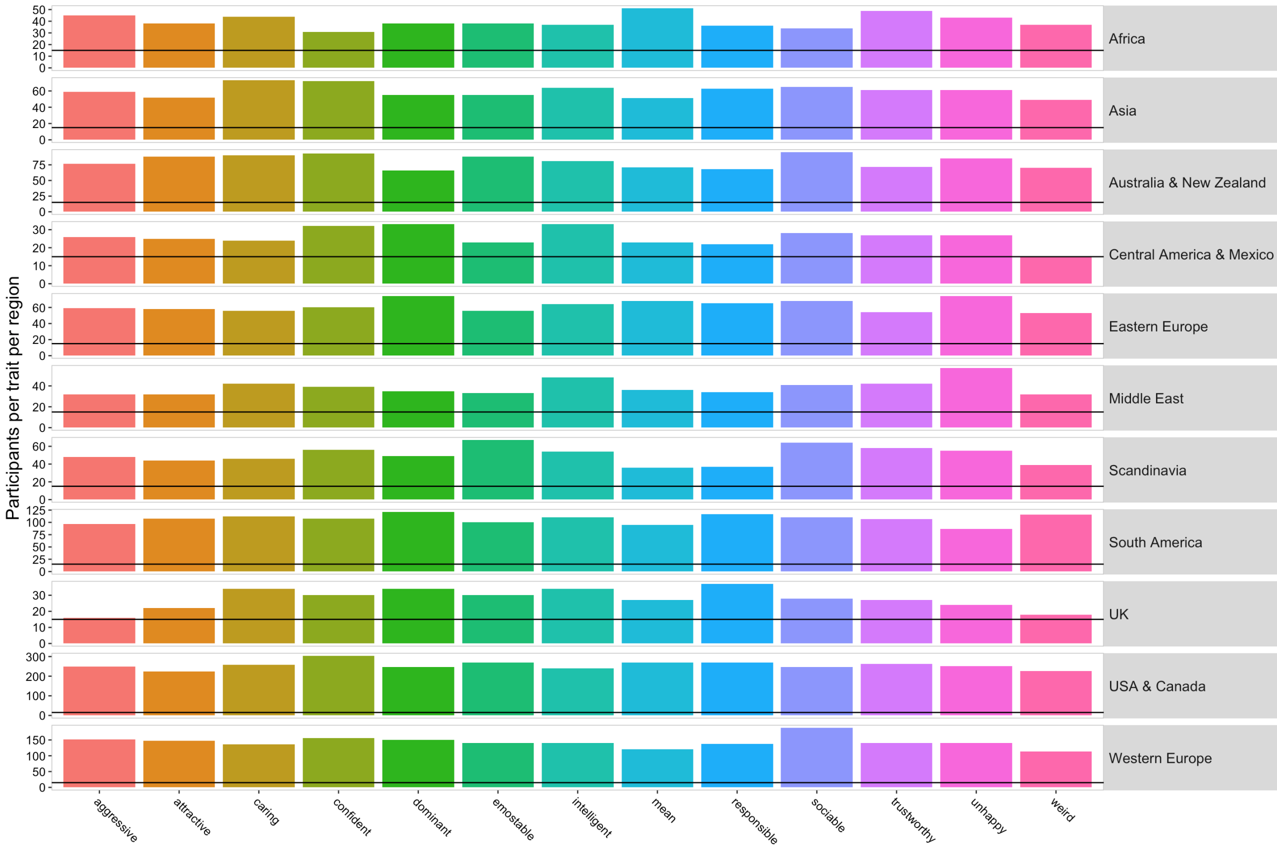
1.6.1 Age and sex distribution per region

Code



1.6.2 Participants per trait per region

Code



Code

1.6.3 Participants per lab

Code

region	lab	n
Africa	KEN_001	166
Africa	KEN_002	181
Africa	NGA_001	43
Africa	RSA_001	47
Africa	RSA_002	83
Asia	CHN_001	45
Asia	CHN_006	121
Asia	CHN_007	56
Asia	CHN_014	94
Asia	MAS_001	70
Asia	MAS_002	53
Asia	MAS_003	44
Asia	MAS_004	114
Asia	TAI_001	31
Asia	TAI_002	71

region	lab	n
Asia	THA_001	81
Australia & New Zealand	AUS_004	80
Australia & New Zealand	AUS_005	88
Australia & New Zealand	AUS_006	193
Australia & New Zealand	AUS_007	251
Australia & New Zealand	AUS_008	103
Australia & New Zealand	AUS_011	84
Australia & New Zealand	AUS_014	86
Australia & New Zealand	NZL_001	20
Australia & New Zealand	NZL_002	139
Central America & Mexico	MEX_002	182
Central America & Mexico	MEX_003	67
Central America & Mexico	SLV_001	89
Eastern Europe	HUN_001	174
Eastern Europe	POL_001	36
Eastern Europe	ROU_001	27
Eastern Europe	RUS_005	240
Eastern Europe	SRB_002	69
Eastern Europe	SVK_001	79
Eastern Europe	SVK_002	184
Middle East	IRI_001	48
Middle East	TUR_001	74
Middle East	TUR_003	102
Middle East	TUR_007	100
Middle East	TUR_009	99
Middle East	UAE_001	80
Scandinavia	DNK_001	92
Scandinavia	FIN_001	46
Scandinavia	NOR_001	139
Scandinavia	NOR_002	111
Scandinavia	NOR_003	63
Scandinavia	NOR_004	50
Scandinavia	SWE_004	59
Scandinavia	SWE_005	40

region	lab	n
Scandinavia	SWE_006	53
South America	ARG_001	91
South America	BRA_001	113
South America	BRA_003	88
South America	CHI_001	163
South America	CHI_003	89
South America	CHI_004	88
South America	CHI_005	90
South America	COL_003	47
South America	COL_004	399
South America	ECU_001	105
South America	PER_001	63
South America	PER_002	52
UK	UK_001	34
UK	UK_005	102
UK	UK_006	44
UK	UK_011	65
UK	UK_018	38
UK	UK_022	34
UK	UK_024	44
USA & Canada	CAN_001	45
USA & Canada	CAN_008	67
USA & Canada	CAN_015	83
USA & Canada	CAN_017	127
USA & Canada	CAN_018	308
USA & Canada	PSA_001	80
USA & Canada	PSA_002	329
USA & Canada	USA_001	90
USA & Canada	USA_003	81
USA & Canada	USA_005	39
USA & Canada	USA_011	35
USA & Canada	USA_014	65
USA & Canada	USA_020	194
USA & Canada	USA_025	91

region	lab	n
USA & Canada	USA_026	81
USA & Canada	USA_030	98
USA & Canada	USA_031	84
USA & Canada	USA_033	55
USA & Canada	USA_036	43
USA & Canada	USA_038	77
USA & Canada	USA_039	123
USA & Canada	USA_042	31
USA & Canada	USA_050	158
USA & Canada	USA_051	142
USA & Canada	USA_054	90
USA & Canada	USA_055	113
USA & Canada	USA_065	51
USA & Canada	USA_067	21
USA & Canada	USA_075	122
USA & Canada	USA_083	17
USA & Canada	USA_113	120
USA & Canada	USA_114	104
USA & Canada	USA_115	52
USA & Canada	USA_116	48
USA & Canada	USA_117	48
Western Europe	AUT_001	100
Western Europe	AUT_002	86
Western Europe	AUT_005	50
Western Europe	BEL_001	86
Western Europe	ESP_001	108
Western Europe	ESP_005	117
Western Europe	ESP_006	45
Western Europe	FRA_003	75
Western Europe	FRA_004	44
Western Europe	FRA_005	32
Western Europe	FRA_006	60
Western Europe	GER_011	50
Western Europe	GER_012	61

region	lab	n
Western Europe	GER_015	97
Western Europe	GER_017	58
Western Europe	GRE_002	99
Western Europe	ITA_001	89
Western Europe	ITA_003	64
Western Europe	NED_008	121
Western Europe	NED_009	235
Western Europe	POR_001	76
Western Europe	SUI_003	109

2 Analyses

2.1 Main Analysis

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.

2.1.1 Calculate Alphas

Code

Code

Code

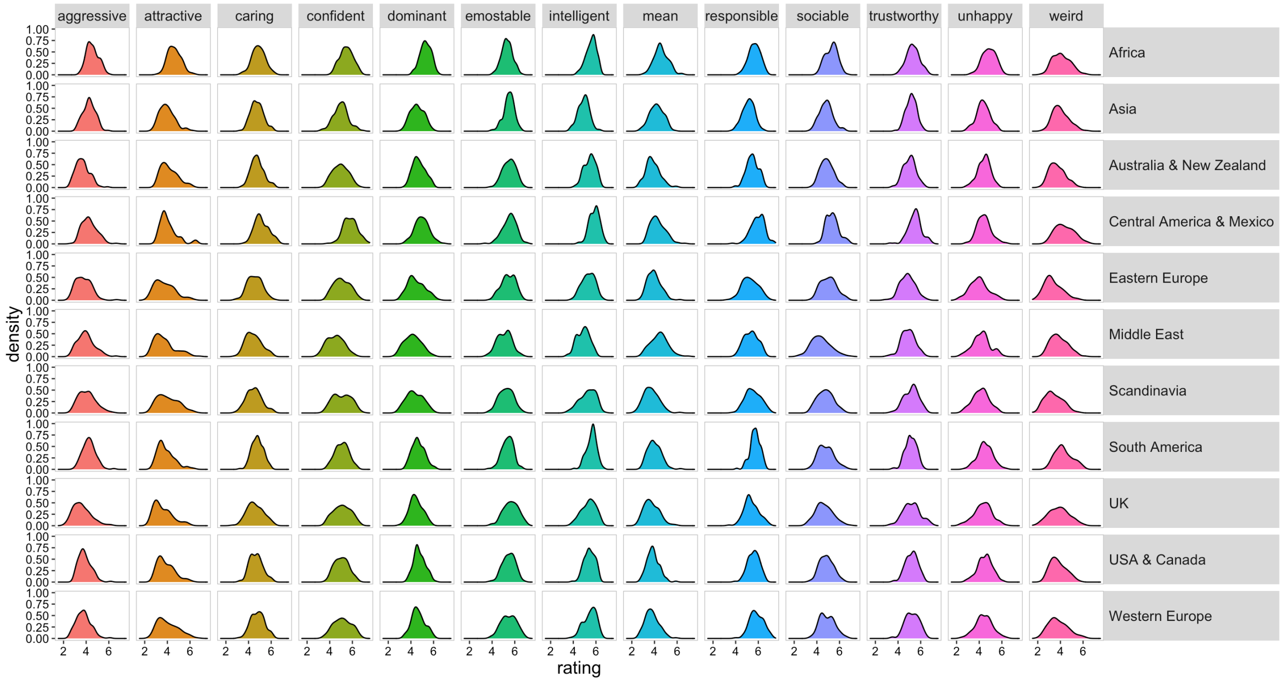
	aggressive	attractive	caring	confident	dominant	emostable	intelligent	mean	responsible	sociable	trustworthy	unhappy	weird
Africa	α = 0.81 n = 45	α = 0.87 n = 38	α = 0.86 n = 44	α = 0.81 n = 31	α = 0.79 n = 38	α = 0.78 n = 38	α = 0.76 n = 37	α = 0.89 n = 51	α = 0.81 n = 36	α = 0.82 n = 34	α = 0.87 n = 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 = 0.84 n = 26	α = 0.93 n = 25	α = 0.79 n = 24	α = 0.89 n = 32	α = 0.86 n = 33	α = 0.83 n = 23	α = 0.83 n = 33	α = 0.82 n = 23	α = 0.82 n = 22	α = 0.88 n = 28	α = 0.85 n = 27	α = 0.77 n = 27	α = 0.84 n = 15
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 = 0.91 n = 32	α = 0.95 n = 32	α = 0.93 n = 42	α = 0.94 n = 39	α = 0.90 n = 35	α = 0.90 n = 33	α = 0.90 n = 48	α = 0.90 n = 36	α = 0.87 n = 34	α = 0.94 n = 41	α = 0.90 n = 42	α = 0.94 n = 57	α = 0.90 n = 32
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 = 97	α = 0.98 n = 108	α = 0.94 n = 112	α = 0.97 n = 108	α = 0.96 n = 121	α = 0.95 n = 100	α = 0.94 n = 110	α = 0.95 n = 95	α = 0.94 n = 117	α = 0.97 n = 110	α = 0.95 n = 107	α = 0.96 n = 87	α = 0.97 n = 116
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

Code

2.1.2 Calculate Aggregate Scores

Code

Code



Code

2.1.3 Principal Component Analysis (PCA)

The number of components to extract was determined using eigenvalues > 1 for each world region. PCA was conducted using the `psych::principal()` function with `rotate="none"`.

Code

Code

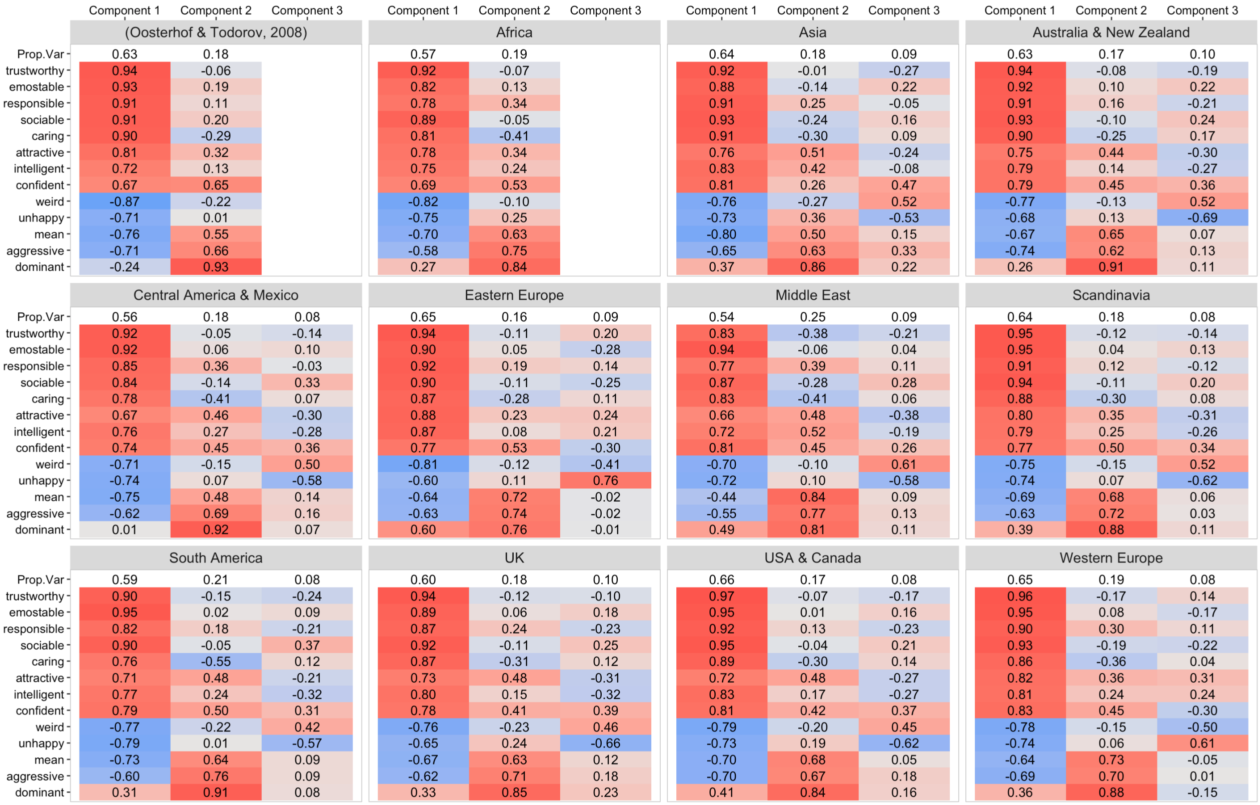
2.1.3.1 Number of Components (and proportion variance) by region

Code

region	nPCs	Component 1	Component 2	Component 3
(Oosterhof & Todorov, 2008)	2	0.633	0.183	
Africa	2	0.566	0.193	
Asia	3	0.644	0.179	0.090
Australia & New Zealand	3	0.627	0.169	0.099
Central America & Mexico	3	0.561	0.183	0.083
Eastern Europe	3	0.646	0.165	0.088
Middle East	3	0.535	0.248	0.087
Scandinavia	3	0.636	0.181	0.081
South America	3	0.589	0.212	0.079
UK	3	0.598	0.180	0.098
USA & Canada	3	0.658	0.172	0.085
Western Europe	3	0.650	0.192	0.077

2.1.3.2 Trait Loadings by Region and Component

Code



Code

2.1.3.3 Replication Criteria (PCA)

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 > .7) 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 > .7) 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.

Code

	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	0.007	0.918	0.915	-0.050	Yes
Eastern Europe	0.599	0.938	0.755	-0.113	No
Middle East	0.490	0.831	0.810	-0.382	Yes
Scandinavia	0.392	0.953	0.881	-0.121	Yes
South America	0.309	0.898	0.905	-0.151	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

2.1.4 Factor Congruence (PCA)

This analysis determines the congruence between the components from Oosterhof & Todorov (2008) and the components in each world region, using the `psych::factor.congruence` function. Congruence is labeled “not similar” for values < 0.85, “fairly similar”, for values < 0.95, and “equal” for values >= 0.95.

Code

	Component 1		Component 2	
Region	**Congruence**	**Conclusion**	**Congruence**	**Conclusion**
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	0.992	equal	0.935	fairly similar
Eastern Europe	0.953	equal	0.948	fairly similar
Middle East	0.952	equal	0.859	fairly similar
Scandinavia	0.973	equal	0.960	equal
South America	0.976	equal	0.953	equal
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

2.2 Robustness Checks

2.2.1 Exploratory Factor Analysis (EFA)

The number of factors to extract was determined using parallel analysis (`paran::paran()`) for each world region. EFA was conducted using the `psych::fa()` function with all default options.

Change to registered analysis:

The analysis for USA & Canada would not converge, so the number of factors was manually reduced from 3 to 4 in the code below.

Code

Calculate for each region

Code

2.2.1.1 Number of Factors (and proportion variance) by region

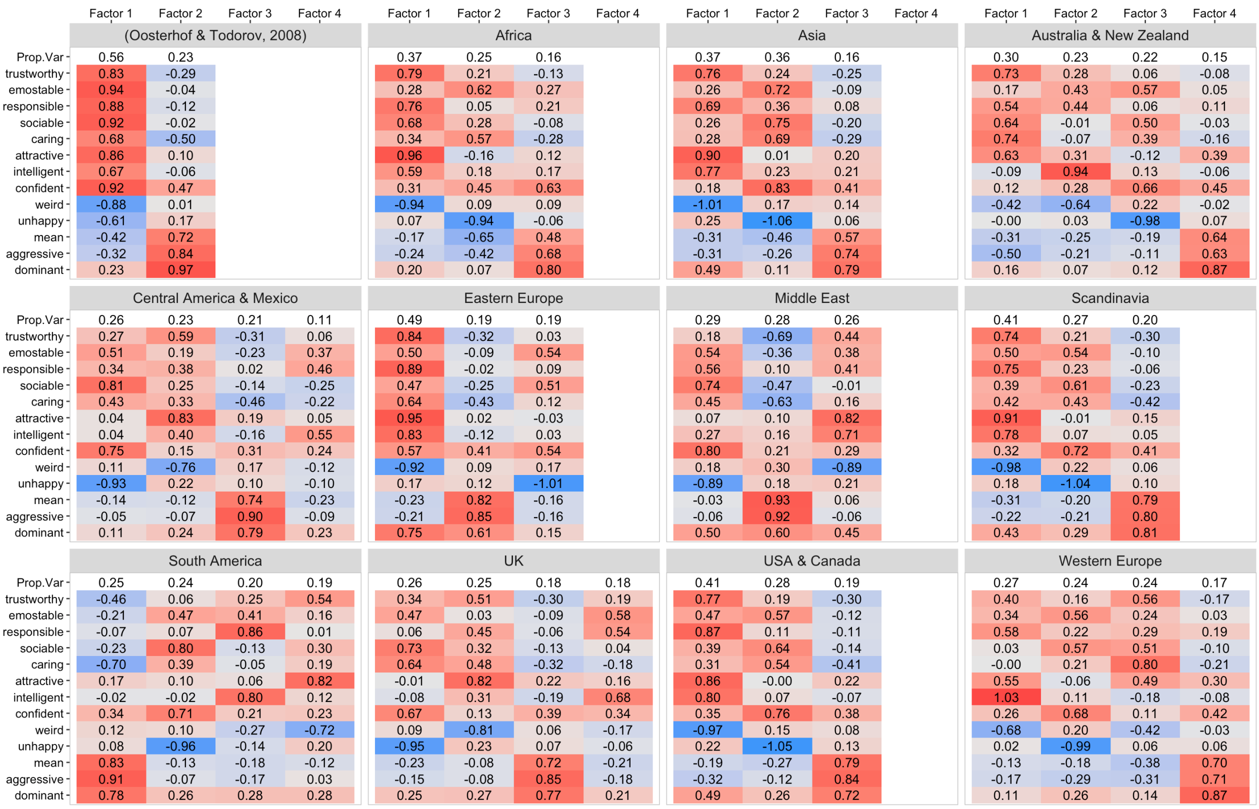
Code

region	nMRs	Factor 1	Factor 2	Factor 3	Factor 4
(Oosterhof & Todorov, 2008)	2	0.564	0.227		
Africa	3	0.371	0.246	0.158	
Asia	3	0.372	0.358	0.157	
Australia & New Zealand	4	0.302	0.225	0.216	0.154
Central America & Mexico	4	0.262	0.230	0.212	0.107
Eastern Europe	3	0.491	0.191	0.191	
Middle East	3	0.293	0.280	0.257	
Scandinavia	3	0.405	0.268	0.196	

region	nMRs	Factor 1	Factor 2	Factor 3	Factor 4
South America	4	0.249	0.243	0.202	0.193
UK	4	0.264	0.246	0.185	0.177
USA & Canada	3	0.414	0.283	0.194	
Western Europe	4	0.269	0.245	0.244	0.171

2.2.1.2 Trait Loadings by Region and Factor

Code



Code

2.2.1.3 Replication Criteria (EFA)

Oosterhof and Todorov's valence-dominance model will be judged to have been replicated in a given world region if the the first factor is correlated strongly (loading > .7) with trustworthiness and weakly (loading < .5) with dominance, and the second factor is correlated strongly (loading > .7) with dominance and weakly (loading < .5) with trustworthiness. All these criteria need to be met to conclude that the model was replicated in a given world region.

Code

	Factor 1		Factor 2		
Region	Trustworthy	Dominant	Dominant	Trustworthy	Replicated
(Oosterhof & Todorov, 2008)	0.826	0.228	0.970	-0.288	Yes
Africa	0.786	0.200	0.069	0.214	No
Asia	0.761	0.487	0.110	0.236	No
Australia & New Zealand	0.730	0.157	0.071	0.281	No
Central America & Mexico	0.268	0.108	0.241	0.591	No
Eastern Europe	0.843	0.750	0.609	-0.322	No

Region	Factor 1		Factor 2		
	Trustworthy	Dominant	Dominant	Trustworthy	Replicated
Middle East	0.177	0.502	0.600	-0.686	No
Scandinavia	0.744	0.428	0.293	0.211	No
South America	-0.458	0.778	0.261	0.058	No
UK	0.338	0.249	0.265	0.510	No
USA & Canada	0.768	0.491	0.264	0.189	No
Western Europe	0.398	0.111	0.256	0.164	No

2.2.2 Factor Congruence (EFA)

This analysis determines the congruence between the factors from Oosterhof & Todorov (2008) and the factors in each world region, using the `psych::factor.congruence` function. Congruence is labeled "not similar" for values < 0.85, "fairly similar", for values < 0.09, and "equal" for values >= 0.95.

Code

Region	Factor 1		Factor 2	
	Congruence	Conclusion	Congruence	Conclusion
Africa	0.894	fairly similar	-0.412	not similar
Asia	0.810	not similar	-0.224	not similar
Australia & New Zealand	0.810	not similar	-0.125	not similar
Central America & Mexico	0.753	not similar	-0.060	not similar
Eastern Europe	0.891	fairly similar	0.957	equal
Middle East	0.763	not similar	0.859	fairly similar
Scandinavia	0.884	fairly similar	-0.078	not similar
South America	-0.318	not similar	0.013	not similar
UK	0.702	not similar	-0.061	not similar
USA & Canada	0.867	fairly similar	-0.082	not similar
Western Europe	0.774	not similar	-0.076	not similar

2.2.3 Replication Criteria for "best" factor (EFA)

Run replication criteria on the two factors that load highest on dominance and trustworthiness, regardless of factor position. Here, "replicated" means that the factor that loads highest on trustworthiness loads has a loading > 0.7 for trustworthy and < .5 for dominant, while the factor that loads highest on dominance loads has a loading > 0.7 for dominant and < .5 for trustworthy.

Code

Region	Trust Factor			Dominance Factor			
	#	Trustworthy	Dominant	#	Dominant	Trustworthy	Replicated
(Oosterhof & Todorov, 2008)	1	0.826	0.228	2	0.970	-0.288	Yes
Africa	1	0.786	0.200	3	0.796	-0.133	Yes
Asia	1	0.761	0.487	3	0.785	-0.251	Yes
Australia & New Zealand	1	0.730	0.157	4	0.873	-0.078	Yes

	Trust Factor			Dominance Factor			
Region	**#**	**Trustworthy**	**Dominant**	**#**	**Dominant**	**Trustworthy**	**Replicated**
Central America & Mexico	2	0.591	0.241	3	0.787	-0.307	No
Eastern Europe	1	0.843	0.750	1	0.750	0.843	No
Middle East	3	0.440	0.447	2	0.600	-0.686	No
Scandinavia	1	0.744	0.428	3	0.806	-0.304	Yes
South America	4	0.544	0.278	1	0.778	-0.458	No
UK	2	0.510	0.265	3	0.766	-0.299	No
USA & Canada	1	0.768	0.491	3	0.720	-0.303	Yes
Western Europe	3	0.560	0.139	4	0.869	-0.172	No

````
---
title: 'PSA001: Stage 2 RR Analyses'
date: 2020-11-05
output:
  html_document:
    code_folding: hide
    number_sections: true
    toc: yes
    toc_depth: 5
---

# Load Data

```{r libraries, messages = FALSE}
# psych 1.9.12.31 has a bug in fa.congruence, so use this version
#install.packages("psych",repos="http://personality-project.org/r",
type ="source")
library(psych) # for SPSS-style PCA
library(paran) # for parallel analyses
library(GPArotation) # for robustness checks
library(kableExtra) # for nice tables
library(tidyverse) # for data cleaning


options(knitr.kable.NA = '')
knitr::opts_chunk$set(echo = TRUE,
                      warning = FALSE,
                      message = FALSE)

R.version.string

set.seed(8675309)
```

## Simulate Study Data (for Stage 1 RR)

See https://osf.io/87rbg/ for Stage 1 RR code. The code below is
modified from the original to account for a different raw data
structure and to add additional tables and graphs. All analysis code
is identical.

## Load Study Data (for Stage 2 RR)

Load study data and demographic questionnaires from the data folder.

```{r data-load}
session <- read_csv("data/session.csv")
dat_quest <- read_csv("data/quest_data.csv")
dat_exp <- read_csv("data/exp_data.csv")

# reshape questionnaire data to make wide
quest <- dat_quest %>%
  select(session_id, endtime, user_id, q_name, dv) %>%
  group_by(session_id, user_id, q_name) %>%
````
````

````
  arrange(endtime) %>%
  filter(row_number() == 1) %>%
  ungroup() %>%
  spread(q_name, dv, convert = TRUE)
```

Join experiment and questionnaire data

```{r exp-quest-join}
ratings_raw <- dat_exp %>%
  left_join(session, by = c("user_id", "session_id")) %>%
  filter(user_status %in% c("guest", "registered")) %>%
  separate(exp_name, c("psa", "language", "trait", "block"),
           sep = "_") %>%
  select(-psa) %>%
  separate(proj_name, c("psa", "lang", "lab1", "lab2"),
           sep = "_", fill = "right") %>%
  filter(lab1 != "test") %>%
  unite(lab_id, c("lab1", "lab2")) %>%
  select(-psa, lang) %>%
  left_join(quest, by = c("session_id", "user_id")) %>%
  select(language, user_id = session_id, trait,
         stim_id = trial_name,
         order, rt, rating = dv,
         country, sex, age, ethnicity, lab = lab_id, block) %>%
  mutate(trait = recode(trait,
                        "Res" = "responsible",
                        "Wei" = "weird",
                        "Old" = "old",
                        "Tru" = "trustworthy",
                        "Dom" = "dominant",
                        "Emo" = "emostable",
                        "Agg" = "aggressive",
                        "Car" = "caring",
                        "Int" = "intelligent",
                        "Unh" = "unhappy",
                        "Soc" = "sociable",
                        "Mea" = "mean",
                        "Con" = "confident",
                        "Att" = "attractive"
  ))

write_csv(ratings_raw, "data/ratings_raw.csv")
```


## Load Auxillary Data

Data on regions and stimuli.

### Load Region Data

```{r load-region}
regions <- read_csv("data/regions.csv")
````

````
```

### Load Stimulus Info

```{r load-stim-info}
stim_info <- read_csv("data/psa_cfd_faces.csv") %>%
  mutate(ethnicity = recode(Race, "A" = "asian", "B" = "black", "L"
= "latinx", "W" = "white"),
         gender = recode(Gender, "M" = "male", "F" = "female")
  )

stim_info %>%
  group_by(ethnicity, gender) %>%
  summarise(
    n = n(),
    mean_age = round(mean(Age), 2),
    sd_age = round(sd(Age), 2)
  ) %>%
  knitr::kable("html") %>%
  kable_styling("striped")

stim_n_male <- sum(stim_info$gender == "male")
stim_n_female <- sum(stim_info$gender == "female")
mean_age <- mean(stim_info$Age) %>% round(2)
sd_age <- sd(stim_info$Age) %>% round(2)
min_age <- min(stim_info$Age)
max_age <- max(stim_info$Age)
```

Stimuli in our study will be an open-access, full-color, face image
set consisting of `r stim_n_male` men and `r stim_n_female` women
(mean age=`r mean_age` years, SD=`r sd_age` years, range=`r min_age`
to `r max_age` years), taken under standardized photographic
conditions (Ma et al., 2015).

### Load O&T 2008 Data

```{r otdata}
# read original OT data and get into same format as data_agg will be

traits <- ratings_raw %>%
  filter(trait != "old", !is.na(trait)) %>%
  arrange(trait) %>%
  pull(trait) %>%
  unique()

ot_data <- readxl::read_excel("data/
Karolinska_14trait_judgmentswithlabels.xls") %>%
  mutate(region = "(Oosterhof & Todorov, 2008)") %>%
  rename(stim_id = `Todorov Label`,
         emostable = `emotionally stable`) %>%
  select(region, stim_id, all_of(traits))
```
````
````

Data Processing

Join Data

```{r join-data}

ratings <- ratings_raw %>%
  rename(qcountry = country) %>%
  separate(lab, c("country", "lab")) %>%
  left_join(regions, by = "country") %>%
  filter(trait != "old")

```

Graph distributions for trait by region
```{r plot-styles}
# plot styles
bgcolor <- "white"
textcolor <- "black"
PSA_theme <- theme(
    plot.background = element_rect(fill = bgcolor, color = NA),
    panel.background = element_rect(fill = NA, color = "grey"),
    legend.background = element_rect(fill = NA),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    text = element_text(color = textcolor, size=15),
    axis.text = element_text(color = textcolor, size=10),
    strip.text.y = element_text(angle = 0, hjust = 0)
  )
```

```{r trait-by-region-plot, fig.width=15, fig.height=6}

ggplot(ratings, aes(rating, fill = trait)) +
  geom_histogram(binwidth = 1, color = "grey", show.legend = F) +
  facet_grid(region~trait, scales = "free_y") +
  scale_x_continuous(breaks = 1:9) +
  PSA_theme
```

Data checks

```{r data-checks}
part <- ratings %>%
  group_by(user_id, sex, age, country, language, trait, region, lab)
%>%
  summarise(trials = n(),
            stim_n = n_distinct(stim_id)) %>%
  ungroup()
```

How many participants completed at least one rating for each of
120 stimuli

````
Participants rated the 120 stimuli in two blocks. Some participants
quit the study before completing both. Our pre-registered exclusion
criteria require that we have at least one rating for each of the
120 stimuli.

```{r rated-1-stim}
part %>%
  mutate(n120 = ifelse(stim_n == 120, "rated all 120", "rated <
120")) %>%
  count(region, n120) %>%
  spread(n120, n) %>%
  knitr::kable("html") %>%
  kable_styling("striped")
```

### Participants who did not complete exactly 240 trials

In rare cases, participants completed more than 240 trials because
they restarted the study before completion (e.g., when a wifi outage
required a page refresh).

```{r rated-lt-240}
part %>%
  mutate(n240 = case_when(
    trials == 240 ~ "rated 240",
    trials > 240 ~ "rated > 240",
    trials < 120 ~ "rated < 120",
    trials < 240 ~ "rated 120-239"
  )) %>%
  count(region, n240) %>%
  spread(n240, n, fill = 0) %>%
  knitr::kable("html") %>%
  kable_styling("striped")
```

### Participants with low-variance responses in block 1

```{r low-variance}

identical_rating_threshold <- 0.75 * 120 # use this for registered
analyses

inv_participants <- ratings %>%
  filter(block == 1) %>%
  count(user_id, region, trait, rating) %>%
  group_by(user_id, region, trait) %>%
  filter(n == max(n)) %>% # find most common rating for each P
  ungroup() %>%
  filter(n >= identical_rating_threshold) # select Ps who gave the
same rating to >= 75% of stimuli

inv <- inv_participants %>%
  count(region, trait) %>%
````
````

````
  spread(region, n, fill = 0) %>%
  mutate(TOTAL = rowSums(select_if(., is.numeric), na.rm = T))

inv_total <-  group_by(inv) %>%
  summarise_if(is.numeric, sum, na.rm = T) %>%
  mutate(trait = "TOTAL")

bind_rows(inv,inv_total) %>%
  knitr::kable("html") %>%
  kable_styling("striped")
```


### Participants with no region

```{r no-region}
part %>%
  filter(is.na(region)) %>%
  select(user_id, country, lab)
```


### Remove excluded data and average ratings

N.B. Oosterhof & Todorov (2008) additionally standardised the means
of each rater before computing the mean for each face. The resulting
scores are correlated >.99, so we retained our pre-registered
protocol.

```{r data-exclusions}
data <- ratings %>%
  group_by(user_id, trait) %>%
  filter(
    # did not complete 1+ ratings for each of 120 stimuli
    dplyr::n_distinct(stim_id) == 120,
    !is.na(region)   # did not specify region (none expected)
  ) %>%
  anti_join(inv_participants, by = "user_id") %>% # exclude Ps with
low variance
  ungroup() %>%
  group_by(user_id, age, sex, ethnicity, language, lab, country,
region, trait, stim_id) %>%
  summarise(rating = mean(rating)) %>% # average ratings across 2
  ungroup()

write_csv(data, "data/psa001_ind.csv")

```



## Participant Demographics

```{r demog-language}
````

````
data %>%
  group_by(user_id, language) %>%
  summarise() %>%
  ungroup() %>%
  group_by(language) %>%
  summarise(n = n()) %>%
  knitr::kable("html") %>%
  kable_styling("striped")
```

```{r demog-region}
by_region <- data %>%
  group_by(user_id, region) %>%
  summarise() %>%
  ungroup() %>%
  group_by(region) %>%
  summarise(n = n()) %>%
  add_row(region = "TOTAL", n = n_distinct(data$user_id)) %>%
  knitr::kable("html") %>%
  kable_styling("striped")

save_kable(by_region, "figures/n_by_region.html")

by_region
```

### Age and sex distribution per region
```{r age-sex-plot, fig.width=12, fig.height=10}
data %>%
  group_by(user_id, sex, age, region) %>%
  summarise() %>%
  ungroup() %>%
  group_by(region) %>%
  mutate(n = n()) %>%
  ungroup() %>%
  ggplot(aes(as.numeric(age), fill = sex)) +
  geom_histogram(binwidth = 5, color = "grey") +
  geom_text(aes(x=0, y=5, label = paste0("n=",n)), color = "black")
+
  labs(title="", y="", x="Participant age in 5-year bins") +
  facet_grid(region~., scales="free_y") +
  PSA_theme
```


### Participants per trait per region

```{r n-trait-region-plot, fig.width=15, fig.height=10.5}
data %>%
  group_by(trait, region) %>%
  summarise(n = n_distinct(user_id)) %>%
  ggplot(aes(trait, n)) +
  geom_col(aes(fill = trait), show.legend = F) +
  geom_hline(yintercept = 15) +
````

~~~~
  facet_grid(region~., scale = "free") +
  labs(title="", x="", y="Participants per trait per region") +
  theme( axis.text.x = element_text(angle = -45, hjust = 0) ) +
  PSA_theme

ggsave("figures/participants_per_trait_per_region.png", width = 15,
height = 8)
```

### Participants per lab

```{r n-per-lab}
labs <- data %>%
  unite(lab, country, lab) %>%
  group_by(region, lab, user_id) %>%
  summarise() %>%
  ungroup() %>%
  count(region, lab) %>%
  arrange(region, lab)

write_csv(labs, "data/labs_post_exclusions.csv")

knitr::kable(labs) %>%
  kable_styling("striped")
```


# Analyses

## Main Analysis

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.

### Calculate Alphas

```{r calc-alphas, eval = FALSE}

# takes a long time, so saves the results and loads from a file in
the next chunk if set to eval = FALSE
data_alpha <- data %>%
  select(user_id, region, stim_id, rating, trait) %>%
  spread(stim_id, rating, sep = "_") %>%
  group_by(trait, region) %>%
  nest() %>%
  mutate(alpha = map(data, function(d) {
    if (dim(d)[1] > 2) {
      # calculate cronbach's alpha
      subdata <- d %>%
        as_tibble() %>%
~~~~

```
        select(-user_id) %>%
        t()

      capture.output(suppressWarnings(a <- psych::alpha(subdata)))
      a$total["std.alpha"] %>% pluck(1) %>% round(3)
    } else {
      NA
    }
  })) %>%
  select(-data) %>%
  unnest(alpha) %>%
  ungroup()

saveRDS(data_alpha, file = "data/alphas.RDS")
```

```{r alpha-table}

data_alpha <- readRDS("data/alphas.RDS")

n_alpha <- data %>%
  select(user_id, region, trait) %>%
  distinct() %>%
  count(region, trait) %>%
  left_join(data_alpha, by = c("region", "trait")) %>%
  mutate(
    trait = as.factor(trait),
    region = str_replace(region, " (and|&) ", " &\n"),
    region = as.factor(region),
    region = factor(region, levels = rev(levels(region)))
  )

n_alpha %>%
  mutate(stat = paste("α =", alpha, "<br>n =", n)) %>%
  select(Region = region, stat, trait) %>%
  spread(trait, stat) %>%
  knitr::kable("html", escape = FALSE) %>%
  column_spec(2:14, width = "7%") %>%
  kable_styling("striped", font_size = 9) %>%
  save_kable("figures/alpha.html")
```

```{r alpha-plot, fig.width=18, fig.height=10}

ggplot(n_alpha) +
  geom_tile(aes(trait, region, fill=alpha >=.7),
            color = "grey20", show.legend = F) +
  geom_text(aes(trait, region, label=sprintf("α = %0.2f\nn = %.0f",
alpha, n)), color = "black", size = 5) +
  scale_y_discrete(drop=FALSE) +
  scale_x_discrete(position = "top") +
  labs(x="", y="", title="") +
  scale_fill_manual(values = c("white", "red")) +
  PSA_theme
```

````
ggsave("figures/alphas.png", width = 18, height = 10)

```


### Calculate Aggregate Scores

```{r calc-agg-scores}

data_agg <- data %>%
  group_by(region, trait, stim_id) %>%
  summarise(rating = mean(rating)) %>%
  ungroup() %>%
  spread(trait, rating)

```


```{r agg-plot, fig.width=15, fig.height = 8}
data_agg %>%
  gather("trait", "rating", aggressive:weird) %>%
  ggplot(aes(rating, fill = trait)) +
  geom_density(show.legend = F) +
  facet_grid(region~trait) +
  PSA_theme

ggsave("figures/agg_scores.png", width = 15, height = 8)
```

### Principal Component Analysis (PCA)

The number of components to extract was determined using eigenvalues
> 1 for each world region. PCA was conducted using the
`psych::principal()` function with `rotate="none"`.

```{r pca-function}
# function to calculate PCA

psa_pca <- function(d) {
  traits <- select(d, -stim_id) %>%
    select_if(colSums(!is.na(.)) > 0) # omits missing traits

  # principal components analysis (SPSS-style, following Oosterhof &
Todorov)
  ev <- eigen(cor(traits))$values
  nfactors <- sum(ev > 1)

  pca <- principal(
    traits,
    nfactors=nfactors,
    rotate="none"
  )

  stats <- pca$Vaccounted %>%
````

````
    as.data.frame() %>%
    rownames_to_column() %>%
    mutate(type = "stat")

  unclass(pca$loadings) %>%
    as.data.frame() %>%
    rownames_to_column() %>%
    mutate(type = "trait") %>%
    bind_rows(stats) %>%
    gather("pc", "loading", 2:(ncol(.)-1))
}
```

```{r pca}

pca_analyses <- data_agg %>%
  bind_rows(ot_data) %>%
  group_by(region) %>%
  nest() %>%
  mutate(pca = map(data, psa_pca)) %>%
  select(-data) %>%
  unnest(pca) %>%
  ungroup() %>%
  mutate(pc = str_replace(pc, "PC", "Component "))

```

#### Number of Components (and proportion variance) by region
```{r pca-components}
pca_analyses %>%
  filter(rowname == "Proportion Var") %>%
  group_by(region) %>%
  mutate(nPCs = n()) %>%
  ungroup() %>%
  spread(pc, loading) %>%
  select(-rowname, -type) %>%
  mutate_if(is.numeric, round, 3) %>%
  knitr::kable("html") %>%
  kable_styling("striped")
```

#### Trait Loadings by Region and Component

```{r pca-trait-loadings, fig.width=15, fig.height=10}

# order traits by P1 loading if loads positively on P1, or by -P2
loading otherwise
trait_order <- pca_analyses %>%
  filter(region == "(Oosterhof & Todorov, 2008)", type == "trait")
%>%
  spread(pc, loading) %>%
  arrange(ifelse(`Component 1`>0,`Component 1`,-`Component 2`)) %>%
  pull(rowname)
````

```
pca_prop_var <- pca_analyses %>%
  filter(rowname == "Proportion Var") %>%
  select(-rowname, -type) %>%
  mutate(loading = round(loading, 2))

pca_analyses %>%
  filter(type == "trait") %>%
  select(-type) %>%
  mutate(
    trait = as.factor(rowname),
    trait = factor(trait, levels = c(trait_order, "Prop.Var")),
    loading = round(loading, 2)
  ) %>%
  ggplot() +
  geom_tile(aes(pc, trait, fill=loading), show.legend = F) +
  geom_text(aes(pc, trait, label=sprintf("%0.2f", loading)), color =
"black") +
  geom_text(data = pca_prop_var, aes(pc, y = 14,
label=sprintf("%0.2f", loading)), color = "black") +
  scale_y_discrete(drop=FALSE) +
  scale_x_discrete(position = "top") +
  scale_fill_gradient2(low = "dodgerblue", mid = "grey90", high =
"#FF3333", limits=c(-1.1, 1.1)) +
  facet_wrap(~region, scales = "fixed", ncol = 4) +
  labs(x = "", y = "", title="") +
  PSA_theme

ggsave("figures/PCA_loadings.png", width = 15, height = 10)
```

Replication Criteria (PCA)

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 > .7) 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 > .7) 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.

```{r pca-replication-criteria}

pca_rep <- pca_analyses %>%
  filter(
    type == "trait",
    rowname %in% c("trustworthy", "dominant"),
    pc %in% c("Component 1", "Component 2")
  ) %>%
  select(-type) %>%
  mutate(rowname = paste(pc, rowname)) %>%
```

````
  select(-pc) %>%
  spread(rowname, loading) %>%
  rename(Region = region) %>%
  mutate(Replicated = ifelse(
    `Component 1 dominant` < .5 & `Component 1 trustworthy` > .7 &
    `Component 2 dominant` > .7 & `Component 2 trustworthy` < .5,
    "Yes", "No"
  )) %>%
  mutate_if(is.numeric, round, 3) %>%
  knitr::kable("html", col.names = c("Region", "Dominant",
"Trustworthy", "Dominant", "Trustworthy", "Replicated")) %>%
  add_header_above(c(" " = 1, "Component 1" = 2, "Component 2" = 2,
"  " = 1)) %>%
  kable_styling("striped")

save_kable(pca_rep, "figures/PCA_rep_criteria.html")

pca_rep
```

### Factor Congruence (PCA)

This analysis determines the congruence between the components from
Oosterhof & Todorov (2008) and the components in each world region,
using the `psych::factor.congruence` function. Congruence is labeled
"not similar" for values < 0.85, "fairly similar", for values <
0.95, and "equal" for values >= 0.95.

```{r pca-factor-congruence}

# get loadings for original O&T2008
ot2008_pca_loadings <- pca_analyses %>%
  filter(region == "(Oosterhof & Todorov, 2008)", type == "trait")
%>%
  select(-region, -type) %>%
  spread(pc, loading) %>%
  column_to_rownames()

# run factor congruence for each region
fc_pca <- pca_analyses %>%
  filter(type == "trait", region != "(Oosterhof & Todorov, 2008)")
%>%
  select(-type) %>%
  spread(pc, loading) %>%
  group_by(region) %>%
  nest() %>%
  mutate(fc = map(data, function(d) {
    loadings <- d %>%
      as.data.frame() %>%
      select(rowname, `Component 1`, `Component 2`) %>%
      arrange(rowname) %>%
      column_to_rownames()

    psych::factor.congruence(loadings,
````

```
                              ot2008_pca_loadings,
                              digits = 4) %>%
      as.data.frame() %>%
      rownames_to_column(var = "regionPC")
  })) %>%
  select(-data) %>%
  unnest(fc) %>%
  ungroup()

pc_fc_table <- fc_pca %>%
  gather(origPC, congruence, `Component 1`:`Component 2`) %>%
  mutate(sig = case_when(
           congruence < .85 ~ "not similar",
           congruence < .95 ~ "fairly similar",
           congruence >= .95 ~ "equal"
         ),
         congruence = sprintf("%0.3f", congruence)) %>%
  filter(regionPC == origPC) %>%
  select(region, PC = regionPC, congruence, sig) %>%
  gather(k, v, congruence, sig) %>%
  unite(PC, PC, k, remove = T) %>%
  spread(PC, v) %>%
  knitr::kable("html", digits = 3, align = 'lrlrl', escape = F,
               col.names = c("Region", "Congruence", "Conclusion",
"Congruence", "Conclusion")) %>%
  add_header_above(c(" " = 1, "Component 1" = 2, "Component 2" = 2))
%>%
  kable_styling("striped")

save_kable(pc_fc_table, "figures/PCA_factor_congruence.html")

pc_fc_table
```

Robustness Checks

Exploratory Factor Analysis (EFA)

The number of factors to extract was determined using parallel analysis (`paran::paran()`) for each world region. EFA was conducted using the `psych::fa()` function with all default options.

Change to registered analysis:

The analysis for USA & Canada would not converge, so the number of factors was manually reduced from 3 to 4 in the code below.

```{r efa-function}
# function to calculate EFA

psa_efa <- function(d, region) {
  cat("\n", region)

  traits <- select(d, -stim_id) %>%
```

````
    select_if(colSums(!is.na(.)) > 0) # omits missing traits

  # Parallel Analysis with Dino's 'paran' package.
  nfactors <- paran(traits, iterations = 5000,
          centile = 0, quietly = TRUE,
          status = FALSE, all = TRUE,
          cfa = TRUE, graph = FALSE)

  # fix overextraction for USA & Canada
  if (region == "USA & Canada") {
    nfactors$Retained <- 3
  }

  efa <- psych::fa(traits, nfactors$Retained)

  stats <- efa$Vaccounted %>%
    as.data.frame() %>%
    rownames_to_column() %>%
    mutate(type = "stat")

  # NEW: rename MRs in order of Vaccounted
  newnames <- stats %>%
    filter(rowname == "Proportion Var") %>%
    select(-type, -rowname) %>%
    gather(mr, pv) %>%
    arrange(desc(pv)) %>%
    mutate(newname = paste0("MR", row_number()))

  names <- newnames$mr
  names(names) <- newnames$newname

  unclass(efa$loadings) %>%
    as.data.frame() %>%
    rownames_to_column() %>%
    mutate(type = "trait") %>%
    bind_rows(stats) %>%
    rename(!!!names) %>%
    gather("mr", "loading", 2:(ncol(.)-1))
}
```


Calculate for each region

```{r efa, results="hide"}

efa_analyses <- data_agg %>%
  bind_rows(ot_data) %>%
  group_by(region) %>%
  nest() %>%
  mutate(efa = map2(data, region, psa_efa)) %>%
  select(-data) %>%
  unnest(efa) %>%
  ungroup() %>%
````

````
  mutate(mr = str_replace(mr, "MR", "Factor "))
```

#### Number of Factors (and proportion variance) by region

```{r efa-factors}
efa_analyses %>%
  filter(rowname == "Proportion Var") %>%
  group_by(region) %>%
  mutate(nMRs = n()) %>%
  ungroup() %>%
  spread(mr, loading) %>%
  select(-rowname, -type) %>%
  mutate_if(is.numeric, round, 3) %>%
  knitr::kable("html") %>%
  kable_styling("striped")
```

#### Trait Loadings by Region and Factor

```{r efa-trait-loadings, fig.width=15, fig.height=10}

efa_prop_var <- efa_analyses %>%
  filter(rowname == "Proportion Var") %>%
  select(-rowname, -type) %>%
  mutate(loading = round(loading, 2))

efa_analyses %>%
  filter(type == "trait") %>%
  select(-type) %>%
  mutate(
    trait = as.factor(rowname),
    trait = factor(trait, levels = c(trait_order, "Prop.Var")),
    loading = round(loading, 2)
  ) %>%
  ggplot() +
  geom_tile(aes(mr, trait, fill=loading), show.legend = F) +
  geom_text(aes(mr, trait, label=sprintf("%0.2f", loading)), color =
"black") +
  geom_text(data = efa_prop_var, aes(mr, y = 14,
label=sprintf("%0.2f", loading)), color = "black") +
  scale_y_discrete(drop=FALSE) +
  scale_x_discrete(position = "top") +
  scale_fill_gradient2(low = "dodgerblue", mid = "grey90", high =
"#FF3333", limits=c(-1.1, 1.1)) +
  facet_wrap(~region, scales = "fixed", ncol = 4) +
  labs(x = "", y = "", title="") +
  PSA_theme

ggsave("figures/EFA_loadings.png", width = 15, height = 10)
```
````
````

Replication Criteria (EFA)

Oosterhof and Todorov’s valence-dominance model will be judged to have been replicated in a given world region if the the first factor is correlated strongly (loading > .7) with trustworthiness and weakly (loading < .5) with dominance, and the second factor is correlated strongly (loading > .7) with dominance and weakly (loading < .5) with trustworthiness. All these criteria need to be met to conclude that the model was replicated in a given world region.

````
```{r efa-replication-criteria}

efa_rep <- efa_analyses %>%
  filter(
    type == "trait",
    rowname %in% c("trustworthy", "dominant"),
    mr %in% c("Factor 1", "Factor 2")
  ) %>%
  select(-type) %>%
  mutate(rowname = paste(mr, rowname)) %>%
  select(-mr) %>%
  spread(rowname, loading) %>%
  rename(Region = region) %>%
  mutate(Replicated = ifelse(
    `Factor 1 dominant` < .5 & `Factor 1 trustworthy` > .7 &
    `Factor 2 dominant` > .7 & `Factor 2 trustworthy` < .5,
    "Yes", "No"
  )) %>%
  select(Region, `Factor 1 trustworthy`, `Factor 1 dominant`,
         `Factor 2 dominant`, `Factor 2 trustworthy`, Replicated)
%>%
  mutate_if(is.numeric, round, 3) %>%
  knitr::kable("html", col.names = c("Region", "Trustworthy",
"Dominant", "Dominant", "Trustworthy", "Replicated")) %>%
  add_header_above(c(" " = 1, "Factor 1" = 2, "Factor 2" = 2, "  " =
1)) %>%
  kable_styling("striped")

save_kable(efa_rep, "figures/EFA_rep_criteria.html")

efa_rep

```
````

Factor Congruence (EFA)

This analysis determines the congruence between the factors from Oosterhof & Todorov (2008) and the factors in each world region, using the `psych::factor.congruence` function. Congruence is labeled "not similar" for values < 0.85, "fairly similar", for values < 0.09, and "equal" for values >= 0.95.

````
```{r efa-factor-congruence}
````
````

```
# get loadings for original O&T2008
ot2008_efa_loadings <- efa_analyses %>%
  filter(region == "(Oosterhof & Todorov, 2008)", type == "trait")
%>%
  select(-region, -type) %>%
  spread(mr, loading) %>%
  column_to_rownames()

# run factor congruence for each region
fc_efa <- efa_analyses %>%
  filter(type == "trait", region != "(Oosterhof & Todorov, 2008)")
%>%
  select(-type) %>%
  spread(mr, loading) %>%
  group_by(region) %>%
  nest() %>%
  mutate(fc = map(data, function(d) {
    loadings <- d %>%
      as.data.frame() %>%
      select(rowname, `Factor 1`, `Factor 2`) %>%
      arrange(rowname) %>%
      column_to_rownames()

    psych::factor.congruence(loadings,
                             ot2008_efa_loadings,
                             digits = 4) %>%
  as.data.frame() %>%
  rownames_to_column(var = "regionMR")
  })) %>%
  select(-data) %>%
  unnest(fc) %>%
  ungroup()

mr_fc_table <- fc_efa %>%
  gather(origMR, congruence, `Factor 1`:`Factor 2`) %>%
  mutate(sig = case_when(
           congruence < .85 ~ "not similar",
           congruence < .95 ~ "fairly similar",
           congruence >= .95 ~ "equal"
         ),
         congruence = sprintf("%0.3f", congruence)) %>%
  filter(regionMR == origMR) %>%
  select(region, MR = regionMR, congruence, sig) %>%
  gather(k, v, congruence, sig) %>%
  unite(MR, MR, k, remove = T) %>%
  spread(MR, v) %>%
  knitr::kable("html", digits = 3, align = 'lrlrl',
               col.names = c("Region", "Congruence", "Conclusion",
"Congruence", "Conclusion")) %>%
  add_header_above(c(" " = 1, "Factor 1" = 2, "Factor 2" = 2)) %>%
  kable_styling("striped")
```

~~~
save_kable(mr_fc_table, "figures/EFA_factor_congruence.html")

mr_fc_table
```
~~~

Replication Criteria for "best" factor (EFA)

Run replication criteria on the two factors that load highest on dominance and trustworthiness, regardless of factor position. Here, "replicated" means that the factor that loads highest on trustworthiness loads has a loading > 0.7 for trustworthy and < .5 for dominant, while the factor that loads highest on dominance loads has a loading > 0.7 for dominant and < .5 for trustworthy.

~~~
```{r efa-replication-criteria-2}

trust <- efa_analyses %>%
  filter(type == "trait", rowname == "trustworthy") %>%
  group_by(region) %>%
  filter(loading == max(loading)) %>%
  ungroup() %>%
  select(region, mr) %>%
  mutate(factor = "trust")

dom <- efa_analyses %>%
  filter(type == "trait", rowname == "dominant") %>%
  group_by(region) %>%
  filter(loading == max(loading)) %>%
  ungroup() %>%
  select(region, mr) %>%
  mutate(factor = "dom")

td <- bind_rows(trust, dom) %>%
  mutate(mr = str_replace(mr, "Factor ", "")) %>%
  spread(factor, mr) %>%
  select(region, dom, trust)

efa_rep2 <- efa_analyses %>%
  left_join(bind_rows(trust, dom), by = c("region", "mr")) %>%
  filter(
    type == "trait",
    !is.na(factor),
    rowname %in% c("trustworthy", "dominant")
  ) %>%
  select(-type) %>%
  mutate(rowname = paste0(factor, " Factor, ", rowname)) %>%
  select(-mr, -factor) %>%
  spread(rowname, loading) %>%
  left_join(td, by = "region") %>%
  rename(Region = region) %>%
  mutate(Replicated = ifelse(
    `trust Factor, dominant` < .5 & `trust Factor, trustworthy` > .7
&
~~~

````
    `dom Factor, dominant` > .7 & `dom Factor, trustworthy` < .5,
    "Yes", "No"
  )) %>%
  mutate_if(is.numeric, round, 3) %>%
  select(Region,
         trust,
         `trust Factor, trustworthy`,
         `trust Factor, dominant`,
         dom,
         `dom Factor, dominant`,
         `dom Factor, trustworthy`,
         Replicated) %>%
  knitr::kable("html", col.names = c("Region",
                                     "#", "Trustworthy", "Dominant",
                                     "#", "Dominant", "Trustworthy",
                                     "Replicated")) %>%
  add_header_above(c(" " = 1, "Trust Factor" = 3, "Dominance Factor"
= 3, "  " = 1)) %>%
  kable_styling("striped")

save_kable(efa_rep2, "figures/EFA_rep_criteria2.html")

efa_rep2

```
````
````