Culture shapes our understanding of others' thoughts and emotions: An investigation across 12 countries

François Quesque, Antoine Coutrot, Leonardo Cruz de Souza, Sandra Baez, Cardona Juan Felipe, Alejandra Neely-Prado, Clarens Maria Florencia, Catalina Trujillo, Grisales Johan Sebastián, Sol Fittipaldi, et al.

Culture shapes our understanding of others' thoughts and emotions: An investigation across 12 countries

Authors

Quesque François[1], Coutrot Antoine[2], Cox Sharon[3], de Souza Leonardo Cruz[4], Baez Sandra[5], Cardona Juan Felipe[6], Mulet-Perreault Hannah[7], Flanagan Emma[8, 9], Neely-Prado Alejandra[10], Clarens Maria Florencia[11], Cassimiro Luciana[12], Musa Gada[13], Kemp Jennifer[14], Botzung Anne[14], Philippi Nathalie[14], Cosseddu Maura[15], Trujillo Catalina[6], Grisales Johan Sebastián[6], Fittipaldi Sol[16, 17], Magrath Guimet Nahuel[11], Calandri Ismael Luis[11], Crivelli Lucia[11], Sedeno Lucas[17], Garcia Adolfo M[16, 17, 18, 19], Moreno Fermin[20], Indakoetxea Begoña[20], Benussi Alberto[21], Brandão Moura Millena Vieira[4], Santamaria-Garcia Hernando [22], Matallana Diana [22], Prianishnikova Galina[23], Morozova Anna[23], Iakovleva Olga[23], Veryugina Nadezda[23], Levin Oleg[23], Zhao Lina[24], Liang Junhua[24], Duning Thomas[25], Lebouvier Thibaud[1], Pasquier Florence[1], Huepe David[10], Barandiaran Myriam[20], Johnen Andreas[25], Lyashenko Elena[23], Allegri Ricardo F[11], Borroni Barbara[21], Blanc Frederic[14], Wang Fen[24], Yassuda Monica Sanches[12], Lillo Patricia[13], Teixeira Antônio Lúcio[4], Caramelli Paulo[4], Hudon Carol[7], Slachevsky Andrea[26, 27, 28, 29], Ibanez Agustin[10, 16, 17, 19, 30*], Hornberger Michael[8, 9*], Bertoux Maxime[1, 9].

Affiliations

[1] Univ. Lille, Inserm, CHU Lille, U1172 - LilNCog - Lille Neuroscience & Cognition, F-59000 Lille, France - LiCEND

[2] Université de Nantes, CNRS, LS2N, F-44000 Nantes, France

[3] Department of Behavioural Science and Health, Institute of Epidemiology and Healthcare, University College London, London, UK

[4] University of Belo Horizonte, Brazil.

[5] Universidad de los Andes, Bogotá, Colombia

[6] Instituto de Psicología, Universidad del Valle, Cali, Colombia

[7] Université Laval and CERVO Brain Research Centre, Québec, Canada

[8] Norwich Medical School, University of East Anglia, UK

[9] Department of Clinical Neurosciences, University of Cambridge, UK

[10] Center for Social and Cognitive Neuroscience, School of Psychology, Adolfo Ibáñez University, Santiago, Chile

[11] FLENI Fondation, Department of Neurology, Buenos Aires, Argentina

[12] School of Arts, Sciences and Humanities, University of São Paulo, Department of Neurology, São Paulo, Brazil

[13] Universidad de Chile, Santiago, Chile

[14] University of Strasbourg, Strasbourg, France

[15] Neurology Unit, Spedali Civili Hospital, Brescia, Italy

[16] Universidad de San Andrés, Buenos Aires, Argentina

[17] National Scientific and Technical Research Council (CONICET), Argentina

[18] Departamento de Lingüística y Literatura, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile.

[19] Global Brain Health Institute (GBHI), University of California-San Francisco (UCSF), San Francisco, California, United States.

[20] Department of Neurology, Unit of Cognitive Disorders, Hospital Universitario Donostia, San Sebastian, Spain

[21] Centre for Neurodegenerative Disorders, Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy

[22] School of Medicine, Neuroscience Doctorate. Aging Institute, Physiology and Psychiatry Department. Pontificia Universidad Javeriana, Bogotá, Colombia

[23] Central Clinic No 1 of the Ministry of Internal Affairs, Moskva, Russia

[24] Innovation center for neurological disorders, Department of Neurology, Xuan Wu Hospital, Capital Medical University, 45 Changchun Street, Beijing

[25] Clinic of Neurology with Institute for Translational Neurology, University Hospital Münster, Münster, Germany

[26] Geroscience Center for Brain Health and Metabolism (GERO), Faculty of Medicine, University of Chile, Santiago, Chile.

[27] Neuropsychology and Clinical Neuroscience Laboratory (LANNEC), Physiopathology Department - ICBM, Neurocience and East Neuroscience Departments, Faculty of Medicine, University of Chile, Santiago, Chile.

[28] Memory and Neuropsychiatric Clinic (CMYN) Neurology Department, Hospital del Salvador and Faculty of Medicine, University of Chile, Santiago, Chile.

[29] Servicio de Neurología, Departamento de Medicina, Clínica Alemana-Universidad del Desarrollo, Santiago, Chile.

[30] Universidad Autónoma del Caribe, Barranquilla, Colombia.

* These authors contributed equally

Corresponding author

Maxime Bertoux: maxime.bertoux@inserm.fr

Abstract

Humans have developed specific abilities to interact efficiently with their conspecifics (social cognition). Despite abundant behavioral and neuroscientific research, the influence of cultural factors on these skills remains poorly understood. This issue is of particular importance as most cognitive tasks are developed in highly specific contexts, not representative of what is encountered by the world's population. Through a large international and multi-site study, we assessed core social cognition aspects using current gold-standard tasks in 587 participants from 12 countries. Age, gender, and education were found to impact emotion recognition as well as the ability to infer mental states. After controlling for these factors, differences between countries accounted for more than 20% of the variance on both abilities. Importantly, it was possible to isolate cultural from linguistic impacts, which classically constitute a major limitation. We suggest important methodological shifts to better represent social cognition at the fundamental and the clinical levels.

Keywords: Social cognition, Emotion recognition, Theory of mind, Cross-Cultural, Gender differences

Humans are a highly social species, characterized by a unique level of cooperation among exceptionally large and genetically heterogeneous groups. Whether our actions are directed toward a social goal or simply involve a social context, there is a social dimension underlying almost every aspect of our lives. Among the different cognitive processes traditionally described as main domains of human cognition, “social cognition” allows us to navigate the social world. Roughly referring to the cognitive abilities supporting efficient interpersonal interactions through accurate social information processing, this domain gained attention within neuroscience where a whole subfield is now aiming to understand “the social brain” (Adolphs, 2009). While the influence of social factors on cognition, such as gender or education, have been considered before, how culture impacts cognitive functioning and individual behaviors remains largely unknown. The impact of cultural dimensions (here referring to a set of shared knowledge, representations and concepts within a country), on neurocognitive functioning has indeed been traditionally ignored by neuroscience (Focus on social neuroscience, 2012). Many cognitive tasks have been validated to quantify human social cognitive and affective dimensions, however, the ability of these tests to be suitable worldwide, without any prior cultural adaptation performed, could be legitimately questioned. Cognitive tests are indeed most often developed in Western, Industrialized and Democratic countries and their norms based on well-Educated and Rich people (referred to as “W.E.I.R.D.” people, Henrich, Heine & Norenzayan, 2010a, 2010b). Yet, it is becoming increasingly clear that cognitive measures no longer have acceptable accuracy when used with individuals from populations that do not fit these specific cultural characteristics (e.g. Boone, Victor, Wen, Razani, & Pontón, 2007; Daugherty, Puente, Fasfous, Hidalgo-Ruzzante & Pérez-Garcia, 2017; Puente et al., 2013). This is far from being a trivial issue, as WEIRD participants represent only 12% of the world’s population (Arnett, 2008), but their assumed representativeness to the human population has only recently been questioned. This narrow sampling of the world's population constitutes a serious threat on the theoretical level, as a conclusion drawn from WEIRD samples in neuroscience studies should not be applied *in extenso* to non-WEIRD populations (Matsumoto et al., 2016). As neuroscience also informs psychological and medical practices, this bias could also have devastating consequences on one's education or health guidance, including misdiagnosis and potentially inadequate treatment prescription. As a striking example, it has been reported that almost half of healthy black Americans were erroneously classified as cognitively impaired on one of the most commonly used memory tests when using the original standardized scores predominately based on data from white individuals (Norman, Ewan, Miller & Heaton, 2000). Such a "universalist" approach should thus be abandoned.

To date, cultural influences on classical neurocognitive functions have been investigated by comparing populations originating from two or sometimes three countries (e.g. Cohn & Statucka, 2019; Daugherty et al., 2017). For each country involved, participants were generally recruited from a single site, which might have generated important confounds (Talhelm, et al., 2014). Moreover, samples were generally composed of young and well-educated people, which constitutes an additional limitation to the generalization of findings. Since cultural differences have been observed in memory or attention tests (Hayden et al., 2014), and recently in spatial navigation across a wide range of countries (Coutrot et al., 2018), one could expect higher variations in tests addressing social cognition, which involve culture-dependent concepts or rules. However, research in social cognition has been historically neglected due to the caricature of "social" as involuntary, irrational and regimented, established by early social psychologists, as well as the apparent threat to liberalism that was posed by socially engaged forms of cognition (Greenwood, 2004). Considering this historical neglect, we believe that there still is a general underestimation of the importance of social cognition and its determinants, in both research and clinical settings.

As deficits in social cognition lead to varied interpersonal difficulties that have been recognized as more incapacitating than traditionally assessed cognitive deficits (Henry et al., 2016), deeply impacting both the patient's (Santamaría-García et al., 2020) and their relatives' quality of life (Spitzer et al., 2019), there is an urgent challenge to explore and quantify the possible cultural variations that could be at stake in this domain. This is especially true for the ability to recognize emotions from faces and to infer other's mental states, which are the core abilities widely explored by validated tests of social cognition (Baron-Cohen et al., 2001; Stone et al., 1998; Ekman & Friesen, 1976).

Some past studies, mostly originating from social psychology, did investigate cultural influences on social cognition (e.g. Awad et al., 2018; Norenzayan et al., 2002), but the methodological approach relied on quick online surveys, and not on a formal (i.e. performance-based) cognitive assessment. This again limits the investigations to opinions, beliefs, and representations, and does not provide reliable findings about cultural effects on cognitive functioning. As a consequence, appeals for international large samples and multi-site approaches have been conveyed over the last years (Barrett, 2020), but answers to these calls have been limited so far, given the complex coordination of efforts required for cross-cultural neuroscientific studies.

Anticipating these recommendations, we developed an international network devoted to the study of social cognition from a clinical perspective (International Network on Social Cognition Disorders, *https://www.scann.fr/inscd*) six years ago. Through this network, we conducted a large international study to quantify cultural variations in social cognitive performance-based tests. Specifically, we investigated cultural variations across 12 countries in the performance on two widely used tests that investigate key social cognition functions: facial emotion recognition and mental state inferences. The "Ekman's faces" from the Picture of Facial Affect Test (Ekman & Friesen, 1976) is the gold-standard to evaluate the ability to recognize others' emotion from facial information (>5280 citations on Google Scholar, May 2020). Similarly, the recognition of "faux pas" (Baron-Cohen et al., 1999; Stone et al., 1998) also constitutes one of the most used tests to assess one's ability to process social information and represent others' thought (>2960 cumulated citations on Google Scholar). On the fundamental level, this study aimed to shed light on the social factors, indivisible from human nature, underlying the variability of social cognition abilities. On a more empirical level, with the increasing development of international projects relying on shared clinical data (e.g. Human Brain Project, ENIGMA Network), it is critical to determine whether test performance is influenced by cultural characteristics or whether outcomes can be compared unequivocally across countries. We hypothesized that important cultural variations would emerge in both tests, alongside a significant influence of age, gender, and education on social cognitive performance.

Results

Data were collected from 587 control participants (339 women, mean age = 58.04, SD = 16.06, mean education = 10.5 years, SD = 5.46), originating from 20 centers across 12 countries through the IN-SCD. Participants' social cognition performance was evaluated through two tests. Their emotion recognition ability was quantified by the Facial Emotion Recognition (FER) score, and their mental state inference ability was quantified by the modified faux pas (mFP) score (see Methods for a full description). Specifically, for emotion recognition abilities, participants were presented with emotional faces and had to choose among seven labels, of which one matched the emotion expressed, for each face. For mental state inference, participants had to read short stories depicting a social interaction and in which one character committed or did not commit a faux pas (i.e. involuntarily said something rude that they would not have said if they had access to the other character's knowledge). This test requires the participant to detect the presence or absence of faux pas, to explain it and to specify the intention, knowledge, and feeling of the characters.

We fit a linear mixed model for each score with age, gender and education as fixed effect, and country as random effect: score ~ age + gender + education + (1 | country). In these models, we controlled for age, gender and education while allowing the intercept to vary by country.

Effect of age, gender, and education

***Facial emotion recognition* -** Independently of the origin of participants, we found that age had a negative effect on the facial emotion recognition score ($t(566)=-6.18$, $p<0.001$), with a monotonic decline over the lifespan. Figure 1A shows the decline of facial emotion recognition score with age, both for men and women. Gender also had an effect on the facial emotion recognition score, with women having higher scores than men ($t(566)=-4.90$, $p<0.001$). We computed the effect size of gender with Hedge's g, g=0.23, 95%CI=[0.06, 0.40], positive values corresponding to an advantage for women. On average, women scored as their 10 years younger men counterparts. We did not however find an effect of education ($t(566)=0.91$, $p=0.36$).

***Mental state inference* -** As with emotion recognition, age had a negative effect on the modified faux pas score ($t(537)=-2.01$, $p=0.03$). We found a small effect of gender ($t(537)=-1.98$, $p=0.048$), but the Hedge's g effect size is almost null: g=0.03, 95%IC=[-0.14 0.20], positive values corresponding to an advantage for women. Contrarily to emotion recognition, we found a positive effect of education on the modified faux pas score ($t(537)=4.00$, $p<0.001$). Figure 1B shows an increase of faux pas detection score with education, both for men and women. We computed the effect size of education with Hedge's g, comparing participants with less than the median education duration (14 years, N=271) to participants with the median education duration or more (N=313), g=0.35, 95%CI=[0.17, 0.51], with positive values corresponding to an advantage for more educated participants.

Intermediary conclusions – We found a continuous decline with age of both facial emotion recognition and mental state inference abilities. Women outperformed men at both tasks. The gender difference was more pronounced for the emotion recognition than for the mental state inference task. Finally, education largely influenced mental state inference only.

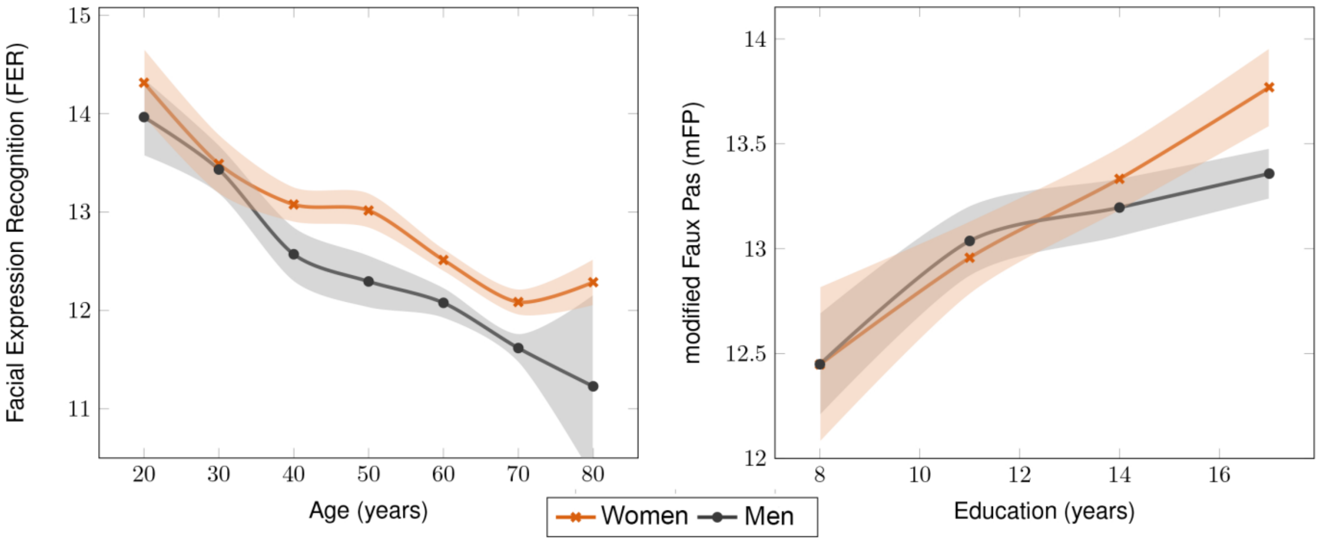

Figure 1. (a) Effect of age and gender on facial emotion recognition score (FER). FER is averaged within 10-year windows. (b) Effect of education and gender on the modified faux pas score (mFP). mFP is averaged within 3-year windows for education. Error bars correspond to standard errors

Effect of culture

To assess the culture effect, we compared the linear mixed model described above to a model only including age, gender and education. The likelihood ratio test statistic is LRStat(1) = 118.53, p<0.001 for mFP, LRStat(1) = 119.01, p<0.001 for FER, indicating that the scores are better predicted when the country random effect was included in the model.

Variation of facial emotion recognition across countries - The variance partition coefficient (VPC) indicates that 20.76% (95%CI = [8.26, 35.69]) of the variance in emotion recognition scores can be attributed to differences among nationalities after controlling for age, gender and education. Figure 2A represents the countries ranked according to their conditional modes, i.e. the difference between the global average predicted response in score and the response predicted for a particular country. We also computed Hedges' g between every pair of countries, see heatmap Figure 2B. Unlike conditional modes, Hedge's g are not corrected for the potential effect of other variables, but they have the advantage of being straightforward and to allow a direct comparison between studies. To mitigate the effect of age (which had the biggest effect on FER scores) in Hedge's g calculations, only participants above 50 years old were considered for this analysis (N=432, i.e. 74% of the participants). Hedge's g between the first and last countries in emotion recognition score (England vs Italy), was g = 1.89, 95%CI = [0.81 2.95]. Since gender had a significant effect on the facial emotion recognition score, we computed the gender effect size for each country (Figure 2c). It went from g=1.17, 95%CI = [0.37, 1.97] in Colombia to g = -0.22, 95%CI = [-1.04, 0.60] in Russia, with positive values corresponding to higher scores for women.

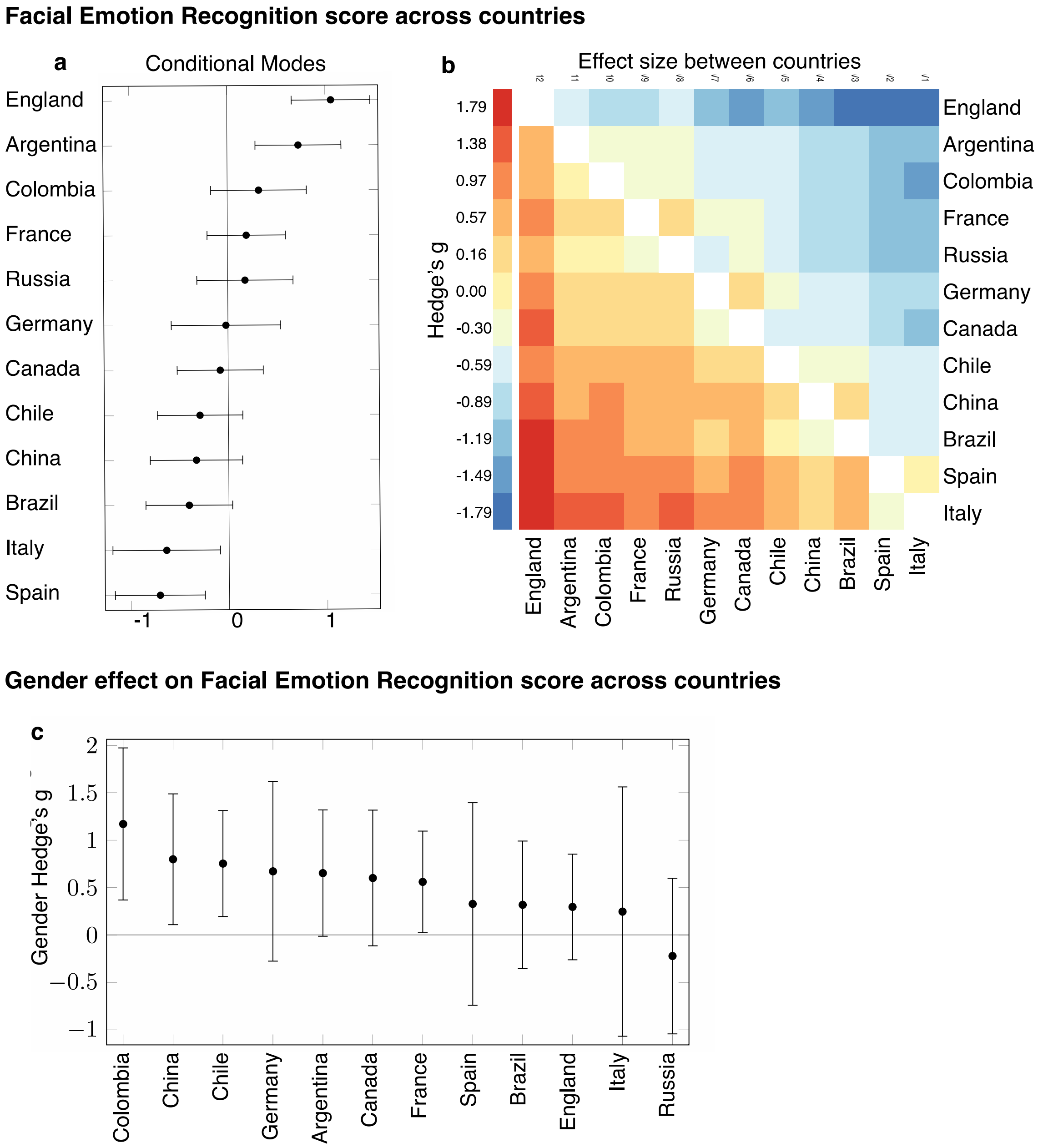

Figure 2 – Facial emotion recognition across countries. (a) Country conditional modes from a linear mixed model controling for age, gender and education. Higher conditional modes correspond to better performances. (b) Heatmap of the difference between every pair of countries measured with Hedge's g. Positive values correspond to a better performance of column countries compared to row countries. (c) Gender effect size in each country, measured with Hedge's g. Positive values correspond to women scoring higher than men. Error bars correpond to 95% confidence intervals.

Figure 3 represents the facial emotion recognition confusion tables for each country. The overall structure was similar across countries, but a number of differences can still be identified. For instance, fear was often misclassified as surprise, but the error rate widely varied across countries, from 25% in Germany to 50% in Canada. Italians misclassified sadness as neutral 21% of the time, but it never happened in China. We report country VPC with 95% confidence intervals for each emotion in Supplemental Material 1.

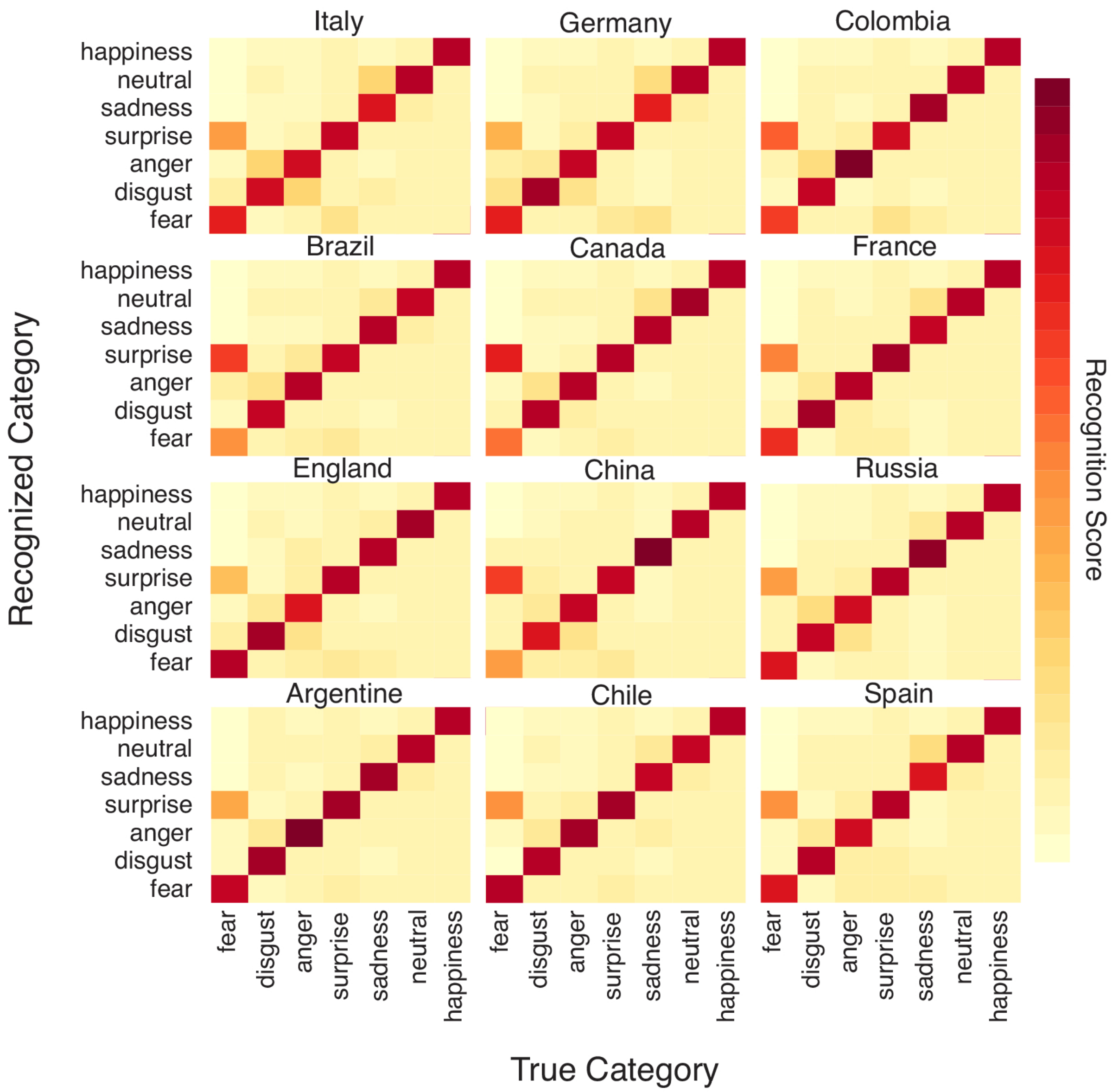

Figure 3 – Confusion tables for emotion recognition in each country. Hotter colors correspond to higher recognition scores.

Variation of mental state inference across countries - The variance partition coefficient (VPC) indicates that 24.52% (95%CI = [9.10, 41.32]) of the variance in the modified faux pas scores can be attributed to differences between nationalities after controlling for age, gender and education. Figure 4a represents the countries ranked according to their conditional modes. As above, we also computed Hedges' g between every pair of countries, see heatmap Figure

4b. Hedge's g between the first and last countries for the faux pas detection score (England vs Spain) was g = 2.17, 95%CI = [1.35 2.97]. We computed the gender effect size for each country (Figure 4c). It went from g=1.42, 95%CI = [0.51, 2.33] in Russia to g = -0.40, 95%CI = [-1.07, 0.28] in China. We also computed the education effect size for each country (Figure 4d, participants with less than the median education duration vs participants with the median education duration or more). It went from g=1.50, 95%CI = [0.35, 2.62] in Russia to g = -0.05, 95%CI = [-0.68, 0.57] in Argentina, with positive values corresponding to a positive effect of education.

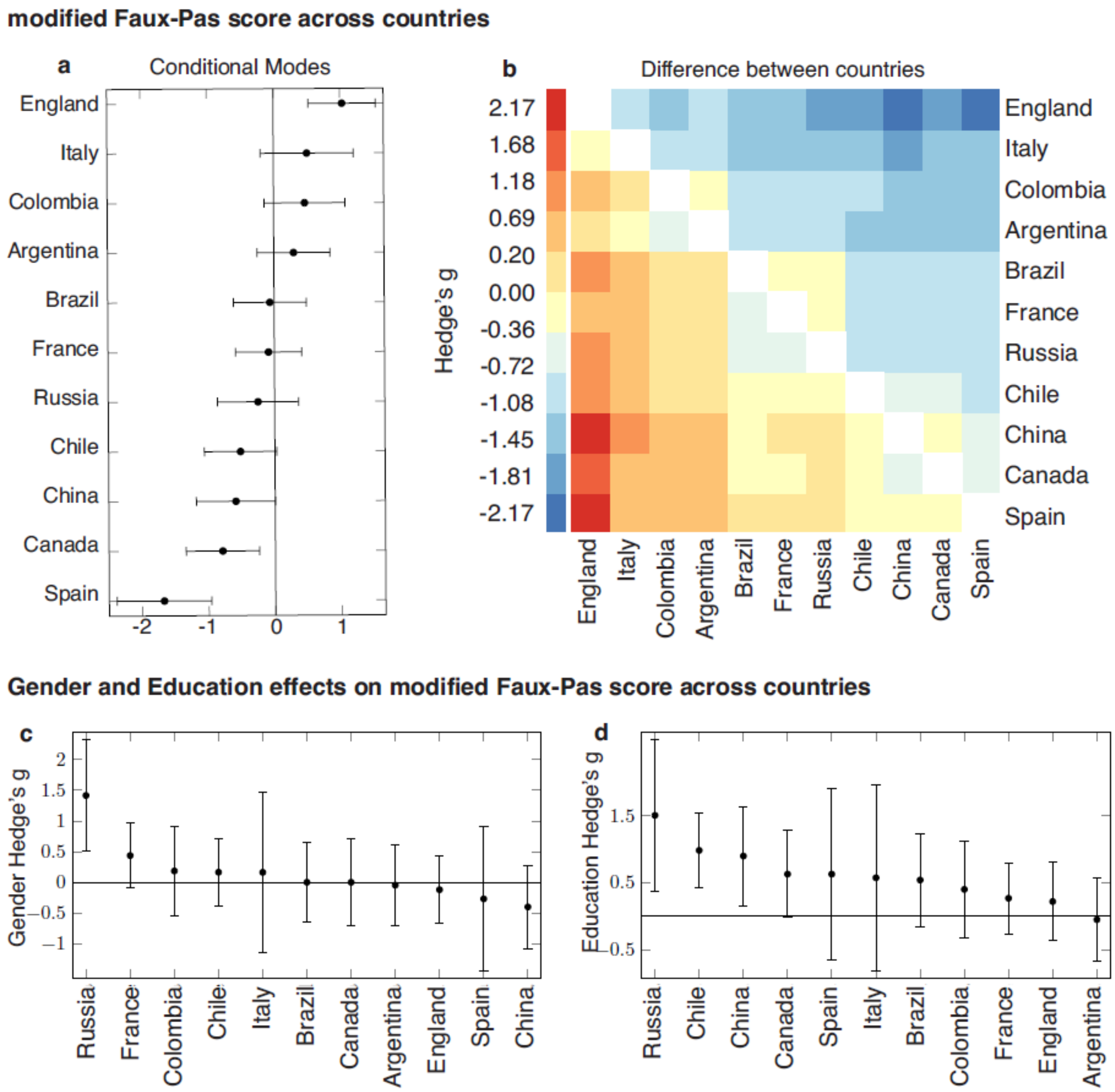

Figure 4 – Mental state inference across countries. (a) Country conditional modes from a linear mixed model controling for age, gender and education. Higher conditional modes correspond to better performances. (b) Heatmap of the difference between every pair of countries measured with Hedge's g. Positive values correspond to a better performance of column countries compared to row countries. (c) Gender effect size in each country, measured with Hedge's g. Positive values correspond to women scoring higher than men. (d) Education effect size in each country, measured with Hedge's g. Positive values correspond to a positive effect of education. Error bars correpond to 95% confidence intervals.

We separated the stories that actually contained a faux pas from the stories that did not, and computed the same LMM as above on the scores resulting from these two groups of stimuli. We found that the country VPC was higher for the stories with a faux pas (VPC_{FP} = 23.8%) than for the stories without (VPC_{NFP} = 11.6%), indicating that the cultural effect is stronger when there actually is a faux pas to identify.

Potential confounding factors – We first ruled out a potential effect of translation. Indeed, the original tests were developed in English, and their translation to other languages could have had an influence on performance. To test this, we added to the linear mixed model previously described "language" as a fixed effect: score ~ age + gender + education + language + (1 | country). Language was a categorical variable corresponding to the language the test was translated into for each country. Interestingly, language did not have an effect on facial emotion recognition performances ($t(559) = 0.93$, $p = 0.54$) nor on faux pas detection performances ($t(531) = 0.74$, $p = 0.77$).

We also compared the variation in scores between countries to the variation in scores between the different centers within countries. We added to our initial linear mixed model a random effect for centers nested within countries: score ~ age + gender + education + (1 | country) + (1 | country:center). For facial emotion recognition, $VPC_{country}$ = 11.70% and $VPC_{country:center}$ = 10.33%. For faux pas detection, $VPC_{country}$ = 16.29% and $VPC_{country:center}$ = 9.00%. This shows that while there is some variance between centers within countries, it is smaller than the variance between countries.

At this stage, we demonstrated that participants ability to recognize emotion from faces and to infer the mental state of others, as measured by current gold-standard respective tests, largely vary across countries. Interestingly, this influence seems to be independent of the language of participants.

Discussion

This study demonstrated that independent demographical and cultural factors significantly impact the ability to recognize emotions and subtle social reasoning (encompassing other's mental state inferences). More specifically, we found a continuous decline across the adult lifespan on both cognitive abilities. Women had better performance than men at both tasks used here, but gender differences were more pronounced for the emotion recognition than for the mental states' inference. We also found an effect of education, which largely influenced mental states inference performance, but not emotion recognition. Finally, we observed important cultural variations on both tests, revealing that the ability to recognize emotion and to infer other's mental states as measured by the current gold-standard tests, drastically vary across countries.

The impact of age retrieved here is congruent with previous cross-sectional studies in which older adults had consistently shown lower accuracy compared with younger adults when labeling facial expressions of emotion (Ruffman, Halberstadt, & Murray, 2009; Ruffman et al., 2008), but also when having to detect social faux pas (Halberstadt, Ruffman, Murray, Taumoepeau, & Ryan, 2011; Wang & Su, 2006). Interestingly, in our study, these effects were independent of the participants' country, suggesting a universal decline of social cognitive abilities with age.

Regarding gender influences, two contradictory views currently co-exist (Lausen & Schacht, 2018). The "gender difference hypothesis" suggests that women have better socio-emotional decoding skills than men, even if this advantage is relatively small according to meta-analyses (Kirkland, Peterson, Baker, Miller & Pulos, 2013; Kret & De Gelder, 2012). Conversely, the "*gender similarity hypothesis"* suggests that most of the time, gender differences only reflect experimental artefacts and concludes that in many cases women and men are rather similar on most psychological dimensions (Baez et al., 2017; Hyde, 2014). In this context, our study, based on a large sample and relying on gold-standards tasks of social cognition, indicates that a gender effect on emotion recognition and social reasoning can be found across a wide variety of cultures (see also Merten, 2005 for similar findings). This gender effect seems moreover consistent over the lifespan. At first sight, it might be tempting to interpret women's systematic advantage as an argument in favor of biological predispositions. This is, however, contradicted by the major variations of the magnitude of gender differences across countries. As previously underlined (Eisenberg, Cumberland & Spinrad, 1998), the meanings of emotion expression and appropriateness of behaviors are largely shaped by socialization since early childhood, through parental reactions and expectations that are gender-differentiated (Denham, Bassett & Wyatt, 2007). As women are expected to be good social decoders (Graham & Ickes, 1997), it is probable that their performances not only reflected their abilities but also some higher motivational aspects (Ickes, Gesn & Graham, 2000) along with the compliance to cultural gender roles (Baez et al., 2017).

The level of formal education was also found to be related to the participants' performance on the modified faux pas task. Interestingly, the advantage for individuals with higher education seems to be common to a wide variety of cultures. To our knowledge, we are the first to report such a finding as developmental studies focusing on mental states inferences classically focus on childhood and are conducted within a single country (e.g. Filippova & Astington, 2008).

The most striking findings of our study are the cultural variations observed after controlling for age, gender, and education. The proportion of variance explained by cultural differences in our population was indeed considerable, reaching 20.76% on emotion recognition scores, and 24.52% on the modified faux pas scores. This is almost ten times higher than the effects reported in previous cross-cultural studies focusing on memory and attention (Hayden et al., 2014) or on spatial navigation abilities (Coutrot et al., 2018). This underlines that, more than other neurocognitive tasks, social stimuli - often relying on local norms or utilizing words with multiple meanings - are highly dependent on culture. The influence of culture was also higher than the influence of gender in both tasks. In the present study, these variations cannot be explained by translation issues as both between-countries analyses controlling for the languages spoken and within-countries multi-site comparisons indicated that language did not have an effect on the tests' performance. This specific result is of particular significance as a major limitation in previous studies was based on the fact that it was not possible to isolate cultural variations from linguistic differences in comparisons over 2 or 3 countries, although social-cognitive abilities are extensively tangled with language pragmatics (Elfenbein & Ambady, 2002; Fiedler, 2008). Moreover, the variations observed in participants' performance between-centers within a single country were smaller than those observed for between-countries variations, suggesting a consistency of our observations within a given culture. Future studies with higher samples will have to deepen the cultural differences reported here, as the total amount of variables (n=90) recorded in our study - which would allow for finer-grained analyses - exceeded the number of participants in most centres. This limit is why we focused on the tests' total score. We however checked the typical patterns of responses to allow a more qualitative reading of cultural differences.

Regarding emotion recognition, cross-country comparisons revealed that some emotions were consensually identified in all countries, while for others, cultural variations existed. Specifically, happy expressions were the easiest to recognize and therefore unambiguously categorized in every cultures (see Nelson & Russel, 2013 for congruent results). By contrast, recognition of fear or disgust items appeared to be more culture-dependent, as the proportions of correct identification significantly varied from a culture to another. Within-countries clasification patterns revealed that the majority of Brazilian and Canadian participants confused “fear” and “surprise” faces, a bias that was almost absent among English and Argentinian participants. About this specific result, it is consensually assumed that given the similarities in their visual configuration, attentional limitation would explain the confusion between both emotions (e.g. Chamberland et al., 2017), but cross-cultural validation of this hypothesis is lacking. Performance for Italian participants also showed higher confusion in the recognition of negative emotions in comparison to other countries. Unfortunately, cultural confusions, such as those we report, have been poorly investigated in the past, and mostly limited to the comparison of Western (i.e. WEIRD) vs Eastern groups (Jack et al., 2009). These Western vs Eastern comparisons have however helped to point out several key cultural differences undermining the universality of the six basic emotions that is generally assumed in the field since the early studies (Ekman & Friesen, 1971; Izard, 1971). For example, a neat differenciation of emotion categories was only observed in WEIRD but not Eastern (Chinese) perceivers (Jack et al., 2012). In addition, the representational structure of emotion expressions in visual face-processing brain regions has been found to be predicted by intra-cultural conceptual similarity between emotions (Brooks et al., 2019). Both findings suggest that conceptual knowledge about emotions scaffolds facial emotion recognition. Stemming from 12 different countries with and without WEIRD participants, our findings support this last notion. The important cultural variability that we retrieved through strong variation of performances across countries and country-specific misclassifying bias, support a constructionist perspective of emotions recognition (e.g. Lindquist, Satput & Gendron, 2015). In that perspective, culture-dependent conceptual knowledge not only provides labels to describe the perception of emotions but deeply shapes it. Using a test conceived by authors from the USA to assess facial emotion recognition may have favored the cultural variations we observed in our study, as the test requires selection within discrete category (thus favoring performance of WEIRD participants) to describe an which are further acted (and not expressed) by white american people only.

Regarding the ability to infer others' mental states as evaluated by the "faux pas" test, the performances of adults (Wu & Keysar, 2007) or children (Liu, Wellman, Tardiff & Sabbagh, 2008) between pairs of countries have already been contrasted in the past. However, the present study is the first to demonstrate a large-scaled cultural effect through a transversal and multi-centric standardized assessment. Interestingly, cross-country comparisons revealed that the best performances were obtained by participants from England, which is the country in which the test has been developed. As for the facial emotion recognition test that we used, this questions the applicability of the faux pas test in other cultures, as some social situations might not be equally interpretated. Specifically, identifying a faux pas requires to represent the mental state of another person, but also to detect that an implicit social rule has been broken. Regarding this second requirement, it is easy to conceive that it can fluctuate from one country to another. As social norms not only drive one's actions but also others' expectations of one's actions, these cultural differences also have the potential to modulate the inference of mental states *per se*. As an example from the task, mistaking a customer for a waiter in a restaurant unequivocally constitutes a faux pas for all participants (100%) from England, but not for Canadians (65.4%). In the same vein, 21.2% of Chinese participants considered that it was a faux pas to give up a seat to an older passenger while riding on a city bus, although 100% of English participants considered this was a normal behavior. Although this very last example was a control item (i.e. originating from a story with no faux pas), we found that cultural variations were stronger when there was actually a faux pas to identify than when it was not the case.

Although we ranked countries accordingly to their performance and discussed the cultural variations observed, it seemed important to specify that we never considered or hypothetized the superiority of one culture on another. While being gold-standard tests of social cognition abilities, both tests originated from and were validated in a specific culture, at a specific period, before being employed all over the world. Our findings do not show that English participants better understand faux pas, but do show that they have a better performance in a task that totally fits their culture, as the aforementionned faux pas originated from the English culture. While it is not clear why a country like Canada, supposedly culturally close to England (Muthukrishna et al., 2020), obtained quite different performances at the test (although they were French-speaking participants from Québec), it still does not mean that they have lower cognitive abilities.

Taken together, our findings suggest that either in suposedly low-level (e.g. emotion recognition) or higher-level (e.g. mental states inference) mechanisms, social cognition is shaped by individual factors as well as culture. Beyond the effect of age, gender and education, categorizing emotions as well as infering others' intentions or beliefs to explain others' behaviors may be reasonnably influenced by local concepts, norms and habits, and thus be differentially apprehended across cultures. This raises questions regarding research practices in human neuroscience, as cultural factors are traditionally neglected. On the contrary, a universalist approach involving the negligence of possible cultural variations is rather favored, as neuroscientific knowledge too often stems from WEIRD samples but aim to be generalized on a worldwide scale and fail to account for the rich cultural diversity (see

Henrich et al., 2010). We believe that beyond the difficulties to form inter-cultural research networks to highlight cultural variations, the question of the generalizability of findings is a pressing and consequential problem in neuroscience (Yarkoni, 2019). Generalizability is applied to many observed results in the field in order to be more appealing and disseminated to a wider audience. What is observed in a given – often small – sample is generalized to the entire humanity. This led us to build most of the neuroscientific knowledge on observations based on WEIRD participants, most often performed by WEIRD researchers. However, highlighting heavy cultural variations, our study underline a strong limit of this approach. Therefore, the field needs to engage in a more epistemologic and critical thinking in order to identify and fight the cultural biases and systemic constraints that impact the design of studies and dissemination of results. As larger-scale multi-cultural studies are also needed to counter-balance this effect, we call states, fundations and charities to better fund these costly initiatives and to favour international collaborations.

These findings could also have important clinical impact. The detection and treatment of cognitive disorders indeed represent one of the biggest challenges in the field of mental health (Alzheimer's Disease International, 2013; Mackin & Arean, 2009; McWhirter et al., 2020). With a rising awareness of these disorders in the general population, higher survival rates of children and elders and the generalization of better diagnosis strategies, their prevalence is expected to increase in the next decades. Because of its impact on social life, social cognition impairments are among the critical cognitive disorders shared by these diverse clinical groups (Kennedy & Adolphs, 2012; Henry et al., 2016). Overall, the present study revealed significant and strong impacts of demographic and cultural factors on two gold-standard measures of facial emotion recognition and mental states inference abilities, which constitute the most tested components of social cognition (Bertoux et al., 2012; Henry et al., 2016; Quesque & Rossetti, 2020). We argue that a mature social cognitive neuroscience aiming to have fundamental relevance to clinical (neuro) psychology, psychiatry or neurology should therefore no longer ignore these social factors for methodological and epistemological convenience. In that way, the development of non-gender-biased instructions could contribute to limit the influence of motivational aspects and expectations on social cognitive tasks. Similarly, setting universality assumptions aside requires innovation in the measurements (Gendron, 2017). For cross-cultural use of measurement tools, contextual adaptations rather than literal translation (which constitutes by far the by default practice) is critical as this ensures that the tool relates to the social context actually encountered by people within their culture (Mehta et al., 2011). Going even further, before developing a new cognitive test, one should first probe whether the underlying principle of the task actually fits to the culture of the targeted population. For example, asking participants to categorize facial expressions into different category of emotions might not be suitable among people that would not spontaneously use emotional terms (e.g. “he is angry”) to describe such material but rather use physical descriptions (e.g. “he frowns”, see Gendron, Roberson, van der Vyver & Barrett, 2014a, 2014b). More generally, when adapting a task implying forced-choice response, researchers should not a priori define the possible responses but rather consider what types of attributions are spontaneously made by perceivers.

As a first step towards the development of more valid theories and tools to quantify social cognition abilities, the IN-SCD represents an important collaborative initiative and seeks to involve new partners. Collecting and sharing data from a larger sample of countries could ultimately lead to study the links between social cognitive performances and classical national indicators (e.g. GDP, PISA, etc.) or between-countries cultural distances estimates (Muthukrishna et al., 2020), given that traditional cultural indicators (e.g. Hofstede’s dimensions) have been questioned (see McSweeney, 2002; Oyserman et al., 2002). Another promising direction for future studies consists in large-scale cross-cultural comparisons of the brain substrates involved while performing classical social cognition tasks, as opposed to classical studies contrasting participants from two cultures (Adams et al., 2010; Kobayashi, Glover & Temple, 2006). This could reveal the existence of culture specific neural differences that might be linked to the variations of behavioral performances. More generally, we believe that the interest in large scale and multi-centric research programs should be generalized to all subfield of human neuroscience. We are aware that this would only constitute a first step towards better practices. A paradigm shift remains however highly needed, and we hope that such collective initiatives will encourage the emergence of more representative theories and measurements of the human cognition in all its diversity.

Methods

Participants

Data of 587 healthy participants (339 women, age range 18–89 years, mean = 58.04, SD = 16.06, mean education = 10.5 years, SD = 5.46) were collected. Centers received local ethics approval (see Supplemental Material 2) and all participants signed informed consent prior to their inclusion. Although inclusion criteria slightly differed between centers, common criteria included (1) no cognitive complaints; (2) no depressive complaints; (3) normal cognitive screening test; (4) no current psychiatric disorder; (5) no past or current neurological disease; (6) native language matching those of the assessment. Additional center-specific inclusion criteria (e.g. normal MRI), sample sizes, gender distribution and age ranges are presented in Supplemental Material 3.

The recruitment of participants and their cognitive assessment was performed through the International Network on Social Cognition Disorders (IN-SCD), a worldwide clinical consortium inaugurated in 2014. Originally, 7 centers from 7 countries formed the IN-SCD; they were based at the Department of Clinical Neurosciences, University of Cambridge (UK), Pitié-Salpêtrière Hospital, Sorbonne Universités, Paris (France), Federal University of Minas Gerais, Belo Horizonte (Brazil), University of Chile, Santiago, (Chile), University of Brescia (Italy), Hospital Universitario Donostia, Donostia-San Sebastian (Spain) and the University Hospital Muenster (Germany). From 2015 to 2019, they were joined by 13 centres from 7 countries, including the University of São Paulo (Brazil), University of East Anglia, Norwich and University College London (UK), Xuan Wu Hospital, Capital Medical University, Beijing (China), FLENI Fundation and INECO/CONICET/Universidad de San Andres, Buenos Aires (Argentina), Universidad Adolfo Ibañez, Santiago (Chile), Universidad del Valle, Cali & Universidad de los Andes, Bogotá (Colombia), CERVO Brain Research Centre, Québec (Canada), Central Clinic No 1 of the Ministry of Internal Affairs of Russia, Moscow (Russian Federation), University of Strasbourg and the University of Lille, Inserm U1172 (France). In this study, 20 centers from 12 countries were involved in the acquisition of data.

Materials

The mini-SEA (mini Social cognition & Emotional Assessment, Bertoux et al., 2012) was administered to all participants. This short battery (approximate administration time <30min) is composed from adaptations of two widely used tests: a modified and reduced version of the faux pas test (Stone et al., 1998) and a reduced version of the Picture of Facial Affect test (Ekman & Friesen, 1976). It allows the computation of a general score (/30) and two subscores (/15). The modified and reduced version of the faux pas test (mFP) is composed of 10 short stories depicting a short social scene in which one-character either commits (in 5 stories) or does not commit (in 5 others) a social faux pas. The test involves the ability to decode social rules and to infer others' knowledge, intention and feelings. The task of the participants is to read the story aloud and to detect the presence or absence of a faux pas (an embarrassing action regarding the context). If a faux pas is detected, the participants have to answer to 5 questions in order to assess their ability to understand who committed the faux pas, what it was, the knowledge and intention of the person who committed the faux pas and the feeling experienced by the person who was victim of it. Two control questions assessed the general understanding of the text for all stories. Participants were invited to read the stories as many times as necessary to answer to the questions.

The reduced version of the Picture of Facial Affect test involves FER. Participants have to choose, among 7 labels (Happiness, Surprise, Neutral, Sadness, Anger, Disgust, Fear), the one that matches the emotion of the 35 faces presented. Gender ratio of the face was maintained constant (50%) and they were all presented in a fixed random order. For all countries except China, where a specific version was used (see Gong, Huang, Wang, et Luo, 2011), the original items were used.

Procedure

As the original tests were published in English, specific translations were used for each language (e.g. the same translation was used in Argentina, Chile, Colombia and Spain, http://www.autismresearchcentre. com/arc_tests) prior to the study or performed when unavailable. Then, specific instructions and scoring procedures applied to the mini-SEA as well as examples of scoring were shared across the network to ensure a rigorous standardization of the assessment.

All participants were recruited through local advertisement. In all centers, a trained senior neuropsychologist supervised or performed the assessment, except in Moskva and Belo Horizonte where it was done by a senior neurologist trained in neuropsychology. The assessment was done in the participants and clinicians' native languages, in compliance with Puente et al. (2013). The items of the tests were printed on A4 pages and presented in the same order.

To ensure standardization of data, procedure for the inclusion of data in center-specific database were shared across the whole network. Data were then fully anonymized. Center-specific databases were then centralized in Cambridge, then Lille. Two quality checks of the data were processed by two individual raters. These checks involved queries and corrections made by the centers in order to match the standardization of the data and the highest quality standard.

Data Analysis

The parameters of the linear mixed models have been estimated with the restricted maximum likelihood method (REML), and the covariance matrix of the random effects have been estimated with the Cholesky parameterization.

The variance partitioning coefficients (VPC), also named intra-class coefficients (ICC), quantifies the proportion of observed variation in the outcome that is attributable to the effect of clustering by country. It is the ratio of the between-cluster variance to the total variance:

$$VPC = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}$$

With σ_0^2 the variance between the countries and σ^2 the residual variance.

To compute the confidence intervals around the VPC, we created a bootstrapped distribution of the VPC (10,000 iterations), then obtained the relevant quantiles from that distribution.

Data & Code availability

The dataset presented in this manuscript and the code allowing to reproduce our analyses are available at https://osf.io/dksv7/?view_only=9386572b0206439ca98d1a03657685ab

Acknowledgements

For their administrative or clinical support related to the acquisition of data, we would like to thank María Goldfeder, Agostina Galiani (INECO Foundation, Argentina), Ma Eugenia Martin, Cecilia Cervino, Ma Eugenia Tabernero, Micaela Arruabarena, Carlos Martinez Canyazo, Yanina Varela (FLENI Foundation, Argentina), Sophia Reul, Sabine Bruchmann (University Hospital Muenster, Germany), Aurélie Funkiewiez, Bruno Dubois, Elisabeth Ryckaert (Pitié-Salpêtrière Hospital, France), Gara Tomé, María de Arriba (Hospital Universitario Donostia, Spain), Clélie Philipps, Thimotée Albasser (University of Strasbourg, France), Alicia Wilcox (University of Cambridge, UK), Stephen Jeffs (University of East Anglia, UK), Ana Carolina Ventura, Andrezza Ferreira dos Santos, Caroline Folli, Emanuel Nery, Gabrielle Oliveira (University of São Paulo, Brazil) and Ângelo Ribeiro Vaz de Faria (Federal University of Minas Gerais, Belo Horizonte, Brazil).

Funding sources

Administrative Department of Science, Technology and Innovation from Colombia (COLCIENCIAS 1106-744-55314), CONICET, FONCyT-PICT 2017-1818, FONCyT-PICT 2017 1820, ANID/FONDAP/15150012, ANID /FONDECYT/1140423 &1100975, GBHI ALZ UK-20-639295, NIH NIA R01 AG057234, and the MULTI-PARTNER CONSORTIUM TO EXPAND DEMENTIA RESEARCH IN LATIN AMERICA (ReDLat; National Institutes of Aging of the National Institutes of Health under award number R01AG057234, Alzheimer's Association grant SG-20-725707-ReDLat, the Rainwater Foundation, and the Global Brain Health Institute), Réseau québécois de recherche sur le vieillissement, the National Natural Science Foundation of China (No. 81100797, 81530036), Beijing high-level health talents training project (No. 2015-3-068), Russian Foundation for Basic Research (No. 18-313-20006), CNPq, Brazil (Ref. 402853/2012-1), Projet Hospitalier de Recherche Clinique (PHRC) inter-régional (IDRCB 2012-A00992-41; HUS N°5330), Alzheimer's Research UK (ARUK_PPG2014B-3; ARUK-DT2016-1), Alzheimer's Society

UK (AS-SF-14-003), Deutsche Telekom, São Paulo Research Foundation (FAPESP - 2016/07967-2), Laboratory of Excellence Development of Innovative Strategies for a Transdisciplinary Approach to Alzheimer's Disease (LabEx DISTALZ) as well as Région Hauts-de-France (Dispositif STaRS) and the Thérèse et René Planiol Foundation. The content is solely the responsibility of the authors and does not represent the official views of the funding agencies.

References

Adams Jr, R. B., Rule, N. O., Franklin Jr, R. G., Wang, E., Stevenson, M. T., Yoshikawa, S., ... & Ambady, N. (2010). Cross-cultural reading the mind in the eyes: an fMRI investigation. Journal of cognitive neuroscience, 22(1), 97-108.

Adolphs, R. (2009). The social brain: neural basis of social knowledge. Annual review of psychology, 60, 693-716.

Alzheimer's Disease International. (2013). Policy brief for heads of Government: The Global Impact of Dementia 2013-2050.

Arnett, J. J. (2008). The neglected 95%: why American psychology needs to become less American.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., ... & Rahwan, I. (2018). The moral machine experiment. Nature, 563(7729), 59-64.

Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., Garcia, A. M., Cetkovich, M., & Ibanez, A. (2017). Men, women… who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. PloS one, 12,(6). e0179336.

Barrett, H. C. (2020). Towards a Cognitive Science of the Human: Cross-Cultural Approaches and Their Urgency. Trends in Cognitive Sciences.

Baron-Cohen, S., O'Riordan, M., Jones, R., Stone, V., & Plaisted, K. (1999). A new test of social sensitivity: Detection of faux pas in normal children and children with Asperger syndrome. Journal of Autism and Developmental Disorders, 29(5), 407-418.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42(2), 241-251.

Bertoux, M., Delavest, M., de Souza, L. C., Funkiewiez, A., Lépine, J. P., Fossati, P., ... & Sarazin, M. (2012). Social cognition and emotional assessment differentiates frontotemporal dementia from depression. J Neurol Neurosurg Psychiatry, 83, 411-416.

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. Archives of clinical neuropsychology, 22, 355-365.

Brooks, J. A., Chikazoe, J., Sadato, N., & Freeman, J. B. (2019). The neural representation of facial-emotion categories reflects conceptual structure. Proceedings of the National Academy of Sciences, 116(32), 15861-15870.

Chamberland, J., Roy-Charland, A., Perron, M., & Dickinson, J. (2017). Distinction between fear and surprise: an interpretation-independent test of the perceptual-attentional limitation hypothesis. Social neuroscience, 12(6), 751-768.

Cohn, M., & Statucka, M. (2019). Origins matter: Culture impacts cognitive testing in Parkinson’s disease. Frontiers in human neuroscience, 13, 269.

Coutrot, A., Silva, R., Manley, E., De Cothi, W., Sami, S., Bohbot, V. D., ... & Spiers, H. J. (2018). Global determinants of navigation ability. Current Biology, 28(17), 2861-2866.

Daugherty, J. C., Puente, A. E., Fasfous, A. F., Hidalgo-Ruzzante, N., & Pérez-Garcia, M. (2017). Diagnostic mistakes of culturally diverse individuals when using North American neuropsychological tests. Applied Neuropsychology: Adult, 24, 16-22.

Denham, S. A., Bassett, H. H., & Wyatt, T. (2007). The socialization of emotional competence. Handbook of socialization: Theory and research, 614-637.

Eisenberg, N., Cumberland, A., & Spinrad, T. L. (1998). Parental socialization of emotion. Psychological inquiry, 9(4), 241-273.

Ekman, P. & Friesen, W. V. (1976). Pictures of facial affect. Consulting psychologists Press. Palo Alto, California.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of personality and social psychology, 17, 124.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. Psychological bulletin, 128(2), 203.

Fiedler, K. (2008). Language: A toolbox for sharing and influencing social reality. Perspectives on psychological science, 3(1), 38-47.

Filippova, E., & Astington, J. W. (2008). Further development in social reasoning revealed in discourse irony understanding. Child development, 79(1), 126-138

Focus on social neuroscience (2012). Nature Neurosciences. 15, 645–645 Editorial. doi:10.1038/nn0512-645

Gendron, M. (2017). Revisiting diversity: Cultural variation reveals the constructed nature of emotion perception. Current opinion in psychology, 17, 145-150.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014a). Cultural relativity in perceiving emotion from vocalizations. Psychological science, 25(4), 911-920.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014b). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. Emotion, 14(2), 251.

Graham, T., & Ickes, W. (1997). When women's intuition isn't greater than men's. In W. Ickes (E.d.), Empathic accuracy (pp. 117-143). New York: Guilford Press.

Greenwood, J. D. (2004). What happened to the “social” in social psychology?. Journal for the theory of social behaviour, 34(1), 19-34.

Halberstadt, J., Ruffman, T., Murray, J., Taumoepeau, M., & Ryan, M. (2011). Emotion perception explains age-related differences in the perception of social gaffes. Psychology and Aging, 26, 133–136.

Hayden, K. M., Makeeva, O. A., Newby, L. K., Plassman, B. L., Markova, V. V., Dunham, A., ... & Roses, A. D. (2014). A comparison of neuropsychological performance between US and Russia: Preparing for a global clinical trial. Alzheimer's & Dementia, 10(6), 760-768.

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

Henrich, J., Heine, S. J., & Norenzayan, A. (2010)b. Most people are not WEIRD. Nature, 466, 29-29.

Henry, J. D., Von Hippel, W., Molenberghs, P., Lee, T., & Sachdev, P. S. (2016). Clinical assessment of social cognitive function in neurological disorders. Nature Reviews Neurology, 12(1), 28.

Hyde, J. S. (2014). Gender similarities and differences. Annual review of psychology, 65, 373-398.

Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation?. Personal Relationships, 7(1), 95-109.

Izard, C. E. (1971). The face of emotion.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. Current biology, 19(18), 1543-1548.

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. Proceedings of the National Academy of Sciences, 109(19), 7241-7244.

Kennedy, D. P., & Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. Trends in cognitive sciences, 16(11), 559-572.

Kirkland, R. A., Peterson, E., Baker, C. A., Miller, S., & Pulos, S. (2013). Meta-analysis Reveals Adult Female Superiority in" Reading the Mind in the Eyes Test". North American Journal of Psychology, 15(1).

Kobayashi, C., Glover, G. H., & Temple, E. (2006). Cultural and linguistic influence on neural bases of 'Theory of Mind': an fMRI study with Japanese bilinguals. Brain and language, 98(2), 210-220.

Kret, M. E., & de Gelder, B. (2012). A review on sex differences in processing emotional signals. Neuropsychologia, 50(7), 1211-1221.

Lausen, A., & Schacht, A. (2018). Gender differences in the recognition of vocal emotions. Frontiers in psychology, 9, 882.

Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion?. Current directions in psychological science, 24(2), 99-108.

Liu, D.,Wellman,H.M.,Tardiff,T.,& Sabbagh, M.A.(2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. Developmental Psychology, 44(2), 523–531.

Mackin, R. S., & Areán, P. A. (2009). Incidence and documentation of cognitive impairment among older adults with severe mental illness in a community mental health setting. The American Journal of Geriatric Psychiatry, 17(1), 75-82.

Matsumoto, D., & Juang, L. (2016). Culture and psychology. Nelson Education.

McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith – A failure of analysis. Human Relations, 55, 89–118.

McWhirter, L., Ritchie, C., Stone, J., & Carson, A. (2020). Functional cognitive disorders: a systematic review. The Lancet Psychiatry, 7(2), 191-207.

Mehta, Urvakhsh M., Thirthalli, J., Naveen Kumar, C., Mahadevaiah, M., Rao, K., Subbakrishna, D.K., Gangadhar, B.N., Keshavan, M.S., 2011b. Validation of social cognition rating tools in Indian setting (SOCRATIS): a new test-battery to assess social cognition. Asian J. Psychiatry, 4, 203-209.

Merten, J. (2005). Culture, gender and the recognition of the basic emotions. Psychologia, 48(4), 306-316.

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) Psychology: Measuring and Mapping Scales of Cultural and Psychological Distance. Psychological Science, 0956797620916782.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. Emotion Review, 5(1), 8-15.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. Cognitive science, 26(5), 653-684.

Norman, M. A., Evans, J. D., Miller, W. S., & Heaton, R. K. (2000). Demographically corrected norms for the California verbal learning test. Journal of Clinical and Experimental Neuropsychology, 22, 80-94.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. Psychological Bulletin, 128, 3–72.

Puente, A. E., Pérez-García, M., Vilar-Lopez, R., Hidalgo-Ruzzante, N., & Fasfous, F. (2013). Neuropsychological assessment of culturally and educationally dissimilar individuals. In F. Paniagua & A. M. Yamada (Eds.) Handbook of multicultural mental health: Assessment and treatment of diverse population (2nd ed., pp. 225–242). New York, NY: Elsevier.

Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. Perspectives on Psychological Science, 15(2), 384-396.

Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of facial, auditory, and bodily emotions I older adults. Journal of Gerontology: Psychological Sciences, 64B, 696-703.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. Neuroscience & Biobehavioral Reviews, 32(4), 863–881.

Santamaría-García H, Baez S, Gómez C, Rodriguez-Villagra O, Portela M, Reyes P, Klahr J, Matallana D, Ibanez A. (2020). The role of social cognition skills (SCS) and social determinants of health (SDH) in predicting symptoms of mental illness. Translational Psychiatry. https://doi.org/10.1038/s41398-020-0852-4

Spitzer, N., Shafir, T., Lerman, Y., & Werner, P. (2019). The Relationship Between Caregiver Burden and Emotion Recognition Deficits in Persons With MCI and Early AD. Alzheimer Disease & Associated Disorders, 33(3), 266-271.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. Journal of cognitive neuroscience, 10, 640-656.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. Science, 344, 603-608.

Wang, Y., & Su, Y. (2006). Theory of mind in old adults: The performance on Happé's stories and faux pas stories. Psychologia, 49(4), 228-237.

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. Psychological science, 18, 600-606.

Yarkoni, T. (2019, November 22). The Generalizability Crisis. https://doi.org/10.31234/osf.io/jqw35