PsyCh Journal
INSTITUTE OF PSYCHOLOGY
Chinese Academy of Sciences

PsyCh Journal (2020)
DOI: 10.1002/pchj.417

Cross-cultural differences in intercultural mindreading: Evidence from a sample of Palestinian, Italian, and German adolescents

Thomas Hünefeldt,[1] Ola Hussein,[2] and Marta Olivetti Belardinelli[1]

[1]Sapienza University of Rome, Rome, Italy, [2]Al-Quds University, Jerusalem, Palestine

Abstract: Despite the fact that the "Reading the Mind in the Eyes" Test (RMET) is now available in more than 20 languages, there are only very few cross-cultural researchers using this test, and these researchers generally focus on North American versus East Asian cultures. Considering that the RMET stimuli were selected and constructed in the United Kingdom, this research explored cross-cultural differences in intercultural mindreading with a large sample of adolescents from Palestine (PAL), Italy (ITA), and Germany (GER). In addition to significant main effects of age (younger < older) and gender (male < female), we found a significant main effect of country (PAL < ITA < GER) and a significant interaction between gender and country. Individualism was not related to mindreading in any of the three countries whereas collectivism was positively related in PAL, but not in ITA or GER, accounting only for a very small amount of the variance. Our results suggest that (a) there may be cultural ingroup effects on mindreading, (b) the known female superiority in mindreading may be moderated by cultural factors, and (c) depending on cultural factors, individualism and collectivism may be differently related to mindreading.

Keywords: age differences; collectivism; cultural differences; gender differences; individualism; theory of mind

Correspondence Dr. Thomas Hünefeldt, Sapienza University of Rome, Rome, Italy. Email: thomas.huenefeldt@uniroma1.it

Received 2 January 2020. Accepted 12 November 2020.

The ability to recognize mental states such as emotions, desires, intentions, beliefs, and so on is an important sociocognitive ability which has been referred to in a variety of different ways, including "mentalization" (Fonagy, Gergely, Jurist, & Target, 2002), "theory of mind" (Premack & Woodruff, 1978), and "mindreading" (Whiten, 1991). This ability comprises a large set of rather loosely related more specific abilities which mainly differ in terms of (a) the kinds of individuals whose mental states are to be recognized (e.g., oneself vs. others, human vs. nonhuman, etc.), (b) the kinds of mental states that are to be recognized (e.g., simple vs. complex, affective vs. cognitive, etc.), and (c) the cues by means of which they are to be recognized (e.g., textual vs. physical, bodily vs. environmental, etc.). In the research presented here, we explore cross-cultural differences in a particular aspect of this ability; namely, the recognition of complex mental states by means of facial cues displayed by other human beings.

This particular aspect of the ability to recognize mental states has so far been investigated mainly with the "Reading the Mind in the Eyes" Test (RMET; Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Unlike typical tests of "emotion recognition," the RMET does not assess the recognition of relatively "simple" affective mental states such as the so-called "basic emotions," but it focuses on more "complex" affective mental states such as doubtful, flirtatious, concerned, and so on; that is, on affective mental states that are supposed to "involve the attribution of a belief or an intention – a cognitive mental state – to the person" (Baron-Cohen et al., 2001, p. 243). And unlike other well-established tests of complex mental state recognition such as the Faux Pas Recognition Test (Stone, Baron-Cohen, & Knight, 1998) or the Strange Stories test (Happé, 1994), the RMET does not rely on textual cues but on a particular kind of facial cues (for details, see the Instruments section).

As the RMET is by far the most frequently used test of complex mental state recognition by means of facial cues, we employed that test in the research presented here. Considering that the RMET stimuli had been selected and constructed in the United Kingdom, we wanted to explore cross-cultural differences in intercultural mental state recognition with a large sample of Palestinian, Italian, and German adolescents. We chose adolescents rather than adults mainly because they can be recruited almost as easily as university students, but they much better represent the general population of their countries. In fact, enrollment in the secondary education levels is ∼85% in Palestine and ∼100% in Italy and Germany whereas enrollment in tertiary education levels is ∼42% in Palestine, ∼66% in Italy, and ∼68% in Germany (UNdata, 2017). Furthermore, we chose adolescents from these three countries for mainly two reasons. First, Palestine, Italy, and Germany represent one Arab versus two different European countries and one Germanic versus two different Mediterranean countries, which differ in several important respects, including their geographical and cultural distance to the cultural context in which the RMET stimuli were selected and constructed (i.e., the United Kingdom). Second, Palestine, Italy, and Germany are the native countries of the three authors of this study. We are thus familiar with the culture of these three countries, especially their languages and educational institutions.

Given that we employed the RMET, we will formulate our research questions and hypotheses with respect to performance on that test, thereby using the term *mindreading* as convenient shorthand for what is assessed by it. In particular, we investigate the following three questions: (1) *Are there cultural ingroup effects on mindreading?* (2) *Are the known age and gender differences in mindreading moderated by culture?* (3) *Is mindreading related to individualism and collectivism?* In the following sections, we will introduce our research hypotheses based on a brief review and discussion of the relevant scientific literature concerning each of these three questions.

Are there cultural ingroup effects on mindreading?

Despite the fact that the RMET is now available in more than 20 languages, including several non-European ones, there are very few cross-cultural studies using this test, all of them with adult samples from Caucasian versus East Asian national or ethnic groups (Adams et al., 2010; Bjornsdottir & Rule, 2016; Franklin, Stevenson, Ambady, & Adams, 2015; Lee et al., 2018; Prevost et al., 2014). These studies have consistently found that mindreading accuracy was higher when mental states were both expressed and recognized by members of the same national or ethnic group, suggesting a cultural ingroup effect on mindreading that is moderated by the amount of exposure to the cultural outgroup. These findings are consistent with the well-established results of research on emotion recognition (Elfenbein, 2007, 2013, 2015; Elfenbein & Ambady, 2002a, 2002b, 2003; Scherer, Clark-Polner, & Mortillaro, 2011).

Based on the research on emotion recognition and mindreading, we thus expected to find evidence of cultural ingroup effects on mindreading in our study. In particular, given that Italy and Germany are both geographically and culturally much closer than Palestine to the cultural context in which the RMET stimuli were selected and constructed (i.e., the United Kingdom), we expected that both Germans and Italians would score higher on that test than would Palestinians. We furthermore expected that Germans would tend to score higher than would Italians because Germany is not only geographically but also culturally somewhat closer to the United Kingdom than is Italy (De Santis, Maltagliati, & Salvini, 2016). In fact, Germany and the United Kingdom are both transalpine, Germanic-speaking countries whereas Italy is a Mediterranean, Romance-language-speaking country.

Are the known age and gender differences in mindreading moderated by culture?

Mindreading is known to be affected by both age and gender. In particular, there is consistent evidence that mindreading is better in younger than in older adults (Henry, Phillips, Ruffman, & Bailey, 2013) and in female than in male adults (Kirkland, Peterson, Baker, Miller, & Pulos, 2013). Furthermore, the few studies investigating the effects of age and gender on adolescents' mindreading have provided evidence of increasing mindreading abilities between early adolescence and early adulthood (Hünefeldt, Laghi, Ortu, & Olivetti Belardinelli, 2013; Vetter, Leipold, Kliegel, Phillips, & Altgassen, 2013) and confirmed the female superiority generally found in studies with adults (Hünefeldt et al., 2013; Rutherford, Baron-Cohen, & Wheelwright, 2002). However, it is an open question whether age and gender differences in mindreading may be moderated by cultural variables. In fact, while potential interactions between culture and age on mindreading have so far not been investigated,

the only study investigating interactions between culture and gender found neural, but no behavioral, evidence of such interactions in a sample of U.S. American and Japanese university students (Franklin et al., 2015).

In the absence of evidence from cross-cultural studies, research on the effects of culture-specific age or gender roles on mindreading could provide valuable suggestions as to whether and how age and gender differences in mindreading may be moderated by cultural variables. To our knowledge, however, there is only one study that has investigated the effects of gender roles on mindreading whereas the effects of age roles on mindreading have thus far not been investigated. Notably, the results of this study suggest that the effects of gender roles extend beyond those of anatomical sex with regard to social cognitive abilities such as mindreading (Vonk, Mayhew, & Zeigler-Hill, 2016). In particular, feminine gender roles, as assessed by the Bem Sex Role Inventory (Bem, 1974), were consistently a strong predictor of higher mindreading scores, even when sex was controlled in the model. Supposing, then, that mindreading may be affected by gender roles and perhaps also by age roles, we expected that the moderating effect of culture on the effects of age or gender on mindreading would be stronger the more that two cultures differ in their age or gender roles.

Regarding our sample of Palestinian, Italian, and German adolescents, we expected that moderating effects of culture would concern the effects of gender rather than those of age. In fact, the difference between the age roles of younger and older adolescents is likely to be rather similar in all three countries whereas the difference between male and female gender roles, as assessed by gender-equality indices such as the Equality subindex of the World Value Survey (Inglehart et al., 2014), the Global Gender Gap Report (2018), and the Gender Inequality Index (2017), is known to be much larger in Arab than in European countries. We therefore expected that gender differences in mindreading would be smaller in Italians and Germans than in Palestinians.

Is mindreading related to individualism and collectivism?

Cultures differ not only on external variables such as geographical location and political institutions but also on internal variables such as values (Hofstede, 1980, 2001; Inglehart et al., 2014). In cross-cultural psychology, the distinction between individualistic and collectivistic values has been particularly influential, but the constructs of individualism and collectivism have been conceptualized in a large variety of different ways (Oyserman, Coon, & Kemmelmeier, 2002; Taras et al., 2014; Voronov & Singer, 2002). It would therefore not be surprising to find a rather large variety of hypotheses and predictions concerning the relationship between individualism and collectivism on one hand, and the ability to recognize mental states on the other hand. However, to our knowledge, only two different kinds of hypotheses and predictions have been made so far.

Following Matsumoto (1989, 1992), Elfenbein and Ambady (2003) argued that "individualist cultures show lower concern about the possible negative consequences of recognizing negative emotions that are potentially damaging to social relationships, and so should not attenuate such communication" (p. 94). They therefore hypothesized that "individualism is associated with the more accurate perception of negative emotions but not with the more accurate perception of positive or neutral emotions" (p. 94). However, the results of their meta-analysis on cross-cultural differences in emotion recognition did not support this hypothesis.

Vu, Finkenauer, Huizinga, Novin, and Krabbendam (2017), however, argued that "collectivism is associated with other-oriented cognition, emotion, and motivation, while individualism is associated with self-oriented cognition, emotion, and motivation" (p. 3/20). They therefore predicted that "collectivism, compared to individualism, might entail faster recognition and more accurate understanding of the thoughts and feelings of other people" (p. 3/20). Backed up by a brief review of research investigating the effects of individualism and collectivism on theory of mind, they further hypothesized that the predicted effect would be found not only at the country level but also at both the individual level and the situational level, with possible interactions between these levels. However, using both the original Caucasian and the East Asian version of the Reading the Mind in Eyes test, they failed to find evidence of the predicted effects.

A possible reason why neither Elfenbein and Ambady (2003) nor Vu et al. (2017) found evidence in favor of their respective hypotheses might be the fact that both groups of researchers assessed individualism and collectivism too broadly. In fact, Elfenbein and Ambady (2003) considered individualism and collectivism at the country level using Hofstede's (2001) unidimensional individualism–collectivism values whereas Vu et al. assessed individualism and collectivism at both the country and the individual levels via a single

individualism–collectivism ratio. As the overly broad ways in which individualism and collectivism are defined and assessed has been repeatedly criticized (Oyserman et al., 2002; Taras et al., 2014; Voronov & Singer, 2002), a more fine-tuned, differential approach to investigating their relationship with the ability to recognize mental states is called for.

In the research presented here, we therefore employed a measure that considers individualism and collectivism as two independent dimensions. In particular, we employed the Auckland Individualism and Collectivism Scale (AICS; Shulruf et al., 2011; Shulruf, Hattie, & Dixon, 2007), which conceives individualism in terms of valuing personal competitiveness, uniqueness and responsibility, and collectivism in terms of valuing interpersonal advice and harmony (for details, see the Instruments section). Given this way of conceiving individualism and collectivism, we expected that the AICS scores of individualism (IND) and collectivism (COL) would be differently related to mindreading and that COL rather than IND would be related to mindreading in the ways predicted by the aforementioned hypotheses. Based on Elfenbein and Ambady's (2003) hypothesis, we expected that mindreading would tend to be negatively related with COL because COL assesses the value of interpersonal harmony that is supposed to be associated with a less accurate perception of negative mental states. Based on Vu et al.'s (2017) hypothesis, by contrast, we expected that mindreading would tend to be positively related with COL because COL assesses the other-related processing that is supposed to be associated with more accurate understanding of the thoughts and feelings of other people. As neither of these two hypotheses refers to conditions of age and culture, we further expected that the predicted effects would be observed cross-culturally regardless of age.

Methods

Participants

For this study, 596 adolescents were recruited at public schools located in urban neighborhoods in Palestine (East Jerusalem), Central Italy (Rome), and Northern Germany (Bremen). These three countries differ on a variety of socioeconomic variables, including affluence, education level, and family size, with Palestinian adolescents generally living in less affluent, less educated, and larger families than Italian or German adolescents (cf. UNdata, 2017). Though the education systems in these three countries are hardly comparable, we took care to select schools that were similar in that they were all tuition-free, thus fairly well representing the general population, and in that they covered different educational profiles (e.g., technical, scientific, linguistic, etc.), potentially leading to final degrees that allow admission to all kinds of public universities in the country. Apart from age and gender, we were unable to obtain individual demographic information from the participants.

In addition to two gender groups (male and female), we defined two age groups: 14–16 years and 17–19 years. The cut-off point was chosen mainly to obtain age groups of roughly the same size. However, in all three countries, the younger group included exclusively students attending compulsory secondary education classes whereas the older group included mainly students attending voluntary secondary education classes. As the age and gender distributions in the subsamples of Palestinian, Italian, and German adolescents were not balanced, we randomly selected participants from these three subsamples to obtain age- and gender-balanced subsamples. The remaining sample consisted of 499 adolescents (see Table 1). Chi-square tests confirmed that the three subsamples did not differ in their distribution by age and gender, $X^2(2, N = 499) = .003$, $p = .96$.

Instruments

Revised adult RMET

The RMET (Baron-Cohen et al., 2001) is a well-established test of advanced mental state recognition that has been successfully used not only with adults but also with adolescents in both clinical and nonclinical studies (e.g., Fossati, Feeney, Maffei, & Borroni, 2014; Naor-Ziv & Glicksohn, 2016; Vetter et al., 2013). It consists of 36 black-and-white photographs featuring the eye region of 18 male and 18 female Caucasian posers. Each photograph is associated with four mental state terms: one target and three foils (for an example of a RMET stimulus, see Figure 1). Performance on the RMET is generally assessed in terms of the number of correctly recognized test items. Accordingly, the total score ranges from 0 to 36, with a score of 9 representing chance level.

Notably, the photographs used in the RMET were taken from commercial magazines available to the British authors of the test. The mental state terms (target and foils) were generated by the first two authors. The target terms were eventually established on the basis of consensus of a large population (N = 225) consisting of normal British adults

Table 1
Final Sample Distribution by Age, Gender, and Country

	14–16 Years		17–19 Years		Total	
	n	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SD*)
Palestinians						
Male	50	15.4 (0.5)	47	17.3 (0.5)	97	16.3 (1.1)
Female	51	15.3 (0.7)	49	17.6 (0.6)	100	16.4 (1.3)
Total *N*	101	15.3 (0.6)	96	17.4 (0.6)	197	16.4 (1.2)
Italians						
Male	46	15.5 (0.7)	44	17.7 (0.8)	90	16.6 (1.3)
Female	47	15.2 (0.7)	45	17.7 (0.8)	92	16.4 (1.4)
Total *N*	93	15.4 (0.7)	89	17.7 (0.8)	182	16.5 (1.4)
Germans						
Male	30	15.1 (0.9)	29	18.1 (0.9)	59	16.6 (1.8)
Female	31	15.2 (0.8)	30	17.8 (0.9)	61	16.5 (1.6)
Total *N*	61	15.1 (0.9)	59	17.9 (0.9)	120	16.5 (1.7)
Total						
Male	126	15.4 (0.7)	120	17.8 (0.8)	246	16.5 (1.4)
Female	129	15.3 (0.7)	124	17.7 (0.7)	253	16.4 (1.4)
Total *N*	255	15.3 (0.7)	244	17.7 (0.8)	499	16.5 (1.4)

drawn from community and education classes in Exeter, from public library users in Cambridge, and from undergraduate students at Cambridge University. For an item to be included in the test, at least 50% of the subjects had to select the target and no more than 25% could select any one of the foils.

Regarding its psychometric properties (for a review and additional evidence, see Olderbak et al., 2015), the RMET typically has acceptable test–retest reliability (e.g., Hallerbäck, Lugnegård, Hjärthag, & Gillberg, 2009; Khorashad et al., 2015; Prevost et al., 2014; Vellante et al., 2013), but poor internal consistency. For example, Cronbach's α was found to be .48 for the English version (Ragsdale & Foley, 2011), .60 for an Italian version (Vellante et al., 2013), .53 for a French version (Prevost et al., 2014), and .37 for a Persian version (Khorashad et al., 2015). As suggested by Olderbak et al. (2015), the low internal consistency may be due to various test attributes, regarding both the photographs (kinds of mental states expressed, angle of face, ratio of dark and light, etc.) and the mental state terms (commonness, frequency, and distribution of the response-option words). Despite these limitations, the RMET can successfully differentiate between groups presumed to differ in their ability to recognize mental states, showing evidence of discriminant validity with measures of autistic traits (Baron-Cohen et al., 2001). Evidence of convergent validity with other standard measures of mental state recognition is somewhat mixed (Olderbak et al., 2015), maybe due to the fact that other measures assess other aspects of this complex ability (e.g., mental state recognition by means of textual or environmental rather than facial cues). In fact, as emphasized by various authors, there is no gold standard against which to validate it (e.g., Hallerbäck et al., 2009; Prevost et al., 2014; Vellante et al., 2013).

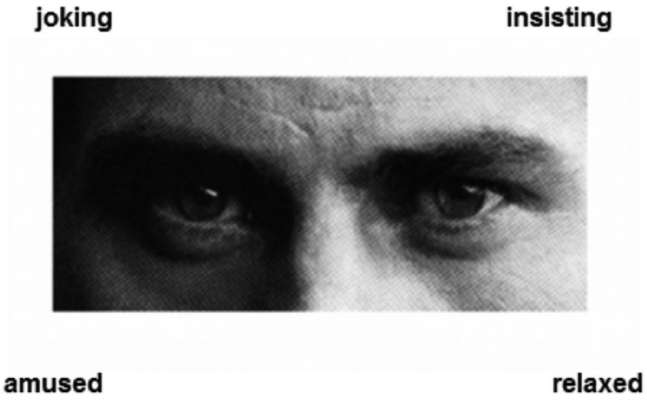

Figure 1. Example of a Reading the Mind in the Eyes Test stimulus.

In the current study, we used an Arabic, an Italian, and a German version of the RMET with the Palestinian, Italian, and German sample, respectively. The Italian version has been validated in an extensive psychometric study (Vellante et al., 2013). The German version (Bölte, 2005) has demonstrated its validity in a number of clinical studies (e.g., Kother et al., 2012; Schilling et al., 2012). It has been authorized by the Autism Research Centre at the University of Cambridge and is one of the standard tests available on its website. The Arab version has been translated for the purpose of this study by using the standard back-translation method, coordinated and supervised by the second author of this study, who is a native Arabic speaker. In line with

the results of previous studies (Khorashad et al., 2015; Prevost et al., 2014; Ragsdale & Foley, 2011; Vellante et al., 2013), the internal consistency was rather low in all three versions; Italian sample: α = .70, Palestinian sample: α = .56, German sample α = .43.

AICS

The AICS (Shulruf et al., 2011; Shulruf et al., 2007) is a short and easy-to-use measurement tool for individualism and collectivism that is based on an extensive meta-analysis of studies on this topic (Oyserman et al., 2002) and that has been found to be highly reliable and valid across a large range of cultures (Shulruf et al., 2011). It consists of 26 items related to two facets of collectivism: advice (seeking advice from people close to oneself, before making decisions) and harmony (seeking to avoid conflict), and three facets of individualism: competitiveness (striving for personal goals is one's prime interest), uniqueness (distinction of the self from the other), and responsibility (acknowledging one's responsibility for one's actions). Participants are requested to rate on a 6-point Likert scale of 1 (*Never*) to 6 (*Always*) how often they would think or behave as described in the test items.

In the current study, we used the existing Italian version (Shulruf et al., 2011) with the Italian sample. With the German and the Palestinian samples, we used a German and an Arab version that we translated for the purpose of this study by using the standard back-translation method. Participants' individualism and collectivism were assessed in terms of their mean scores of IND and COL. In all three samples, internal reliability of the scales was satisfactory for both IND (PAL: α = .74, ITA: α = .74, GER: α = .77) and COL (PAL: α = .69, ITA: α = .66, GER: α = .72). However, IND and COL were uncorrelated only in the Italian and the German samples whereas they were positively correlated in the Palestinian sample, r = .243, p = .001, two-tailed, suggesting that the constructs might not be fully equivalent in the three samples.

Procedure

After obtaining ethical approval from the review board of Sapienza University of Rome and from local school authorities, we obtained informed consent from the adolescent participants and their parents. The RMET and the AICS were administered in the adolescents' classrooms during regular class time, in the presence of their teacher. The same procedure was used in all classrooms. Participants received a booklet containing the RMET and AICS instructions and answer sheets. They were asked to read through the RMET instructions, the experimenter explained the test procedure with an example of a practice stimulus, and participants were given the opportunity to ask questions. The experimenter emphasized that the test must be taken individually and in silence. Once there were no more questions, participants were asked to follow the RMET instructions, and then RMET stimuli were presented on a large screen. Each stimulus was presented for 25 s, followed by a blank slide presented for 5 s. Immediately following the RMET, the AICS was administered. Participants were asked to read through the AICS instructions and to ask questions about the procedure. Once there were no more questions, participants were asked to follow the AICS instructions and to complete the AICS individually and in silence. On average, it took ~40 to 45 min to administer the two tests. Afterwards, participants and their teacher discussed the research with the experimenter. In general, participants were eager to learn about the research and provided positive feedback.

Design and statistical analyses

To address our first two research questions (Are there cultural ingroup effects on mindreading? Are the known age and gender differences in mindreading moderated by culture?), we performed an analysis of variance (ANOVA) with Age (younger vs. older), Gender (male vs. female), and Country (Palestine vs. Italy vs. Germany) as between-subject factors on the participants' RMET scores. To test our hypotheses concerning our third research question (Is mindreading related to individualism and collectivism?), we performed for each of the three samples a Pearson correlation analysis between the participants' RMET scores, on one hand, and their AICS scores of IND and COL, on the other. In addition to these main analyses, we performed two supplementary analyses to assess if and in how far the country differences evidenced by the main analyses could be due to country-level differences in individualism or collectivism. In particular, we first performed a multivariate analysis of variance (MANOVA) with Country (Palestine vs. Italy vs. Germany) as the between-subject factor on the AICS scores of IND and COL, and second, a hierarchical regression analysis including Age, Gender, Country, IND, and COL as predictors of mindreading. All statistical analyses were performed using SPSS Statistics Version 19 (IBM, Armonk, NY).

Results

Effects of age, gender, and country on mindreading

The ANOVA with Age (younger vs. older), Gender (male vs. female), and Country (Palestine vs. Italy vs. Germany) as between-subject factors on the participants' RMET scores evidenced significant main effects of Age, Gender, and Country as well as a significant two-way interaction between Gender and Country whereas the other interactions were not statistically significant (see Table 2).

Regarding the significant main effects of Age and Gender, older adolescents (M = 25.5, SE = 0.3) scored higher than younger adolescents (M = 23.9, SE = 0.3), and females (M = 26.0, SE = 0.2) scoring higher than males (M = 23.3, SE = 0.3).

For the significant main effect of Country, pairwise comparisons revealed that Palestinians (M = 22.6, SE = 0.3) scored significantly lower than both Italians (M = 25.6, SE = 0.3), $p < .001$, and Germans (M = 26.7, SE = 0.3), $p < .001$, with Italians scoring significantly lower than Germans, $p < .05$.

For the significant interaction between Gender and Country (see Figure 2), a follow-up ANOVA revealed that in all three countries, females scored significantly higher than males, but the gender difference was much smaller in Palestinians, $p < .05$, than in Italians, $p < .001$, and Germans, $p < .001$.

Relation of mindreading to individualism and collectivism

Results of the Pearson correlation analyses between mindreading, as assessed by the RMET, and individualism and collectivism, as assessed by the AISC scores of IND and COL, in the three samples of Palestinian, Italian, and German adolescents are reported in Table 3. As evident from the table, the correlation patterns differed substantially between the Palestinian sample on one hand, and the Italian and German samples on the other hand. In the Palestinian sample, RMET was positively correlated with COL whereas its correlation with IND was not statistically

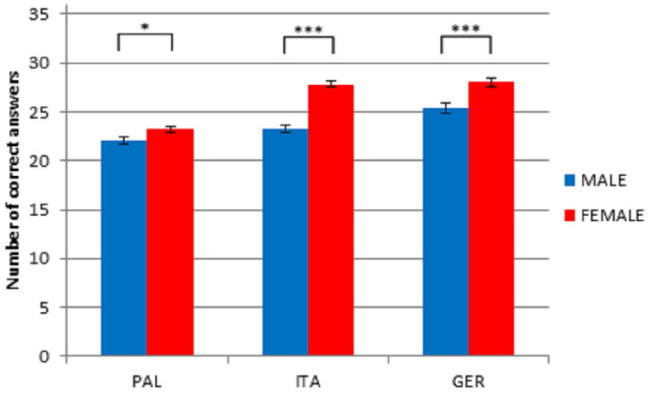

Figure 2. Mean Reading the Mind in the Eyes Test scores of male and female Palestinian, Italian, and German participants. Vertical bars denote *SE*s. *$p < .05$. ***$p < .001$.

Table 3
Results of the Pearson Correlation Analyses between RMET and the AICS Scores of IND and COL in the Palestinian, Italian, and German Samples

Sample	Variable	*M*	*SD*	1	2
Palestinian	1. IND	4.25	0.67		
	2. COL	3.92	0.74	.243*	
	3. RMET	22.61	3.77	.126	.220*
Italian	1. IND	3.90	0.70		
	2. COL	3.57	0.69	−0.75	
	3. RMET	25.57	4.26	−.102	.59
German	1. IND	4.20	0.65		
	2. COL	3.51	0.74	.75	
	3. RMET	26.74	3.03	−.114	−.010

Note. RMET = Reading the Mind in the Eyes Test; AICS = Auckland Individualism and Collectivism Scale; IND = individualism; COL = collectivism. *$p < .01$, two-tailed.

Table 2
Results of the Analysis of Variance with Age, Gender, and Country as Between-Subject Variables on the Reading the Mind in the Eyes Test

Predictor	*df*	*SS*	*MS*	*F*	*p*	η_p^2
(Intercept)	1	296720.76	296720.76	25687.46	.000	
Age	1	225.33	225.33	19.51*	.000	.039
Gender	1	864.15	864.15	74.81*	.000	.133
Country	2	1475.77	737.89	63.88*	.000	.208
Age × Gender	1	0.44	0.44	0.04	.845	.000
Age × Country	2	64.33	32.17	2.78	.063	.011
Gender × Country	2	262.08	131.04	11.34*	.000	.045
Age × Gender × Country	2	33.95	16.97	1.47	.231	.006
Error	487	5625.43	11.55			

Note. SS = sums of squares; MS = mean squares.
*$p < .001$.

significant. In the Italian and German samples, by contrast, neither IND nor COL was significantly correlated with RMET.

Effect of Country on individualism and collectivism

The MANOVA with Country (Palestinian vs. Italian vs. German) as the between-subjects factor on the AICS scores of IND and COL evidenced significant main effects on both IND, $F(2, 496) = 13.6$, $p < .001$, $\eta_p^2 = .05$, and COL, $F(2, 496) = 16.1$, $p < .001$, $\eta_p^2 = .06$. Pairwise comparisons revealed that IND was significantly lower in Italians ($M = 3.90$, $SE = 0.05$) than in Palestinians ($M = 4.25$, $SE = 0.05$), $p < .001$, and Germans ($M = 4.20$, $SE = 0.06$), $p < .01$, whereas it did not differ significantly between Palestinians and Germans. By contrast, COL was significantly higher in Palestinians ($M = 3.92$, $SE = 0.05$) than in Italians ($M = 3.57$, $SE = 0.05$), $p < .001$, and Germans ($M = 3.51$, $SE = 0.06$), $p < .001$, whereas it did not differ significantly between Italians and Germans.

Individualism and collectivism as predictors of mindreading

Results of the hierarchical regression analysis predicting mindreading, as assessed by the RMET, in our sample of Palestinian, Italian, and German adolescents are reported in Table 4. Age and Gender were entered as categorical variables in the first block. In the second block, we entered the two-level categorical country variables "European" (0 = *no*, 1 = *yes*) and "Germanic" (0 = *no*, 1 = *yes*), thus dummy-coding our three-level country variable for the scope of the regression. In the third block, we entered the AICS scores of IND and COL.

As evident from Table 4, the whole model accounted for 31.3% of the variance. Age and Gender alone accounted for 13.5% of the variance. The two country variables accounted for 17.2% of the variance, strongly increasing the fit of the model, $p < .001$. Finally, IND and COL accounted only for 0.5% of the variance without significantly increasing the fit of the model. In the overall model, Age, Gender, and both country variables were significantly related to mindreading whereas the effects of IND and COL were not statistically significant.

Discussion

We introduced our hypotheses in the context of three questions that we wanted to address in the research presented here. Accordingly, we will discuss our results with reference to these three questions.

Are there cultural ingroup effects on mindreading?

As predicted, both Italians and Germans scored higher on the RMET than Palestinians, and Germans scored higher

Table 4
Results of the Hierarchical Regression Analysis Predicting Mindreading

	B	*SE B*	*β*	R^2	ΔR^2	ΔF
Model 1	–	–	–	.135	.135	38.8**
(Intercept)	22.59	0.30	–	–	–	–
Age (0 = *younger*, 1 = *older*)	1.52	0.35	0.18**	–	–	–
Gender (0 = *male*, 1 = *female*)	2.65	0.35	0.32**	–	–	–
Model 2				.308	.172	61.5**
(Intercept)	20.53	0.33	–	–	–	–
Age (0 = *younger*, 1 = *older*)	1.51	0.31	0.18**	–	–	–
Gender (0 = *male*, 1 = *female*)	2.66	0.31	0.32**	–	–	–
European (0 = *no*, 1 = *yes*)	2.96	0.36	0.35**	–	–	–
Germanic (0 = *no*, 1 = *yes*)	1.16	0.41	0.12*	–	–	–
Model 3	–	–	–	.313	.005	1.8
(Intercept)	19.65	1.28	–			
Age (0 = *younger*, 1 = *older*)	1.52	0.31	0.18**	–	–	–
Gender (0 = *male*, 1 = *female*)	2.60	0.31	0.31**	–	–	–
European (0 = *no*, 1 = *yes*)	3.05	0.37	0.36**	–	–	–
Germanic (0 = *no*, 1 = *yes*)	1.23	0.41	0.13*	–	–	–
IND	−0.16	0.23	−0.03	–	–	–
COL	0.40	0.22	0.07	–	–	–

Note. IND = individualism; COL = collectivism.
*$p < .01$. **$p < .001$.

than Italians, confirming the hypothesis that there are cultural ingroup effects on mindreading. To be sure, the findings potentially admit alternative explanations, as we used translations of the original RMET constructed in the United Kingdom and did not check whether analogous effects would be found using ethnic versions of the RMET, with stimuli selected and constructed in Palestine, Italy, and Germany, respectively. However, there are at least two circumstances that let alternative explanations appear less parsimonious. First, we found that all three predicted effects; that is, we found the predicted effect concerning Europeans versus Palestinians with two different European samples (one from Germany and the other from Italy), and we also found the predicted effect concerning Germans versus Italians. Second, the observed effects could not be accounted for in terms of the value dimensions that are most often referred to that would account for cultural differences (i.e., in terms of individualism and collectivism). Overall, the hypothesized cultural ingroup effects on mindreading seem therefore to be the most parsimonious explanation of the observed effects.

On a critical note, note that the dichotomic distinctions between ingroups and outgroups are actually rather arbitrary and ambiguous because they depend on choosing one often fuzzy reference category rather than some other category (e.g., "Caucasian" rather than "European"). Another, perhaps more appropriate way to account for cross-cultural differences might be to use measures of cultural distance. Unfortunately, however, comprehensive measures of cultural distance are available thus far only for a rather limited number of culture-bearing units, such as European countries (De Santis et al., 2016).

Are the known age and gender differences in mindreading moderated by culture?

As predicted, the moderating effects of culture concerned the effects of gender rather than those of age. In fact, only the effects of gender, but not those of age, differed significantly between Palestinians, Italians, and Germans. However, our expectation that gender differences in mindreading would be smaller in Italians and Germans than in Palestinians was not confirmed. Contrary to our expectations, in fact, gender differences in mindreading were smaller in Palestinians than in Italians and Germans.

Nevertheless, this unexpected result does not necessarily falsify our hypothesis that larger differences in gender roles are associated with larger gender differences in mindreading. In fact, there seem to be at least three ways to explain the discrepancy between predicted and observed results. First, the gender-equality measures we employed to assess differences in gender roles might not capture those aspects of gender roles that affect mindreading. In fact, gender-equality measures typically assess either subjective opinions or objective facts about economic, political, and educational gender equality, but do not directly assess gender roles concerning interpersonal behavior. Second, supposing that gender-equality measures appropriately reflect the relevant differences in gender roles, gender roles might rather affect ingroup than outgroup mindreading. In fact, gender roles concerning interpersonal behavior are likely to determine experience and expertise with ingroup rather than with outgroup people. Finally, cultural bias toward the target culture (United Kingdom) might negatively affect mindreading in both male and female adolescents, and due to larger cultural distance, it might be stronger in Palestinians than in Italians or Germans, so that gender differences are not so obvious.

In any case, our results provide the first behavioral evidence that the effects of gender on mindreading may be moderated by cultural variables. However, more research is needed to better understand the mechanisms underlying the moderating effect of cultural variables, especially the role of gender roles and outgroup bias within and across cultures.

Is mindreading related to individualism and collectivism?

Any interpretation of our results concerning the relation of mindreading to individualism and collectivism is limited by the fact that the constructs of IND and COL might not be fully equivalent in the three samples because they were uncorrelated only in the Italian and the German samples, but positively correlated in the Palestinian sample (see the Instruments section). Keeping this limitation in mind, our results partly confirm our expectation that IND and COL would be differently related to mindreading and that rather COL than IND would be related to mindreading, but they contradict our expectation that the effects of IND and COL on mindreading would be observed in each of our three samples. In fact, we found that COL was related to mindreading only in the Palestinian sample, but unrelated in the Italian and German samples, whereas IND was unrelated in all three samples.

Given that mindreading was positively correlated with COL in the Palestinian sample, our results do not support

the prediction based on Elfenbein and Ambady's (2003) hypothesis, according to which mindreading would tend to be negatively correlated with COL whereas they confirm the prediction based on Vu et al.'s (2017) hypothesis, according to which mindreading would tend to be positively correlated with COL. However, as it stands, Vu et al.'s hypothesis does not account for cross-cultural differences in the relation between COL and mindreading. Our finding that COL was positively correlated with mindreading only in the Palestinian sample, but unrelated in the Italian and German samples, needs therefore to be explained in terms of some additional factor.

Vu et al. (2017) themselves suggested that potential cross-cultural differences in the relation between COL and mindreading might be due to country-level differences in COL. Given that in our study Palestinians scored overall higher on COL than Italians or Germans, our finding that COL was positively correlated with mindreading only in the Palestinian sample might therefore be explained in terms of country-level differences in COL. However, there seems to be no straightforward reason for why the relation between COL and mindreading should be more evident at higher country-level COL.

Overall, our results thus suggest that mindreading is related rather to collectivistic than to individualistic values and that this relation is moderated by cultural variables; however, more research is needed to better understand the underlying mechanisms. In particular, future research should investigate the effects of specific individualistic and collectivistic values on mindreading within and across cultures.

Conclusion

Overall, our results suggest that (a) there are cultural ingroup effects on mindreading; (b) the female superiority in mindreading is moderated by cultural factors; and (c) depending on cultural factors, individualism and collectivism may be differently related to mindreading. However, in addition to the specific limitations already mentioned in the Discussion, we note four more general limitations of our study. First, our results were obtained with adolescent samples and therefore cannot be easily generalized to adults, even though there seem to be no straightforward reasons to expect different results with adults. Second, though we took care that our samples fairly well represented the general population, we were unable to obtain individual demographic information besides age and gender, so we could not control for potentially confounding factors such as socioeconomic status. Third, the RMET is a rather imperfect measure of the ability it assesses, as evidenced by its cross-culturally poor internal consistency and unclear convergent validity. Finally, our results concern intercultural mindreading; namely, how well adolescents from three different countries recognize the mental states displayed in the eyes of people from a fourth country. Accordingly, they cannot simply be generalized to intracultural mindreading. Despite these limitations, our study provides a number of interesting new results that can give important indications for future research.

Disclosure of conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Adams, R. B., Rule, N. O., Franklin, R. G., Jr., 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. https://doi.org/10.1162/jocn.2009.21187

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or asperger syndrome. *The Journal of Child Psychology and Psychiatry*, *38*(7), 813–822. https://doi.org/10.1111/j.1469-7610.1997.tb01599.x

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*, *42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*, 155–162. https://doi.org/10.1037/h0036215

Bjornsdottir, R. T., & Rule, N. O. (2016). On the relationship between acculturation and intercultural understanding: Insight from the reading the mind in the eyes test. *International Journal of Intercultural Relations*, *52*, 39–48. https://doi.org/10.1016/j.ijintrel.2016.03.003

Bölte, S. (2005). Reading mind in the eyes test. Erwachsenen version. Retrieved from https://www.autismresearchcentre.com/arc_tests

De Santis, G., Maltagliati, M., & Salvini, S. (2016). A measure of the cultural distance between countries. *Social Indicators Research*, *126*(3), 1065–1087. https://doi.org/10.1007/s11205-015-0932-7

Elfenbein, H. A. (2007). It takes one to know one better: Controversy about the cultural ingroup advantage in communicating emotion as a theoretical rather than methodological issue. In U. Hess & P. Philippot (Eds.), *Group dynamics and emotional expression (Studies in Emotion and Social Interaction)* (pp. 51–68). Cambridge, MA: Cambridge University Press. https://doi.org/10.1017/CBO9780511499838.004

Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review*, *5*(1), 90–96. https://doi.org/10.1177/1754073912451332

Elfenbein, H. A. (2015). In-group advantage and other-group bias in facial emotion recognition. In M. K. Mandal & A. Awasthi (Eds.), *Understanding facial expressions in communication: Cross-cultural and multidisciplinary perspectives* (pp. 57–71). New York, NY: Springer Science + Business Media. https://doi.org/10.1007/978-81-322-1934-7pass:[_]4

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

Elfenbein, H. A., & Ambady, N. (2002b). Is there an in-group advantage in emotion recognition? *Psychological Bulletin*, *128*(2), 243–249. https://doi.org/10.1037//0033-2909.128.2.243

Elfenbein, H. A., & Ambady, N. (2003). Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition. *Journal of Cross-Cultural Psychology*, *34*(1), 92–109. https://doi.org/10.1177/0022022102239157

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self* (). New York, NY: Other Press. https://doi.org/10.4324/9780429471643

Fossati, A., Feeney, J., Maffei, C., & Borroni, S. (2014). Thinking about feelings: Affective state mentalization, attachment styles, and borderline personality disorder features among Italian nonclinical adolescents. *Psychoanalytic Psychology*, *31*(1), 41–67. https://doi.org/10.1037/a0033960

Franklin, R. G., Jr., Stevenson, M. T., Ambady, N., & Adams, R. B., Jr. (2015). Cross-cultural reading the mind in the eyes and its consequences for international relations. In J. E. Warnick & D. Landis (Eds.), *Neuroscience in intercultural contexts* (pp. 117–141). New York, NY: Springer Science + Business Media. https://doi.org/10.1007/978-1-4939-2260-4pass:[_]5

Gender Inequality Index (2017). Human development reports. *United Nations Development Programme*. Retrieved from http://hdr.undp.org/en/content/gender-inequality-index-gii

Global Gender Gap Report (2018). *World Economic Forum*. Retrieved from http://www3.weforum.org/docs/WEF_GGGR_2018.pdf

Hallerbäck, M. U., Lugnegård, T., Hjärthag, F., & Gillberg, C. (2009). The Reading the mind in the eyes test: Test-retest reliability of a Swedish version. *Cognitive Neuropsychiatry*, *14*(2), 127–143. https://doi.org/10.1080/13546800902901518

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*(2), 129–154. https://doi.org/10.1007/BF02172093

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, *28*, 826–839. https://doi.org/10.1037/a0030677

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values* (). Beverly Hills, CA: SAGE Publications.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Hünefeldt, T., Laghi, F., Ortu, F., & Olivetti Belardinelli, M. (2013). The relationship between 'theory of mind' and attachment related anxiety and avoidance in Italian adolescents. *Journal of Adolescence*, *36*, 613–621. https://doi.org/10.1016/j.adolescence.2013.03.012

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., … Puranen, B. (2014). *World Values Survey: All Rounds - Country-Pooled Datafile Version*. Madrid: JD Systems Institute. Retrieved from http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp

Khorashad, B. S., Baron-Cohen, S., Roshan, G. M., Kazemian, M., Khazai, L., Aghili, Z., … Afkhamizadeh, M. (2015). The "Reading the mind in the eyes" test: Investigation of psychometric properties and test-retest reliability of the Persian version. *Journal of Autism and Developmental Disorders*, *45*(9), 2651–2666. https://doi.org/10.1007/s10803-015-2427-4

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), 121–146.

Kother, U., Veckenstedt, R., Vitzthum, F., Roesch-Ely, D., Pfueller, U., Scheu, F., & Moritz, S. (2012). "Don't give me that look"—Overconfidence in false mental state perception in schizophrenia. *Psychiatry Research*, *196*, 1–8. https://doi.org/10.1016/j.psychres.2012.03.004

Lee, H.-S., Corbera, S., Poltorak, A., Park, K., Assaf, M., Bell, M. D., … Choi, K.-H. (2018). Measuring theory of mind in schizophrenia research: Cross-cultural validation. *Schizophrenia Research.*, *201*, 187–195. https://doi.org/10.1016/j.schres.2018.06.022

Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology*, *20*(1), 92–105. https://doi.org/10.1177/0022022189201006

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology*, *23*(1), 72–84. https://doi.org/10.1177/0022022192231005

Naor-Ziv, R., & Glicksohn, J. (2016). Investigating cognitive deficits as risk factors for developing eating disorders during adolescence. *Developmental Neuropsychology*, *41*(1–2), 107–124. https://doi.org/10.1080/87565641.2016.1170129

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric

analysis of the reading the mind in the eyes test: Toward a brief form for research and applied settings. *Frontiers in Psychology*, *6*(1503), 1–14. https://doi.org/10.3389/fpsyg.2015.01503

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

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Prevost, M., Carrier, M.-E., Chowne, G., Zelkowitz, P., Joseph, L., & Gold, I. (2014). The Reading the mind in the eyes test: Validation of a French version and exploration of cultural variations in a multi-ethnic city. *Cognitive Neuropsychiatry*, *19*(3), 189–204. https://doi.org/10.1080/13546805.2013.823859

Ragsdale, G., & Foley, R. A. (2011). A Maternal influence on reading the mind in the eyes mediated by executive function: Differential Parental Influences on Full and Half-Siblings. *PLoS ONE*, *6*(8), e23236. https://doi.org/10.1371/journal.pone.0023236

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, *32*(3), 189–194. https://doi.org/10.1023/A:1015497629971

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, *46*(6), 401–435. https://doi.org/10.1080/00207594.2011.626049

Schilling, L., Wingenfeld, K., Lowe, B., Moritz, S., Terfehr, K., Kother, U., & Spitzer, C. (2012). Normal mind-reading capacity but higher response confidence in borderline personality disorder patients. *Psychiatry and Clinical Neurosciences*, *66*, 322–327. https://doi.org/10.1111/j.1440-1819.2012.02334.x

Shulruf, B., Hattie, J., & Dixon, R. (2007). Development of a new measurement tool for individualism and collectivism. *Journal of Psychoeducational Assessment*, *25*(4), 385–401. https://doi.org/10.1177/0734282906298992

Shulruf, B., Alesi, M., Ciochină, L., Faria, L., Hattie, J., Hong, F., … Watkins, D. (2011). Measuring collectivism and individualism in the third millennium. *Social Behavior and Personality*, *39*(2), 173–188. https://doi.org/10.2224/sbp.2011.39.2.173

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*, 640–656. https://doi.org/10.1162/089892998562942

Taras, V., Sarala, R., Muchinsky, P., Kemmelmeier, M., Singelis, T. M., Avsec, A., … Sinclair, H. C. (2014). Opposite ends of the same stick? Multi-method test of the dimensionality of individualism and collectivism. *Journal of Cross-Cultural Psychology*, *45*(2), 213–245. https://doi.org/10.1177/0022022113509132

UNdata (2017). Popular statistical tables, country (area) and regional profiles. United Nations Statistical Division. Retrieved from http://data.un.org/Default.aspx

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the mind in the eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry*, *18*(4), 326–354. https://doi.org/10.1080/13546805.2012.721728

Vetter, N. C., Leipold, K., Kliegel, M., Phillips, L. H., & Altgassen, M. (2013). Ongoing development of social cognition in adolescence. *Child Neuropsychology*, *19*(6), 615–629. https://doi.org/10.1080/09297049.2012.718324

Vonk, J., Mayhew, P., & Zeigler-Hill, V. (2016). Gender roles, not anatomical sex, predict social cognitive capacities, such as empathy and perspective-taking. In D. F. Watt & J. Panksepp (Eds.), *Psychology of emotions, motivations and actions. Psychology and neurobiology of empathy* (pp. 187–209). Hauppauge, NY: Nova Biomedical Books.

Voronov, M., & Singer, J. A. (2002). The myth of individualism-collectivism: A critical review. *The Journal of Social Psychology*, *142*(4), 461–480. https://doi.org/10.1080/00224540209603912

Vu, T.-V., Finkenauer, C., Huizinga, M., Novin, S., & Krabbendam, L. (2017). Do individualism and collectivism on three levels (country, individual, and situation) influence theory-of-mind efficiency? A cross-country study. *PLoS One*, *12*(8), e0183011. https://doi.org/10.1371/journal.pone.0183011

Whiten, A. (Ed.). (1991). *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (). Cambridge, MA: Basil Blackwell.