Are women the more empathetic gender? The effects of gender role expectations

Charlotte S. Löffler^{1,2} · Tobias Greitemeyer²

Accepted: 3 December 2020 / Published online: 2 February 2021
© The Author(s) 2021

Abstract

The present research aimed to extend the state of knowledge regarding the relationship between self-perceived empathy and traditional gender roles and placed particular focus on the contextual conditions under which gender differences in empathy are present, can be created, or eliminated. Across two studies, women rated themselves higher in empathy than men in all experimental conditions, whereas an objective female superiority in emotion recognition was only evident in one condition. In Study 1 ($n = 736$), using the term ‘social-analytic capacity’ instead of ‘empathic capacity’ increased gender differences in self-reported empathy and resulted in women performing better in the Eyes-test than men. In a neutral task (verbal intelligence), gender differences (in this case, a male superiority), were only found when participants believed that this task had an association with empathy. In Study 2 ($n = 701$), gender differences in self-reported empathic capacity, but not in performance in emotion recognition, increased when motivation for empathy was raised. Further, gender-role orientation mediated the association between gender and self-reported empathic capacity, whereas it did not account for the association between gender and emotion recognition. Overall, the present studies provide strong support for the idea that empathy is influenced by contextual factors and can be systematically biased by gender roles and stereotypical beliefs.

Keywords Empathy · Emotion recognition · Gender differences · Traditional gender role orientation · Stereotypes

For centuries, people have been just as fascinated as they have been convinced by the idea that men and women are different not only in physiological characteristics but also in their psychological functions (Hyde 2013). The assumption that men and women differ in their ability to be empathic—to understand and share the mental and affective states of others (Singer and Lamm 2009)—is one of these widespread stereotypical beliefs. According to this stereotype, women are depicted as more interpersonally oriented and empathetic than men (Christov-Moore et al. 2014).

Regarding the question to which extent this characterization is valid, scientific studies show inconsistent results. A large number of studies reported no significant gender difference (e.g., Kim and Lee 2010; Lamm et al. 2011), some research demonstrated higher female competence (e.g., Baron-

Cohen and Wheelwright 2004; Hall and Matsumoto 2004; Kirkland et al. 2013), whereas some studies even found male superiority (e.g., Lennon et al. 1986). In sum, the evidence to suggest a higher female competence in empathy and empathy-related constructs is scarce.

Eisenberg and Lennon’s (1983) meta-analysis identified large gender differences favoring women when empathy was measured on self-report scales, but when more objective methods were used to assess empathic abilities, such as physiological or unobtrusive observations, these differences were no longer evident. On this basis, they conclude that there may rather be a female tendency to report a stronger empathic response than an actual gender difference in empathic capacity. In sum, it appears that gender differences in self-reported empathy are largely driven by the influence of gender stereotypes. This reasoning is in line with approaches, such as social role theory (Eagly 1987), that highlight the importance of social roles and expectations for gender-typed behaviors, skills, attributes, and beliefs.

Despite the fact that there are undeniable biological differences between men and women—for example, physiology or hormonal balance—other authors have suggested that a number of gender differences, observed in psychological research,

✉ Charlotte S. Löffler
charlotte.loeffler@uni-koeln.de

¹ Present address: Social Cognition Center Cologne, University of Cologne, Richard-Strauss-Str. 2, 50931 Köln, Germany

² University of Innsbruck, Innsbruck, Austria

could rather arise from cultural expectations and stereotypical beliefs than from innate attributes (e.g., Berman 1980; Hodges et al. 2011; Ickes et al. 2000; Rand et al. 2016; Thomas and Maio 2008). According to these authors, there is evidence suggesting that the division into masculine and feminine characteristics gives rise to a set of expectations that may affect the motivation to report and demonstrate certain behaviors related to empathy. Consequently, psychological gender differences could be a result of the tendency to imitate same-gender models because there are gender-specific rewards and punishments for certain behaviors (Hyde 2013).

Evidence for this hypothesis comes from several studies in the field of empathic accuracy, in which participant's precision to assess the thoughts and feelings of others is evaluated. After Graham and Ickes (1997) conducted seven studies without detecting any gender differences, three subsequent studies indicated a higher empathic accuracy of female participants. These findings were probably the result of a small change in the instrument they used: The revised instrument required the participants to estimate the degree of accuracy in inferring the other person's thoughts and feelings instead of just inferring the general emotional tone as positive, negative, or neutral. Graham and Ickes concluded that the first instrument was less obvious in focusing on the measurement of empathy than their slightly modified version. A further meta-analysis by Ickes et al. (2000) demonstrated that women's higher accuracy was only present in situations where the participants had to estimate their level of accuracy after each task. Based on the theory of Eisenberg and Lennon (1983), Ickes et al. (2000) concluded that the observed gender differences would rather reflect the differential motivation of men and women than actual disparities in empathic ability.

Subsequently, Klein and Hodges (2001) demonstrated that a female advantage within empathic accuracy could only be determined if the participants had been presented with the task to estimate their empathetic sympathy towards the target person. When the participants were rewarded financially for their accuracy, the performance of both genders increased, and gender differences were eliminated completely. Considering these results, it seems possible that women and men need different motivators to develop their full empathic potential.

In another study (Thomas and Maio 2008), men were not motivated solely by monetary rewards, but also by a specific reframing of the desirability and usefulness of empathic abilities. The authors were able to demonstrate that male participants assessed the emotions of others more accurately if they believed that higher empathy would make them more attractive to women. On the other hand, there was no effect when men were asked to disprove the stereotype that depicts them as poor mind-readers. It can be assumed that the accuracy with which people can assess the feelings and thoughts of others is further affected by other motivational factors, such as the

attractiveness of the target and the degree of personal interest in it (Ickes et al. 1990).

Other motivational aspects of expressing higher empathic capacity are deeply rooted beliefs about men and women that are referred to as stereotypes. There is evidence to suggest that women are more concerned than men about how empathic they appear to others—and probably also themselves (Eisenberg and Lennon 1983). Since being nurturant and interpersonally oriented are both part of the stereotypical feminine role, women likely have a higher motivation to present themselves as empathic irrespective of their actual traits (Thomas and Maio 2008). This implies that, in situations when gender-role expectations are made salient, a female superiority may reflect the higher feminine motivation to present oneself consistently with the stereotype (Ickes et al. 2000).

In this regard, the magnitude of many gender-specific differences identified through psychological research varies greatly and is determined by several influencing factors. Accordingly, large gender differences can be both created and erased by the context (Hyde 2013). As a prominent example, Spencer et al. (1999) could demonstrate that a gender difference in mathematical abilities, which is often assumed stereotypically, was only evident if the stereotype was specifically referred to in the experiment. When, on the other hand, participants were informed that previous experiments had not yielded any difference between men and women, no gender effects were observed.

How important the context can be in the assessment of empathy demonstrates an experiment by Nanda (2013). In line with the results of Eisenberg and Lennon (1983), a gender difference in self-reported empathy was only conspicuous when participants were aware that they were evaluated in their empathic capacity. However, no significant gender difference was found when participants were led to believe that the task assessed social abilities in general.

Another crucial factor was shown to be the traditional gender-role orientation. In a study by Karniol et al. (1998), self-reported empathy had no association with masculinity, but was highly correlated with femininity. When gender-role orientation was included as a covariate, the main effect of biological sex on empathy was no longer significant. Accordingly, the authors concluded that it may not be the biological sex that determines the level of empathic capacity, but the female gender-role orientation.

Based on these previous results and theoretical considerations (such as social role theory), the present research focuses on the relationship between traditional gender-role orientation and a potential gender-specific motivation for empathy, in that we aimed to systematically evaluate the interactions between gender, self-reported empathic capacity, and objective performance in emotion recognition under differing experimental variations. In Study 1, we explicitly informed the participants about the fact that empathy was assessed or led them to

believe that the task measured ‘social-analytic abilities’—a term that we expected to be less deterrent for male participants. Additionally, we used a task that evaluated verbal intelligence as a neutral instrument to see if participants would demonstrate higher motivation, respectively higher empathy scores, when they believed that there is a relationship between empathy and verbal intelligence. If our reasoning is correct, gender differences in self-reported empathic capacity (Hypothesis 1a) and objective emotion recognition (Hypothesis 1b) will be smaller in the social-analytic condition, whereas there will be gender differences in the neutral task when participants believe that verbal intelligence is associated with empathy (Hypothesis 1c).

The aim of Study 2 was to induce a different motivation for empathy across experimental conditions. Participants in one condition received a framing of empathy as an essential skill that yields numerous benefits, whereas the other condition functioned as a control group. We expected to observe smaller gender differences in both self-reported empathic capacity (Hypothesis 2a) and objective emotion recognition (Hypothesis 2b) in the motivation condition. In addition, we hypothesized that the relationship between biological sex and both measurements of empathy would be mediated by traditional gender-role orientation (Hypothesis 3a, Hypothesis 3b).

The data were collected in two online questionnaires, conducted at the same time. One link to the two online questionnaires, interconnected by a random generator, was distributed via Facebook and SurveyCircle. Additionally, students from other disciplines than psychology were invited via a university mailing list. No compensation for participation was provided, but participants could register anonymously to enter a prize draw for three Amazon vouchers. No formal approval from an ethics committee was required at our university, under the provision that the research was in line with the guidelines of the German Psychological Society. Participation in the study was voluntary, and the data were collected and analyzed anonymously. Only participants older than 18 years were allowed to participate.

Since effect sizes vary greatly in previous research, a small effect size of $f = .16$ was used as an estimator of the expected effect size. Using G*Power (Faul et al. 2009), the required sample size for a two-way ANOVA with a power of .80 was $N = 644$. Hence, for the final samples, the number of participants required was calculated to be 644 for each study. When testing a priori predictions, as it is the case in the present research, planned comparisons provide the best statistical test of possible mean differences across conditions (Rosenthal and Rosnow 1985; Steiger 2004). Therefore, aside from employing an overall 2 X 2 analysis of variance, we additionally relied on the statistically more powerful planned contrasts. In both studies, the majority of the data were assessed in a forced-choice format. However, for the performance tasks, 75% of missing answers were tolerated under the

assumption that a partial completion of the test could reflect not just a differential capacity but also a differential motivation of the participants to present themselves as empathic. For the same reason, there was no time limit set to complete the tasks.

Study 1

The aim of Study 1 was to examine to which degree gender-role expectations influence self-reports of empathy and objective performance in emotion recognition. Inasmuch as the stereotype depicts women as more empathetic and better mind-readers, we presumed that females have a motivational advantage when it comes to displaying empathic behaviors.

To test this assumption, participants were randomly assigned to two experimental conditions. In both conditions, self-reports of empathy and performance in emotion recognition were assessed. Whereas participants in the empathy condition were explicitly informed about the fact that the tasks assessed empathy (“in the subsequent tasks, your empathic capacity will be assessed”), participants in the second condition were informed that the tasks assessed ‘social-analytic abilities’ (“in the subsequent tasks, your social-analytic capacity will be assessed”), because we expected this term to be less deterrent for male participants. Based on the results of Eisenberg and Lennon (1983), we predicted that gender differences in self-reported empathic capacity (Hypothesis 1a) and objective performance in emotion recognition (Hypothesis 1b) would be smaller when individuals are not aware that empathy is measured. In both conditions, verbal intelligence was assessed as a neutral capacity and was not expected to have a reliable association with empathy (e.g., Koch et al. 2007). Whereas participants in the social-analytic condition were informed about the true nature of this task, participants in the empathy condition were told that empathy and verbal intelligence had a strong positive association. We expected to observe differences in gender-specific performance between the two conditions if women indeed have a higher motivation in proving their empathic capacity (Hypothesis 1c). To examine the possible mediating influence of gender-role expectations, we assessed the traditional gender-role orientation of all participants.

Method

Participants In Study 1, 80 participants were excluded from the analyses because they were non-native speakers of German—a condition that had to be fulfilled to complete the verbal intelligence task. Further, nine participants were excluded because they defined their gender as neither male nor female. The final sample comprised 736 participants (494 females, 242 males; mean age = 25.5, $SD = 8.3$).

Procedure and Materials Participants were randomly assigned to either be instructed that their empathic capacity or their social-analytic capacity would be assessed. Subsequently, empathy was assessed on a self-report scale. To this end, the German short version (Samson and Huber 2010) of the Empathy Quotient (EQ) by Baron-Cohen and Wheelwright (2004) was employed. It contained 13 items ($\alpha = .85$). Sample item: “I can easily tell if someone else wants to enter a conversation.” All items were assessed on a scale from 1 (“strongly agree”) to 4 (“strongly disagree”). According to Baron-Cohen and Wheelwright (2004), “strongly agree” responses scored 2 points, “slightly agree” responses scored 1 point, and the remaining options scored 0 points. These scores were then summed. Performance in emotion recognition was assessed with the German version of the “Reading the Mind in the Eyes” Test (Baron-Cohen et al. 1997; revised version Baron-Cohen et al. 2001; German version Bölte 2005). It contains 36 black and white photographs of the eye region ($\alpha = .61$), providing four different mental state terms to select the correct answer from (e.g., “playful”, “comforting”, “irritated”, “bored”).

As a manipulation check, participants were asked to report if either empathic capacity or social-analytic capacity had been assessed in the previous tasks. The manipulation check was partially successful, in that 66.8% in the empathy condition and 69.6% in the social-analytic condition were able to provide the expected answer. Given that the pattern of the main findings remained unchanged when excluding participants that had not passed the manipulation check, the following analyses are based on all participants.

Subsequently, participants were again provided with condition-specific information. Whereas participants in the empathy condition were informed that the following task would assess their verbal intelligence and there would be a strong association between empathy and verbal intelligence, participants in the social-analytic condition were told that the following task assessed verbal intelligence. Subsequently, verbal intelligence was assessed with the Mehrfach-Wortschatz-Intelligenztest (MWT-B) by Lehrl (2005), which is exclusively available in German. This performance test originally contains 37 items in the course of which the degree of difficulty rises gradually ($\alpha = .64$). To keep the task as short as possible, only the last 15 items were used in the present research. Every item in the MWT-B consists of five words (e.g., “Tuhl – Lar – Lest – Dall – Lid”), from which only one word is an existing term in German; the other four are nonsense words. Participants are required to identify the true term, in this case, “Lid” (also translated as a/the lid). For the subsequent assessment of gender role orientation, the Traditional Masculinity and Femininity Scale (TMF) by Kachel et al. (2016) was employed. Published both in English and in German, this scale contains 6 items ($\alpha = .94$), which can be completed on a scale from 1 (“very masculine”) to 7 (“very feminine”).

Sample item: “I consider myself as...”. After providing their demographic information, participants were asked to guess the purpose of the study. (None of the participants was able.) They were also asked to voluntarily leave their e-mail addresses for the prize draw and pointed to the possibility of contacting the study leader.

Results

Descriptive statistics and intercorrelations of the measures are reported in Table 1. Overall, across both conditions, gender was weakly negatively associated with self-reported empathy and emotion recognition, suggesting a female superiority. In contrast, verbal intelligence was positively associated with gender, suggesting a higher male performance. As expected, gender-role orientation showed a strong negative association with gender. Self-reported empathy was positively associated with emotion recognition and feminine gender-role orientation, whereas emotion recognition was associated with verbal intelligence and, again, feminine gender-role orientation, although this second correlation was smaller than the first one. In turn, gender-role orientation was negatively associated with verbal intelligence, suggesting an association with masculinity.

To examine whether gender effects in self-reported empathy would be smaller when empathy was assessed as a social-analytic capacity (Hypothesis 1a), a two-way ANOVA was performed on the data. There was a significant main effect of gender, $F(1, 732) = 14.69, p = .000, \eta_p^2 = .02$; female participants ($N = 494, M = 14.05, SD = 5.32$) reported higher empathic capacities than male participants ($N = 242, M = 12.44, SD = 5.13$). The main effect of condition was non-significant, $F(1, 732) = .16, p = .688, \eta_p^2 = .00$, reflecting that the empathic condition ($N = 371, M = 13.60, SD = 5.31$) and the social-analytic condition ($N = 365, M = 13.44, SD = 5.31$) did not differ, as well as there was a non-significant interaction effect between gender and condition, $F(1, 732) = .64, p = .425, \eta_p^2 = .00$.

As noted above, the statistically more powerful planned comparisons provide a better test of our hypotheses (Rosenthal and Rosnow 1985; Steiger 2004). To this end, we computed a new variable, distinguishing between the groups female/empathic ($N = 262, M = 13.97, SD = 5.35$), male/empathic ($N = 109, M = 12.72, SD = 5.13$), female/social-analytic ($N = 232, M = 14.14, SD = 5.30$), and male/social-analytic ($N = 133, M = 12.22, SD = 5.13$). There was homogeneity of the error variances, as assessed by Levene’s test ($p > .05$). The four groups significantly differed, $F(3, 732) = 5.28, p = .001, \eta^2 = .02$. Planned contrasts demonstrated a significant difference between female/empathic and male/empathic, $t(732) = 2.10, p = .036, r = .12$. We also found a significant difference between female/social-analytic and

Table 1 Means, Standard Deviations, Bivariate Correlations with internal consistency reliabilities (Cronbach's Alpha) in the diagonal (Study 1)

	<i>M</i>	<i>SD</i>	1	2	3	4	5
1. Gender ^a	—	—					
2. Self-reported empathy	13.52	5.31	-.14**	(.85)			
3. Emotion recognition	24.24	4.23	-.13**	.20**	(.61)		
4. Verbal intelligence	8.29	2.61	.13**	.01	.08*	(.64)	
5. Gender-role orientation ^b	4.48	1.61	-.85**	.17**	.10**	-.14**	(.94)

N = 736. ^a 1 = female, 2 = male. ^b 1 = very masculine – 7 = very feminine

* $p < .05$, ** $p < .01$

male/social-analytic, $t(732) = 3.35$, $p = .001$, $r = .18$. Therefore, women rated themselves significantly higher in empathy across both conditions. Unexpectedly, however, when participants believed that their social-analytic capacities were assessed, gender differences tended to be more pronounced (see Fig. 1).

To further examine whether gender effects on performance in emotion recognition would be smaller when empathy was assessed as a social-analytic capacity (Hypothesis 1b), a two-way ANOVA was conducted. Female participants ($N = 494$, $M = 24.62$, $SD = 4.00$) had higher scores than male participants ($N = 242$, $M = 23.47$, $SD = 4.56$), $F(1, 732) = 11.54$, $p = .001$, $\eta_p^2 = .02$. There was a non-significant main effect of condition, $F(1, 732) = 0.57$, $p = .449$, $\eta_p^2 = .00$, reflecting that the empathic condition ($N = 371$, $M = 24.33$, $SD = 4.14$) and the social-analytic condition ($N = 365$, $M = 24.14$, $SD = 4.31$) did not differ. The interaction effect between gender and condition was non-significant, $F(1, 732) = 1.61$, $p = .205$, $\eta_p^2 = .00$.

To further test for possible mean differences across conditions, planned contrasts were performed. Again, we distinguished between the four groups female/empathic ($N = 262$, $M = 24.54$, $SD = 3.90$), male/empathic ($N = 109$, $M = 23.83$, $SD = 4.65$), female/social-analytic ($N = 232$, $M = 24.71$, $SD = 4.12$), and male/social-analytic ($N = 133$, $M = 23.17$, $SD = 4.48$). There was homogeneity of the error variances, as assessed by Levene's test ($p > .05$). The four groups significantly differed, $F(3, 732) = 4.65$, $p = .003$, $\eta^2 = .02$. Planned contrasts demonstrated a non-significant difference between female/empathic and male/empathic, $t(732) = 1.47$, $p = .142$, $r = .08$. In contrast, there was a significant difference between female/social-analytic and male/social-analytic, $t(732) = 3.38$, $p = .001$, $r = .18$. Accordingly, the pattern of the first analysis, evaluating gender effects on self-reported empathy (Hypothesis 1a), was replicated for the performance in emotion recognition. In this case, there was no significant difference in the performance of males and females when participants were informed that their empathic capacities would be assessed, whereas women performed significantly better in

emotion recognition than men when they believed that their social-analytic capacities were assessed. Therefore, contrary to our expectations, gender differences were not smaller when empathy was assessed as a social-analytic capacity, but actually more pronounced (Fig. 1).

To determine gender effects on a neutral performance task (verbal intelligence) when participants believed that this task was strongly associated with empathy (Hypothesis 1c), a further two-way ANOVA was performed on the data. There was a significant main effect of gender, $F(1, 732) = 12.51$, $p = .000$, $\eta_p^2 = .02$; female participants ($N = 494$, $M = 8.06$, $SD = 2.59$) had lower scores than male participants ($N = 242$, $M = 8.76$, $SD = 2.59$). The main effect of condition was non-significant, $F(1, 732) = 1.13$, $p = .288$, $\eta_p^2 = .00$, reflecting that the empathy condition ($N = 371$, $M = 8.30$, $SD = 2.71$) and the verbal intelligence condition ($N = 365$, $M = 8.28$, $SD = 2.50$) did not differ. There was a significant interaction effect between gender and condition, $F(1, 732) = 4.06$, $p = .044$, $\eta_p^2 = .01$.

To further test for possible mean differences across conditions, planned contrasts were analyzed. We distinguished between the groups female/empathy ($N = 262$, $M = 7.97$, $SD = 2.66$), male/empathy ($N = 109$, $M = 9.10$, $SD = 2.68$), female/verbal intelligence ($N = 232$, $M = 8.16$, $SD = 2.51$), and male/verbal intelligence ($N = 133$, $M = 8.47$, $SD = 2.49$). There was homogeneity of the error variances, as assessed by Levene's test ($p > .05$). The four groups significantly differed, $F(3, 732) = 5.32$, $p = .001$, $\eta^2 = .02$. Planned contrasts demonstrated a significant difference between female/empathy and male/empathy, $t(732) = -3.84$, $p = .000$, $r = .21$. In contrast, there was a non-significant difference between female/verbal intelligence and male/verbal intelligence, $t(732) = -1.10$, $p = .271$, $r = .06$. Hence, male participants performed better than female participants on the verbal intelligence task when they were led to believe that empathy and verbal intelligence were associated. On the other hand, there were no significant gender differences when participants were solely told that we would assess their verbal intelligence. While we originally expected women to outperform men

Fig. 1 The influence of the experimental variations on the magnitude of gender differences across measures of self-reported empathy, performance in emotion recognition, and performance on a neutral task (verbal intelligence)

when they believed that empathy and verbal intelligence were associated, men actually achieved higher results in this condition (Fig. 1).

Analyses for the possible mediating impact of traditional gender-role expectations (Hypothesis 3a, Hypothesis 3b) will be reported for Studies 1 and 2 combined (see below).

Discussion

Study 1 lends initial support for the hypothesis that self-reported empathy and even performance in emotion recognition are subject to contextual influences. While women rated themselves significantly higher in empathic capacity, there was no significant gender difference in performance in emotion recognition when participants were told that empathy was measured. However, when they were told that their social-analytic capacity would be assessed, we observed more pronounced gender differences both in self-reported empathy and performance in emotion recognition, indicating a female superiority. This result is remarkable, insofar as we expected the term social-analytic to be less deterrent for male participants and less influenced by the subtle gender stereotype that is associated with the term empathy. In fact, using the word social-analytic increased and even created gender differences not only in self-reported capacity but also in performance in emotion recognition. In addition, we exclusively detected gender differences in the neutral task (verbal intelligence) when participants believed that it was linked to empathy. Unexpectedly, in this case, men outperformed women by solving significantly more items on the verbal intelligence task. The fact that verbal intelligence was weakly associated with masculine gender role orientation across the full sample cannot provide an explanation for this effect because gender differences were only evident in the condition that had received the manipulation. Overall, Study 1 provides evidence that both self-reported empathy and objective performance in emotion recognition can be influenced through a subtle experimental variation and that even a presumed association with the concept of empathy might induce gender differences.

Study 2

Study 2 examined the impact of an experimentally induced external motivation on empathy. As noted above, we assumed that women have a higher motivation to demonstrate their

empathic capacity, in order to present themselves consistently with the common stereotype (Eisenberg and Lennon 1983). Therefore, it seems possible that men need different motivators to demonstrate their full empathic potential, as it was shown for financial rewards (Klein and Hodges 2001) and reframing empathy as a desirable and useful ability (Thomas and Maio 2008). To test this idea, in one condition, participants received a short text that emphasized the importance of empathy in daily life and depicting numerous benefits empathic persons would have. We expected this condition to produce smaller gender differences in both self-reported empathy (Hypothesis 2a) and objective performance in emotion recognition (Hypothesis 2b), compared to a control condition where no benefits of being empathic were mentioned.

Method

Participants The structure of Study 2 was similar to Study 1, except for the employment of different stimuli and control tasks. Further, the MWT-B (to assess verbal intelligence) was not employed. Two participants were excluded from the analyses because they defined their gender as neither male nor female. Furthermore, three participants were excluded because they provided fewer than 25% of the answers for the Eyes-Test. The final sample of Study 2 comprised 701 participants (478 females, 223 males; mean age = 25.9, $SD = 8.6$).

Procedure and Materials Participants were randomly assigned to one of two experimental conditions. In the motivation condition, participants got informed that the following task was meant to assess their empathic capacity. In the subsequent briefing, empathy was framed as a highly important skill in social interaction. Participants were told that studies had shown empathic individuals to be more attractive to the opposite gender, to have closer friendships, less mental problems, and even receive a higher salary. In contrast, participants in the control condition were solely told that the following task assessed their empathic capacity. Subsequently, participants completed the 13 items from the EQ ($\alpha = .85$) and the 36 items from the Eyes-Test ($\alpha = .56$).

After the tasks, we included a manipulation check, consisting of 3 items ($\alpha = .72$). Participants were asked if they believed that (1) empathy was a desirable feature, (2) empathy would yield benefits for their life, and (3) empathic persons would be perceived more positively. These items were assessed on a scale from 1 (“*strongly disagree*”) to 4 (“*strongly agree*”). As there were no significant differences between the motivation ($M = 3.61$, $SD = 0.45$) and the control condition ($M = 3.56$, $SD = 0.49$), it is unclear whether the manipulation had the intended effects. We will return to the issue of the failed manipulation check in the discussion of this study.

The control-scale was followed by the TMF ($\alpha = .94$) and the assessment of demographics. Participants were asked to

Table 2 Means, Standard Deviations, Bivariate Correlations with internal consistency reliabilities (Cronbach’s Alpha) in the diagonal (Study 2)

	<i>M</i>	<i>SD</i>	1	2	3	4
1. Gender ^a	–	–				
2. Self-reported empathy	13.66	5.23	-.18**	(.85)		
3. Emotion recognition	24.51	3.93	-.09*	.16**	(.56)	
4. Gender-role orientation ^b	4.44	1.54	-.83**	.19**	.10*	(.94)

N = 736. ^a 1 = female, 2 = male. ^b 1 = very masculine – 7 = very feminine

* $p < .05$, ** $p < .01$

guess the purpose of the study. (None of the participants did it correctly.) Subsequently, they could leave their e-mail address for the prize draw and were pointed to the possibility of contacting the study leader.

Results

Descriptive statistics and intercorrelations of the measures are reported in Table 2. As in Study 1, gender was negatively associated with self-reported empathy, emotion recognition, and gender role orientation. Again, self-reported empathy was associated with performance in emotion recognition and feminine gender role orientation. Further, emotion recognition was associated with female gender role orientation.

To examine whether gender effects in self-reported empathy would be smaller in the motivation condition (Hypothesis 2a), a two-way ANOVA was conducted. There was a significant main effect of gender, $F(1, 697) = 23.48, p < .001, \eta_p^2 = .03$; female participants ($N = 478, M = 14.30, SD = 5.16$) reported higher empathic capacities than male participants ($N = 223, M = 12.28, SD = 5.12$). The main effect of condition was non-significant, $F(1, 697) = 0.38, p = .538, \eta_p^2 = .00$, reflecting that the motivation condition ($N = 353, M = 13.85, SD = 4.97$) and the control condition ($N = 348, M = 13.47, SD = 5.48$) did not differ. The interaction between gender and condition was also non-significant, $F(1, 697) = 0.79, p = .373, \eta_p^2 = .00$.

To test for possible mean differences across conditions, planned comparisons were performed. As in Study 1, a new variable was computed, distinguishing between the four groups female/motivation ($N = 240, M = 14.62, SD = 4.86$), male/motivation ($N = 113, M = 12.22, SD = 4.84$), female/control ($N = 238, M = 13.99, SD = 5.45$), and male/control ($N = 110, M = 12.34, SD = 5.41$). There was homogeneity of the error variances, as assessed by Levene's test ($p > .05$). The

four groups significantly differed, $F(3, 697) = 8.45, p < .001, \eta^2 = .04$. Planned contrasts demonstrated that female/motivation was significantly different from male/motivation, $t(697) = 4.08, p < .001, r = .22$. There was also a significant difference between female/control and male/control, $t(697) = 2.78, p = .006, r = .15$. Therefore, there were gender effects in self-reported empathic capacity in both conditions and gender effects were not smaller but slightly more pronounced in the motivation condition.

To examine whether gender effects in emotion recognition would be smaller in the motivation condition compared to the control condition (Hypothesis 2b), a two-way ANOVA was conducted. There was a significant main effect of gender, $F(1, 697) = 5.80, p = .016, \eta_p^2 = .01$; female participants ($N = 478, M = 24.75, SD = 3.84$) performed better than male participants ($N = 223, M = 23.99, SD = 4.07$). There was a non-significant main effect of condition, $F(1, 697) = 0.90, p = .343, \eta_p^2 = .00$, reflecting that the motivation condition ($N = 353, M = 24.68, SD = 3.87$) and the control condition ($N = 348, M = 24.34, SD = 3.98$) did not differ. The interaction between gender and condition was also non-significant, $F(1, 697) = 0.19, p = .660, \eta_p^2 = .00$.

To further test for possible mean differences across conditions, planned contrasts were analyzed. Again, it was distinguished between the groups female/motivation ($N = 240, M = 24.98, SD = 3.66$), male/motivation ($N = 113, M = 24.07, SD = 4.26$), female/control ($N = 238, M = 24.53, SD = 4.01$), and male/control ($N = 110, M = 23.91, SD = 3.88$). There was homogeneity of the error variances, as assessed by Levene's test ($p > .05$). The four groups did not significantly differ, $F(3, 697) = 2.47, p = .061, \eta^2 = .01$. However, planned contrasts demonstrated that female/motivation was significantly different from

Fig. 2 Mediation of the relationship between gender and self-reported empathy (top) and emotion recognition (bottom) by gender role orientation. Coefficients in parentheses are parameter estimates containing both predictors. * $p < .05$, ** $p < .01$, *** $p < .001$

male/motivation, $t(697) = 2.03$, $p = .043$, $r = .12$. In contrast, there was a non-significant difference between female/control and male/control, $t(697) = 1.38$, $p = .167$, $r = .08$.

To test whether gender role orientation would account for the association of gender and self-reported empathy (Hypothesis 3a), the data of both studies were combined, because the structure of Study 1 and 2 were nearly identical. To this end, the PROCESS macro for SPSS (Hayes 2018) was employed. This analysis revealed a significant negative total effect between gender and self-reported empathy ($\beta = -1.81$, $p < .001$). Gender was significantly negatively associated with gender role orientation, while gender role orientation was also associated with self-reported empathy. Whereas the direct effect between gender and self-reported empathy was no longer significant ($\beta = -.44$, $p = .41$), the indirect effect of gender on self-reported empathy via gender role orientation was significant $\beta a * b = -1.37$. The bootstrap confidence interval of the indirect effect (95% CI $[-2.26; -.45]$) does not include 0, suggesting that the association of gender and self-reported empathy was mediated by gender role orientation. This mediation effect, based on regression analyses, is shown in Fig. 2.

To test if this pattern could be replicated for emotion recognition (Hypothesis 3b), a further mediation analysis was conducted. It revealed a significant negative total effect between gender and the number of solved items in the Eyes-Test ($\beta = -.97$, $p < .001$). Gender was significantly negatively associated with gender role orientation, while gender role orientation was not associated with performance in emotion recognition. The direct effect between gender and performance in emotion recognition was significant ($\beta c' = -.86$, $p = .039$), whereas the indirect effect of gender on performance in emotion recognition via gender role orientation was non-significant ($\beta a * b = -.11$). Since, the bootstrap confidence interval of the indirect effect (95% CI $[-.90; .65]$) includes 0, the association of gender and performance in emotion recognition was not mediated by gender role orientation (Fig. 2).

Discussion

Study 2 provides evidence that implicit gender role orientation influences self-reported empathy and that a motivational reframing of empathy as a desirable and useful ability can have an impact on gender differences in empathy. In this regard, it is important to keep in mind that our manipulation check was not successful in that participants in the motivation condition were *not* more likely to perceive empathic capacity to be an important skill in social interactions than were participants in the control condition. Given that the effectiveness of our motivation manipulation could not be established, we have to concede that no strong causal conclusions are warranted and that future research employing other ways to induce a motivation to appear empathic would be very welcome.

In addition, we found the relationship between gender and self-reported empathy fully mediated by gender role orientation, whereas gender role orientation did not account for the relationship between gender and emotion recognition. Therefore, it indeed appears that it is not the biological sex that determines how empathic people rate themselves but the gender role orientation and the expectations that come with feminine and masculine gender identities. In contrast, gender role orientation appears to have no detectable influence on the ability to recognize and determine emotions in others.

General Discussion

The present studies advance our knowledge regarding the relationship between the concept of empathy and traditional gender roles and demonstrates how a slight linguistic variation in one term (Study 1) or a motivational reframing of empathy (Study 2) can effectively create more pronounced gender differences. As previous research has shown, gender differences are most evident when empathy is assessed on self-report scales or when gender role expectations are made salient, but these differences become smaller or completely undetectable when more objective measurements are used (Eisenberg and Lennon 1983; Ickes et al. 2000). In line with these results, women rated themselves as significantly more empathic than men in all four conditions, while a female superiority in emotion recognition was only evident in the condition where empathy was referred to as ‘social-analytic capacity’. On this basis, the present studies lend strong support for the idea that there is a female tendency to report a stronger empathic response rather than an actual difference in male and female ability, as a number of authors have already suggested (e.g., Berman 1980; Eisenberg and Lennon 1983; Hodges et al. 2011; Ickes et al. 2000; Thomas and Maio 2008).

However, the assumption that gender differences in self-reported empathic capacity and performance in emotion recognition would be smaller when participants were not aware of the true nature of the tasks (Hypothesis 1a, Hypothesis 2a) could not be verified, as our experimental setup could not conceal the fact that empathy was measured by using the term ‘social-analytic capacity’. It is conceivable that, in the present case, the term ‘social-analytic’ appeared too sophisticated or even artificial and, as a consequence, had a deterrent effect on some participants, whereas, in the female sample, it apparently raised motivation for empathy. On the other hand, it is also conceivable that the term ‘social’ has a higher emotional connotation than the term ‘analytic’—so it might have overshadowed it. As a result, gender differences in both self-reported empathic capacity and objective emotion recognition were more pronounced when we used the term ‘social-analytic capacity’ compared to the term ‘empathic capacity’. Against this background, it seems reasonable to suppose that

the term ‘social-analytic’, which was originally meant to be more neutral and less influenced by stereotypical beliefs than the term ‘empathy’, had the opposite effect and created gender differences in the performance in emotion recognition that were not observable when using the term ‘empathy’.

Regarding our hypothesis that gender differences on a neutral task (verbal intelligence) are more pronounced when participants believe that it is related to empathy (Hypothesis 1c), unexpectedly, we detected a male superiority in verbal intelligence when we evoked the association with empathy. Hence, it seems possible that even a presumed association with empathy might induce differences in the performance of men and women. But remarkably, in this case, men might have had a higher motivation and outperformed women when they were led to believe that a concept, with which they were familiar with, was related with empathy. The fact that verbal intelligence was weakly associated with masculine gender role orientation across the full sample cannot provide an explanation for this effect, because gender differences were only evident in the condition that had received the manipulation.

While in Study 1 we were able to manipulate emotion recognition by using an alternative term for empathy, emotion recognition was not significantly influenced in Study 2 by using external motivators (Hypothesis 2b). This result is contrary to some previous research in the field of empathic accuracy (Klein and Hodges 2001; Thomas and Maio 2008) that demonstrated that appropriate motivators could indeed increase the performance in emotion recognition. However, regarding self-reported empathic capacity, we did demonstrate more pronounced gender-differences in the condition that had received the motivation (Hypothesis 2a). This result suggests that external motivations can indeed manipulate self-reports and lends support for the notion that the context can play a key role in self-perception. But at this point, we have to concede that the stimuli we used to raise motivation for empathy turned out to be weak, as suggested not only by the failed manipulation check, but also the fact that we could only demonstrate a small motivational effect in females.

Apart from this, our research managed to demonstrate that the association between gender and self-reported empathy was fully mediated by gender role orientation (Hypothesis 3a), whereas gender role orientation did not account for the relationship between gender and emotion recognition (Hypothesis 3b). Together with the finding that a female superiority in emotion recognition was detected in only one case when the context had been manipulated successfully, these results provide strong evidence that a female superiority in empathy and related constructs does not reflect the differential ability of men and women and may indeed be a stereotype—a stereotype that causes women to present themselves as empathic, because being caring and interpersonally oriented are part of the traditional feminine role. On the other hand, men may tend to underestimate their full empathic potential in the absence of appropriate external motivators. It is also important to point

out that in both studies there was only a moderate correlation between the self-reported empathy measure and the performance in the emotion recognition task. Taken together, the belief to be empathic may not be reflected in actual empathy.

As noted above, an important limitation of the present research is the failed manipulation check in Study 2. Hence, no strong conclusions are warranted how the motivation to appear empathic has an impact on gender differences in empathy. Furthermore, most of the present findings were small in terms of their effect sizes. In fact, analyses of variance did not reveal significant interaction effects (with one exception), but only the more statistically powerful planned comparisons yielded significant effects.

In conclusion, the present studies provide evidence that self-reported empathy and even objective performance in emotion recognition can both be influenced by the contextual setting, and that even a presumed association with the concept of empathy might induce gender differences. In addition, it was demonstrated that there is indeed a female tendency to report stronger empathic responses, while our results did not suggest a major female superiority in emotion recognition. We find it remarkable that at the present time that is characterized by reshaping traditional gender roles and societal structures empathy still appears to be perceived as a typical feminine trait. Therefore, it is questionable to use self-reports of empathy as a measure for actual empathic capacity in research. This is not only suggested by the fact that the association between gender and self-reported empathy was fully mediated by gender role orientation, but also by the weak correlation between self-reported empathy and performance in emotion recognition and that self-reported empathy was shown to be highly dependent on the experimental context. Against this background, some scientific results in this field might have been systematically biased by implicit gender stereotypes and that differences between males and females had been overestimated.

Regarding the present research, a female superiority in emotion recognition was only found in one of our experimental conditions. But even if there is indeed such a small female advantage, as Kirkland et al. (2013) and Warrier et al. (2018) suggested in their meta-analyses, it is important to keep in mind that the concept of gender differences is too narrow to map and explain the huge variety of inter-individual differences that are observable in psychological research and that a female advantage in empathy and related constructs could rather reflect a combination of biological factors, differing experience, socialization, and cultural expectations, which in turn appear to be mediated by some form of motivation (Hodges et al. 2011). The specific interactions between these factors remain to be determined by future studies. Another important question concerns the implementation of alternative instruments for measuring objective empathic responses, such as physiological or unobtrusive observations. Further, it would be of interest to address whether other constructs, that are likewise afflicted by gender stereotypes (e.g.,

emotionality, dominance, or intuitive processing), are also context-dependent and are influenced by gender role orientation in a similar way. Until then, it is important not to overemphasize these potential differences because, as Hyde (2013) has pointed out, gender similarities are as interesting and as important as gender differences.

Funding Open Access funding provided by University of Innsbruck and Medical University of Innsbruck.

Data Availability The data for both studies are publicly available here: <https://osf.io/w38bv/>.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethics Approval The authors declare that the study was performed in accordance with the ethical standards of the German Psychological Society. In such a case (when ethical guidelines are met), our university requires no formal approval from an ethics committee.

Informed Consent Participants were informed that the data are analyzed anonymously, that participation is voluntary, and that they can withdraw at any time.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34, 163–175. <https://doi.org/10.1023/b:jadd.0000022607.19833.00>.
- 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. *Journal of Child Psychology and Psychiatry*, 38, 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. *Journal of Child Psychology and Psychiatry*, 42, 241–251. <https://doi.org/10.1017/S0021963001006643>.
- Berman, P. W. (1980). Are women more responsive than men to the young? A review of developmental and situational variables. *Psychological Bulletin*, 88, 668–695. <https://doi.org/10.1037/0033-2909.88.3.668>.
- Bölte, S. (2005). *Reading mind in the eyes test: Erwachsenenversion*. Available online at www.as-tt.de/assets/applets/Augentest_Erwachsene.pdf
- Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: Gender effects in brain and behavior. *Neuroscience & Biobehavioral Reviews*, 46, 604–627. <https://doi.org/10.1016/j.neubiorev.2014.09.001>.
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale: Erlbaum.
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94, 100–131. <https://doi.org/10.1037/0033-2909.94.1.100>.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>.
- Graham, T., & Ickes, W. (1997). When women's intuition isn't greater than men's. In W. Ickes (Ed.), *Empathic accuracy* (pp. 117–143). New York: Guilford Press.
- Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, 4, 201–206. <https://doi.org/10.1037/1528-3542.4.2.201>.
- Hayes, A. F. (2018). The PROCESS macro for SPSS and SAS [Software]. PROCESS macro. <http://www.processmacro.org/download.html>.
- Hodges, S. D., Laurent, S. M., & Lewis, K. L. (2011). Specially motivated, feminine, or just female: Do women have an empathic accuracy advantage? In J. L. Smith, W. Ickes, J. A. Hall, & S. D. Hodges (Eds.), *Psychology of emotions, motivations and actions. Managing interpersonal sensitivity: Knowing when and when not to understand others* (pp. 59–74). New York: Nova Science Publishers.
- Hyde, J. S. (2013). Gender similarities and differences. *Annual Review of Psychology*, 65, 373–398. <https://doi.org/10.1146/annurev-psych-010213-115057>.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59, 730–742. <https://doi.org/10.1037/0022-3514.59.4.730>.
- Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, 7, 95–109. <https://doi.org/10.1111/j.1475-6811.2000.tb00006.x>.
- Kachel, S., Steffens, M. C., & Niedlich, C. (2016). Traditional masculinity and femininity: Validation of a new scale assessing gender roles. *Frontiers in Psychology*, 7, 1–19. <https://doi.org/10.3389/fpsyg.2016.00956>.
- Karni, R., Gabay, R., Ochion, Y., & Harari, Y. (1998). Is gender or gender-role orientation a better predictor of empathy in adolescence? *Sex Roles*, 39, 45–59. <https://doi.org/10.1023/A:1018825732154>.
- Kim, J., & Lee, S. J. (2010). Reliability and validity of the Korean version of the empathy quotient scale. *Psychiatry Investigation*, 7, 24–30. <https://doi.org/10.4306/pi.2010.7.1.24>.
- 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, 121–146.
- Klein, K. J., & Hodges, S. D. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin*, 27, 720–730. <https://doi.org/10.1177/0146167201276007>.
- Koch, K., Pauly, K., Kellermann, T., Seifert, N. Y., Reske, M., Backes, V., Stöcker, T., Shah, N. J., Amunts, K., Kircher, T., Schneider, F., & Habel, U. (2007). Gender differences in the cognitive control of

- emotion: An fMRI study. *Neuropsychologia*, 45, 2744–2754. <https://doi.org/10.1016/j.neuropsychologia.2007.04.012>.
- Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, 54, 2492–2502. <https://doi.org/10.1016/j.neuroimage.2010.10.014>.
- Lehrl, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B* (5. Auflage ed.). Balingen: Spitta Verlag.
- Lennon, R., Eisenberg, N., & Carroll, J. (1986). The relation between nonverbal indices of empathy and preschoolers' prosocial behavior. *Journal of Applied Developmental Psychology*, 7, 219–224. [https://doi.org/10.1016/0193-3973\(86\)90030-4](https://doi.org/10.1016/0193-3973(86)90030-4).
- Nanda, S. (2013). Are there gender differences in empathy. *Undergraduate Journal of Psychology at Berkeley*, 7, 22–42.
- Rand, D. G., Brescoll, V. L., Everett, J. A., Capraro, V., & Barcelo, H. (2016). Social heuristics and social roles: Intuition favors altruism for women but not for men. *Journal of Experimental Psychology: General*, 145, 389–396. <https://doi.org/10.1037/xge0000154>.
- Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the analysis of variance*. Cambridge: Cambridge University Press.
- Samson, A. C., & Huber, O. W. (2010). Short German versions of empathizing and systemizing self-assessment scales. *Swiss Journal of Psychology*, 69, 239–244. <https://doi.org/10.1024/1421-0185/a000028>.
- Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, 1156, 81–96. <https://doi.org/10.5167/uzh-25655>.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4–28. <https://doi.org/10.1006/jesp.1998.1373>.
- Steiger, J. H. (2004). Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods*, 9, 164–182. <https://doi.org/10.1037/1082-989X.9.2.164>.
- Thomas, G., & Maio, G. R. (2008). Man, I feel like a woman: When and how gender-role motivation helps mind-reading. *Journal of Personality and Social Psychology*, 95, 1165–1179. <https://doi.org/10.1037/a0013067>.
- Warrier, V., Grasby, K. L., Uzefovsky, F., Toro, R., Smith, P., Chakrabarti, B., Khadake, J., Mawbey-Adamson, E., Litterman, N., Hottenga, J.-J., Lubke, G., Boomsma, D. I., Martin, N. G., Hatemi, P. K., Medland, S. E., Hinds, D. A., Bourgeron, T., & Baron-Cohen, S. (2018). Genome-wide meta-analysis of cognitive empathy: Heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Molecular Psychiatry*, 23, 1402–1409. <https://doi.org/10.1038/mp.2017.122>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.