scientific reports

Check for updates

OPEN

Genetic variations as predictors of dispositional and dyadic empathy—a couple study

Maria Kaźmierczak[1], Magda Rybicka[2] & Paweł Syty[3,4]✉

Biological drivers of empathy have been explored in an interdisciplinary manner for decades. Research that merges the psychological and genetic perspectives of empathy has recently gained interest, and more complex designs and analyses are needed. Empathy is a multidimensional construct that might be regarded both dispositionally (as a personality trait) and contextually (experienced and/or expressed in a particular relationship/situation). This study analyzed genetic variations associated with genes encoding oxytocin, arginine vasopressin, and receptors that regulate their secretion as predictors of the empathic dimensions of emotional (empathic concern and personal distress) and cognitive (perspective taking) dyadic factors of partners in heterosexual intimate relationships. Machine learning methods to capture both linear and nonlinear relationships between SNPs, RS1 and RS2 repeat polymorphisms and dimensions of empathy in couples were employed. A total of 442 individuals (221 couples) participated in this study. Empathy was measured by the Polish version of the Interpersonal Reactivity Index and the Interpersonal Reactivity Index for Couples. The MassARRAY® 4 instrument, which combines mass spectrometry with endpoint PCR, was used for genotyping all 14 genetic variations. Microsatellite fragment analysis was performed by denaturing polyacrylamide gel electrophoresis. The results confirmed the significance of certain genetic alterations linked to oxytocin, vasopressin, serotonin and estrogen for dispositional and dyadic empathy (mainly rs1884051, rs6311, RS1, rs4686302, and rs1042778) in couples. The effects were stronger for the prediction of emotional and dyadic empathy than for perspective taking. Separate analyses for women and men indicated different predictive effects of genes for empathy (for example, effects of rs53576 were indicated only in women), which are also experienced and expressed in couples. Different dimensions of empathy should be included when the genetic predictors of empathy are examined.

Keywords Empathy, Genetic variability, Couples, Machine learning

Empathy has been linked to bonding, parental caregiving, relationship maintenance and intimacy[1–4]. It facilitates mutual support, cooperation, and prosocial behaviors, which have also been observed in nonhuman primates[5,6]. Thus, the mechanisms responsible for empathy and its outcomes have been sought. Apart from studies on sociocultural influences on the development and expression of empathy, its biological predictors, including neuroendocrine and genetic factors, have been well documented and interpreted as markers of evolutionary processes[7–10]. In mammals, empathy reflects approach tendencies to social cues, both positive and negative, in various relational contexts. Nevertheless, neurophysiological processes have been indicated as possible mediators of links between genes and empathic behaviors as well as traits, with different effects on emotional and cognitive empathy[7]. One of the factors behind empathic reactions might be the neuropeptide oxytocin, which is related to heightened attunement to social stimuli[11]. It has been suggested that oxytocin facilitates approach-related prosocial behaviors and decreases perceptions of social threats, and depending on the modulating role of the context, it might promote empathy[9,12,13].

The link between oxytocin and empathy has also been explored in genetic research. In this paper, we examined genetic predictors of dispositional and dyadic empathy in partners in heterosexual intimate relationships. The role of single nucleotide polymorphisms (SNPs) as predictors of emotional and cognitive

[1]Institute of Psychology, Faculty of Social Sciences, Division of Family Studies and Quality of Life, University of Gdańsk, Bażyńskiego 4, 80-309 Gdańsk, Poland. [2]Department of Photobiology and Molecular Diagnostics, Intercollegiate Faculty of Biotechnology, University of Gdansk and Medical University of Gdansk, Abrahama 58, 80-307 Gdansk, Poland. [3]Institute of Physics and Applied Computer Science, Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland. [4]BioTechMed Center, Gdańsk University of Technology, Narutowicza 11/12, 80-233 Gdańsk, Poland. ✉email: pawel.syty@pg.edu.pl

dispositional empathy in predicting individual differences in empathy has attracted some attention and has been debated, with oxytocin receptor gene (*OXTR*) polymorphisms being the most commonly studied[14]. However, to our knowledge, no studies have been conducted on the predictive effects of SNPs on empathy experienced toward a romantic partner. Researchers have suggested that empathic tendencies linked to oxytocin might be only enhanced in supportive, cooperative social contexts[11,15], and intimate, stable relationships might be one of them. It has recently been confirmed that the oxytocin receptor (OXTR) rs53576 polymorphism, which is linked to the empathic dispositions of partners, predicts increased attachment security and relationship satisfaction in intimate couples[16]. It has also been well documented that empathy is a predictor of high relationship quality (e.g.,[17]). Nevertheless, associations between partners' empathy and relationship quality are complex[1] and are not the focus of this paper. Instead, we uniquely explored whether the emotional and cognitive dimensions of empathy, as well as empathic tendencies oriented toward a romantic partner, are predicted by SNPs related mainly to oxytocin. We also innovatively tested the predictive effects of arginine vasopressin (AVP), which has been linked to social functioning and affiliative tendencies[14], and genetic variation within genes encoding receptors that regulate oxytocin and vasopressin secretion (*ESR1* and *HTRA2*).

Dimensions of empathy

Empathy is a multidimensional concept that might be viewed as a disposition rather than a personality trait. In accordance with this approach, we followed the definition by Davis[18–20], who differentiated emotional and cognitive dimensions of trait empathy, all of which were measured with the Interpersonal Reactivity Index (IRI). In detail, we explored the genetic predictors of three dispositional empathy dimensions: (1) empathic concern (EC; sometimes called empathy or sympathy[21,22])—other-oriented emotional empathy related to feelings of warmth, compassion, care and support of others in need; (2) personal distress (PD) —self-oriented emotional empathy linked to emotional contagion, distress or anxiety while facing others' plight, which ultimately leads to a focus on one's own well-being; and 3) perspective taking (PT) —other-oriented cognitive empathy, attempts to imagine and take others' point of view in social situations[23].

Additionally, the dyadic approach has been introduced to the measurement of self-reported empathy, which is subjectively experienced in the context of a romantic relationship and shown to a partner. Thus, dispositional (experienced independently of specific relationships) and dyadic empathy are not the same constructs. Personal involvement in a couple's dynamics might make it difficult to experience or express empathy toward a romantic partner depending on the quality of this relationship. Thus, self-reported dyadic empathy might reflect the strength and intimacy of connections between partners more than general tendencies to empathize with others[24,25]. This type of empathy translates into behaviors that promote higher relationship quality, such as displaying care and affection or adopting a partner's point of view[24,26]. It has been repeatedly confirmed that dyadic empathy predicts or reflects highly satisfactory relationships[24–26].

It should be noted that dyadic empathy has been shown to be independent of relationship duration[24,25]. However, partners' empathy plays a crucial role in providing emotional support within couples, particularly during significant role transitions. Research indicates that higher levels of dyadic empathy predict better relationship adjustment for both women and men during the transition to parenthood, by reducing stress and fostering intimate communication and care[26]. Partners' perceived empathy has also been linked to better adjustment to parenthood[2]. In this context, empathy acts as a protective factor against the negative effects of relationship strain and serves as a foundation for maintaining emotional intimacy[27]. As couples face various life-role challenges, supportive behaviors may increasingly depend on each partner's empathic tendencies, to a greater extent than in early stages of dating[28]. From a lifespan perspective, studies largely show that empathy remains stable or undergoes only small changes—whether slight increases or declines—across adulthood, with mixed results concerning emotional and cognitive empathy dimensions[29]. These variations are often attributed to cohort differences[30]. Still, more empathic individuals rate their interpersonal interactions more positively throughout adulthood, suggesting that empathy contributes to sustaining healthy relationships well into later life[30,31]. Therefore, considering the relational context in genetic studies on empathy seems valuable.

In this paper, we aimed to expand the knowledge not only about genetic predictors of dyadic empathy but also about congruence in declared partner-oriented empathic reactions. Schneiderman et al.[32] previously added a genetic perspective to the mutual empathy of partners by finding that certain SNPs related to the oxytocin receptor gene (*OXTR*) predicted levels of reciprocity and congruence in partners' behaviors during couple communication, which are important aspects of dyadic empathy[25]. Such findings fall within the broader area of research on the benevolent effects of similarity or congruence in perceptions of empathy between romantic partners for their relationships[1,33].

Genetic predictors of empathy

Twin studies have indicated that empathy is partially heritable[34]. Greater effects have been obtained for emotional empathy linked to temperamental emotional reactivity or regulation than for cognitive empathy, which is more strongly associated with environmental influences throughout individual development[7,35,36]. Hypotheses have been formulated about the heritability of so-called "affective drive," which is reflected in emotional empathy, and about the crucial role of learning processes in taking the perspective of others in social situations[36]. Nevertheless, across various studies and methodologies, there have been inconsistencies regarding gender differences in genetic and environmental influences on empathy, with research results pointing to genetic roots of either female[10] or male[37] empathy; no gender differences in the heritability of empathy have been shown[38].

Since oxytocin increases prosocial motivation and salience of social signals, SNPs related to oxytocin receptor genes have been the most frequently analyzed in studies on predictors of empathy[7]. Meta-analyses of *OXTR* rs53576, which is commonly studied in this context, have indicated that GG homozygotes often have greater empathic tendencies than do individuals with AA genotypes[39], especially with respect to emotional empathy[40].

However, in their meta-analysis, Chander et al.[41] reported that research on a general empathy construct and SNPs indicated a lack of genetic effects. Studies on general sociality seem to confirm this conclusion. In their meta-analysis, Bakermans–Kranenburg and Van Ijzendoorn[42] concluded that *OXTR* rs53576 is not linked to any social behaviors or personality traits such as empathy. Nevertheless, Li et al.[43] also performed a meta-analysis and found that *OXTR* rs53576 GG homozygotes displayed greater general sociality (emotional support seeking, empathy, or extraversion) but not in the context of close relationships (attachment, dealing with relational conflicts). *OXTR* rs53576 might be more strongly associated with individual differences in behaviors and attitudes in sexual behaviors but not with dispositional empathy[44]. For example, the inclusion of emotional and cognitive empathy better reflected the diverse predictive effects of rs53576, especially in European cohorts. Rs53576 GG homozygotes scored higher in EC[41], and this effect was sometimes confirmed only in women (not in men)[45]. Lower emotional or other-oriented empathy has been confirmed in adults with *OXTR* rs53576 A alleles (AA/AG)[14,46].

Furthermore, it has been recommended that multiple SNPs effects within candidate genes should be tested simultaneously for effects on different aspects of sociality, including empathy[44]. As the *OXTR* gene is a determinant of oxytocin sensitivity, some SNPs have been analyzed in relation to not only empathy but also functioning in social relations, sometimes in couples. Israel et al.[47] reported that rs1042778 is significantly linked to altruism and autism and that TT homozygotes were also less prosocial than individuals with the G allele. Regarding empathy dimensions, Chist et al.[48] reported that female GG carriers had greater PT scores than TT carriers, but no effects of this SNP on empathy were reported[49]. Rs4686302 has also been linked to empathy. Research has indicated that men with the CT and TT rs4686302 genotypes score lower than women do in terms of emotional empathy[49]. Individuals with C alleles scored lower than others (TT homozygotes or CT genotypes, depending on the study) in PT[49,50]. Additionally, a recent study on children with ADHD indicated that the rs4686302 C allele corresponded to greater facial emotional recognition[51], which is linked to empathy and emotional intelligence[23]. Thus, the results on the associations between rs4686302 and empathy dimensions are inconclusive.

Many other SNPs were found to be inconsistently associated with empathy and related social variables[40]. For example, a frequently tested SNP, rs2254298, was variously linked to either emotional or cognitive empathy as measured by the IRI, which was moderated or not moderated by gender[48,49], and the results differed in different clinical samples, such as in patients with schizophrenia[52]. Rs237887 was linked to prosocial behaviors[47] and, inconsistently, to cognitive or emotional empathy[48–50]. In terms of AVP, polymorphisms within *AVPR1a* were linked to cognitive empathy[44], with a long RS3 promoter repeat length predicting lower scores[14]. Zhang et al.[53] concluded that short RS3 and RS1 repeats are typically linked with less sociality. Furthermore, both oxytocin and vasopressin were correlated with the estrogen receptor alpha (*ERα*) mRNA, suggesting that not only more SNPs within a gene of interest but also genes related to the analyzed pathway should be examined simultaneously[54]. In addition, changes in mitochondrial function have recently been linked to psychiatric disorders, including those related to social functioning and aggressive behavior. In addition, a serine protease in the mitochondria, HTRA2, has been implicated in promoting cell death, which could indirectly affect social behavior. Although there is no specific research on its direct involvement in empathy, studies investigating a possible link between HTRA2 and empathy are worthwhile[55,56].

The findings described above led to the conclusion that empathy dimensions and the context of relationships should be included in studies on the predictive effects of the above-described SNPs. For example, the question might be posed whether partners with the rs53576 GG genotype, especially women, experience and express more empathy in their relationship and whether they are more similar in dyadic empathy. In studies on couples, the existing results indicated that rs1042778 (the TT allele as a risk factor) exerted effects on interactions between partners depending on[57] or irrespective of[32] their gender. Schneiderman et al.[32], in a sample of 120 couples, showed that the rs1042778 T allele was linked to lower communicative empathy and lower verbal and nonverbal signs of affection, responsiveness and harmony in dyads. According to the analyses of 79 couples by Mattson et al.[57], men's rs1042778 T allele was linked to less support provided and perceived in couples, which reflected the dyadic context of this effect. Another SNP examined in the same study (rs4686302), the CT genotype, was associated with less support in intimate relationships than CC homozygotes (effects for women). Further exploration of multiple SNPs' effects on empathic dimensions, and in couples, would add new insights into this complex but socially important area of individual differences.

The present study

The aim of this paper was to explore the effects of multiple SNPs within genes linked to empathy and social functioning. We went beyond *OXTR* and *AVPR1a* polymorphisms and examined genetic variation within genes encoding receptors that regulate oxytocin and vasopressin secretion (*ESR1*, *HTRA2*). Taking into consideration previous research, we analyzed selected genetic predictors of empathy within *OXTR* (rs53576, rs1042778, rs13316193, rs2228485, rs2254298, rs2268494, rs4686302, rs237887, rs7632287), two in *HTRA2* (rs6311, rs6314), one in *ESR1* (rs1884051), and two *AVPRA1* microsatellites (RS1 and RS3). Based on earlier research, we expected the lowest scores in other-oriented empathy in participants with rs53576 A alleles (H1), with stronger effects for women (H1a), and in participants with T rs1042778 alleles (H2). We further asked whether rs4686302 heterozygotes would be less empathic and whether other SNPs predict levels of dispositional empathy in our sample. Additionally, as empathy is linked to gender stereotypes and femininity[58], we explored the above effects for women and men separately.

We added the dyadic perspective to the link between genetic factors and empathy and included intimate partners' assessments of their empathy experienced and expressed in their relationships in a relatively large sample of couples. To explore the dyadic data more extensively, we analyzed whether congruence in partner-

Measures	1	2	3	Men		Women		$t(440)$	Cohen's d
				M	*SD*	*M*	*SD*		
Empathic concern	–			36.68	6.18	40.85	5.72	− 7.363**	− 0.70
Personal distress	0.35**	–		19.40	4.78	24.16	5.36	− 9.853**	− 0.94
Perspective taking	0.37**	0.04	–	33.34	5.02	34.42	4.46	− 2.385*	− 0.23
Dyadic empathy	0.29**	0.04	0.43**	40.29	4.96	40.93	4.61	− 1.421	− 0.14

Table 1. Intercorrelations, means, standard deviations of study variables and differences in means between women and men. *$p < 0.05$, **$p < 0.001$.

Genetic variation	Allele order
rs53576	[G, GA, A]
rs2254298	[G, GA, A]
rs237887	[A, AG, G]
rs4686302	[C, CT, T]
rs13316193	[T, TC, C]
rs1042778	[G, GT, T]
rs2268494	[T, TA, A]
rs7632287	[G, GA, A]
rs2228485	[A, AG, G]
rs1884051	[A, AG, G]
rs6311	[C, TC, T]
rs6314	[G, AG, A]
RS1	[L, M, S]
RS3	[L, M, S]

Table 2. Encoding scheme for alleles of analyzed genetic variations (IVs). For numerical purposes, all IVs were encoded as [0, 1, 2] in order.

oriented empathy between women and men, which reflects mutual empathy in intimate relationships, was predicted by SNPs.

Following recent trends in studies on individual differences and relationship quality[59], we employed machine learning methods to capture both linear and nonlinear relationships between SNPs and dimensions of empathy in couples. Machine learning allows the shift of deep domain knowledge toward constructing and training machine models[60]. Such models allow us to look for relationships and interactions between variables that might go unnoticed during typical statistical analysis.

Linear mixed effects (LMEs) and random forest (RF) modeling

Data preprocessing

The data have been preprocessed in a standard way[61]; e.g., missing (or incorrect) values have been imputed (or replaced) by the most frequent value in each genetic variation. The total percentage of erroneous and missing values was less than 0.3% of the full data, so it had no practical impact on the results.

Intercorrelations, means, standard deviations of study variables and differences in means between women and men are collected in Table 1. Analysis of the table indicates the need to model, in addition to the group as a whole, subgroups composed of individuals of different genders. The statistical power of the sample, assuming a significance level of $\alpha = 0.05$, sample size of $n = 442$, and standardized effect size of $f = 0.23$ (calculated using data from[62]), is > 90%.

To create numerical models for studying the significance of each individual genetic variation (independent variable, referred to as IV), their alleles were encoded to [0, 1, 2] using the order [*major homozygote*, *heterozygote*, *minor homozygote*][63] (see Table 2). For example, alleles [G, GA, A] in rs53576 were encoded as [0, 1, 2]. For vasopressin (RS1, RS3), a fixed encoding order [L, M, S] of [0, 1, 2] was used instead, where L is the long, M is the medium and S is the short repeat length. We stress that encoding has been introduced as a requirement of the numerical libraries used in calculations, but these IVs were still treated as categorical data.

Overview of the machine learning modeling for explanatory analysis

To determine the significance of each IV in predicting empathy (including testing H1, H1a, and H2), two modeling methods were employed in the current work: linear mixed effects (LMEs) and random forests (RFs). Several models based on these methods were constructed to predict the values of the dependent variables (EC—Empathic concern, PD—Personal distress, PT—Perspective taking, DYADIC—Dyadic empathy and dyadic

empathy congruence—DYADIC CONGRUENCE) as a function of the IVs. These models were trained and then subjected to explainability to determine the significance of each SNP.

The LME method, in general, allows for the modeling of structured linear relationships between variables[64]. In that approach, data points consist of inputs of varying types, which may be categorized into groups, and the output of a real value. LME models are hierarchical ones: they share statistical strength across groups, which, in turn, allows for improving inferences about any single data point[65]. Each model is composed of fixed effects (held constant across the population of data points) and random effects (varied across the population of data points). This method (as with the other linear methods) allows for easy determination of the contribution of particular variables to the model performance (by analyzing the regression coefficient), together with the statistical significance (p values). However, although simple and effective, it does not automatically detect interactions between variables—they need to be included explicitly in the model.

The RF method, in contrast, allows for modeling complex, strong nonlinearities between variables[66] while being resistant to overfitting. Unfortunately, this comes at the cost of a loss of model explainability. However, several techniques can be used for explanatory analysis of trained RF models[67]. Moreover, RFs are also suitable for studying interactions between predictors[68].

Details of the LME models

In the current analysis, fixed effects were chosen as the individual genetic variations (14 IVs). Random effects were introduced to the model in the sense of random intercept terms only. In the case of the whole sample, random effects were connected with the grouping of the population by (A) couple ID and (B) person ID. This allows for checking for couple effects in the gene analysis. For the separate analyses of women and men, random effects were introduced by grouping the population by person ID only. For each measure, three models were studied:

Model 1: baseline model with no predictors (1A—grouping by couple ID, 1B—grouping by person ID),
Model 2: all IVs as predictors (incl. testing of H1 and H2; 2A—grouping by couple ID, 2B—grouping by person ID),
Model 3: all IVs as predictors, including the interaction between rs53576 and gender (H1a; 3A—grouping by couple ID, 3B—grouping by person ID; for the whole group only, since including an interaction with gender is inappropriate when the analyzed population contains only women or men).

Additionally, for the DYADIC CONGRUENCE measure, Model 3 of type A (grouping by couple) was omitted because DYADIC CONGRUENCE is calculated as the difference between the DYADIC values within a couple, so grouping of the population by couple is inappropriate.

In order to test whether expecting a child affects the performance of the above models, a random effect was introduced for each model for both groups (individuals/couples expecting, and not expecting a child).

The importance of each individual genetic variation (IV) was determined by analyzing the calculated regression coefficient, together with the corresponding p value. To facilitate the comparison of coefficients and the assessment of their significance, standardized coefficients were also calculated. They were obtained by standardizing the data prior to using LME analysis. To compensate for multiple tests, adjusted p values were calculated using the Holm–Šidák method[69]. To estimate the size of the effects, Cohen's f^2 was calculated for variables with p values < 0.05.

Details of the RF models

The optimal set of hyperparameters of RF models has been found to maintain the trade-off between performance and the required computing time. The number of independent trees (number of estimators) was restricted to 1500 without restricting their depths. Each model for each dependent variable was trained; then, along with the study aim (including testing H1, H1a, and H2), the significance of the variables was then calculated for each model using explanatory methods[67]. The out-of-bag estimate[70] has been used, allowing for efficient overfitting prevention.

Results

LME

To train the models, which were subsequently used to calculate the significance of each SNP, the Python package *statsmodels* was used (https://www.statsmodels.org). The performances of the models are summarized in Table 3, which shows that grouping by couple ID (Models 1A, 2A, 3A) did not result in a sufficiently high intraclass correlation (ICC < 0.05). Therefore, for further analysis, models introducing grouping by person ID were selected. Model 3B (including the interaction between rs53576 and gender) was the best model when the full sample was analyzed, but for females and males, Model 2B showed the best fit to the data.

For each model, the inclusion of grouping by expecting (or not) a child was found to have no effect on performance (the coefficient for the random effect was zero or negligibly small in each case).

A summary of the results for Model 2B (including all SNPs, without an interaction term, grouping by person ID) is shown in Table 4 (for the whole group) and Table 5 (separately for females and males). The list of the most important IVs was selected as those with p value (adjusted for multiple testing) < 0.05 and at least a small effect size ($f^2 > 0.02$). For variables with adjusted p value ≥ 0.05, IVs with unadjusted p value < 0.05 and at least a small effect size were also selected. The full results for Models 1B, 2B and 3B are shown in Supplementary Tables S1–S5.

Thus, H1 and H1a were partially confirmed, as no strong effects on dispositional empathy occurred for the whole sample or for women. For women, rs53576 predicted dyadic empathy levels and dyadic congruence scores (the more A alleles there were, the lower the score and the greater the incongruence in empathy between women

Variable	Whole sample		Women		Men	
	− 2LL	RMSE	− 2LL	RMSE	− 2LL	RMSE
EC (model 1A)	2881.10	6.30	–	–	–	–
EC (model 1B)*	2881.10	2.99	1396.94	2.58	1431.45	2.79
EC (model 2A)	2866.96	6.20	–	–	–	–
EC (model 2B)*	2866.96	2.80	1377.60	2.47	1419.46	2.72
EC (model 3A)	2813.02	5.43	–	–	–	–
EC (model 3B)*	2814.15	2.64	–	–	–	–
PD (model 1A)	2777.01	5.60	–	–	–	–
PD (model 1B)*	2777.01	2.66	1368.12	2.42	1317.94	2.19
PD (model 2A)	2750.51	5.43	–	–	–	–
PD (model 2B)*	2750.51	2.46	1340.83	2.27	1308.55	2.11
PD (model 3A)	2668.44	4.95	–	–	–	–
PD (model 3B)*	2668.44	2.24	–	–	–	–
PT (model 1A)	2634.60	4.77	–	–	–	–
PT (model 1B)*	2634.60	2.16	1287.00	2.01	1338.88	2.26
PT (model 2A)	2625.04	4.71	–	–	–	–
PT (model 2B)*	2625.04	2.13	1271.08	1.94	1327.52	2.21
PT (model 3A)	2617.91	4.62	–	–	–	–
PT (model 3B)*	2617.94	2.12	–	–	–	–
Dyadic (model 1A)	2638.80	4.76	–	–	–	–
Dyadic (model 1B)*	2638.81	2.17	1301.95	2.08	1333.69	2.24
Dyadic (model 2A)	2613.51	4.65	–	–	–	–
Dyadic (model 2B)*	2613.51	2.11	1275.66	1.96	1316.63	2.15
Dyadic (model 3A)	2607.43	4.58	–	–	–	–
Dyadic (model 3B)*	2607.45	2.20	–	–	–	–
D. C. (model 1A)	–	–	–	–	–	–
D. C. (model 1B)*	2938.11	3.04	1469.05	3.19	1469.05	3.19
D. C. (model 2A)	–	–	–	–	–	–
D. C. (model 2B)*	2926.51	3.00	1453.09	2.93	1459.78	2.98
D. C. (model 3A)	–	–	–	–	–	–
D. C. (model 3B)*	2924.96	2.99	–	–	–	–

Table 3. Performance of the LME models: − 2 times the log of the likelihood (− 2LL) and root-mean-square error (RMSE). *D.C.* dyadic congruence. *Indicates ICC > 0.05.

Variable	IVs	Coefficient	Coefficient (standardized)	p value	p value (adjusted)	f^2
EC	–	–	–	–	–	–
PD	RS1	− 1.193	− 0.213	0.001	0.005	0.131
	rs4686302	1.494	0.267	0.015	0.059	0.021
	rs6311	0.892	0.159	0.018	0.087	0.020
PT	–	–	–	–	–	–
DYADIC	rs1042778	− 0.969	− 0.202	0.007	0.035	0.024
	rs4686302	− 1.372	− 0.286	0.009	0.044	0.022
	rs237887	− 0.938	− 0.196	0.035	0.163	0.020
Dyadic congruence	rs2254298*	− 2.230	− 0.332	0.019	0.091	0.013

Table 4. List of most important IVs (with $p < 0.05$) as result of LME modeling for the whole group (Model 2B). For EC and PT no genetic variation with statistical significance is found. Coefficient is the slope of the regression straight line, higher absolute value implies higher importance. Sign of the coefficient is the trend, according to the scheme in Table 2. Standardized coefficients, p values adjusted for multiple testing (using the Holm–Šidák method), and Cohen's f^2 effect size are also presented. *Excluded from further analysis, due to highly imbalanced data and too small effect size ($f^2 < 0.02$).

Variable	IV	Coefficient	Coefficient (standardized)	p value	p value (adjusted)	f^2
Women						
EC	rs1884051	− 1.837	− 0.321	0.001	0.005	0.178
PD	rs6311	2.037	0.380	0.000	0.000	0.073
	RS1	− 1.086	− 0.203	0.024	0.114	0.023
PT	–	–	–	–	–	–
Dyadic	rs13316193	− 1.699	− 0.368	0.005	0.025	0.035
	rs53576	− 1.277	− 0.277	0.026	0.123	0.022
Dyadic congruence	rs53576	− 1.669	− 0.248	0.049	0.182	0.020
Men						
EC	rs4686302	− 2.526	− 0.408	0.010	0.049	0.031
PD	–	–	–	–	–	–
PT	–	–	–	–	–	–
Dyadic	rs4686302	− 1.867	− 0.377	0.016	0.062	0.026
	rs1042778	− 1.092	− 0.220	0.038	0.176	0.021
Dyadic congruence	–	–	–	–	–	–

Table 5. List of most important IVs (with $p < 0.05$) as result of LME modeling separately for women and men (model 2B). For PT (both genders) as well as for PD and Dyadic congruence for men no IV with statistical significance is found. Coefficient is the slope of the regression straight line, higher absolute value implies higher importance. Sign of the coefficient is the trend, according to the scheme from Table 2. Standardized coefficients, p values adjusted for multiple testing (using the Holm–Šidák method), and Cohen's f^2 effect size are also presented.

	Whole group		Women		Men	
Variable	RMSE	R^2	RMSE	R^2	RMSE	R^2
EC	2.55	0.84	2.24	0.85	2.51	0.83
PD	2.33	0.83	2.06	0.85	1.99	0.83
PT	2.03	0.82	1.83	0.83	2.07	0.83
Dyadic	1.97	0.83	1.87	0.84	2.03	0.83
Dyadic congruence	2.90	0.81	2.81	0.82	2.81	0.83

Table 6. Root-mean-square error (RMSE) and R^2 for the trained RF models.

and men). The rs1042778 T allele seemed to be a risk factor for dyadic empathy, particularly for men, which partially confirmed H2, as no effects on dispositional empathy were shown. Several other predictive effects for emotional and dyadic empathy also occurred.

Random forests

The basic metrics obtained during the training of the RF models are listed in Table 6. The results for all of the models are similar and are at an acceptable level.

A summary of the importance of each IV calculated from the nonlinear RF models, is presented in Table 7. For the calculation, the *dalex* Python package was used[67], which implements the idea described in[71]. Basically, how much a model's performance changes is measured if the effect of a selected predictor is removed (the so-called drop-out loss). Since there is no natural cutoff that can be used to discriminate between important and nonimportant variables when using variable importance measures[72], we created a criterion based on the drop-out loss value. The most important IVs were chosen as those for which the drop-out loss was above the threshold chosen as the mean + 0.2*range of the drop-out losses for all IVs. This choice of cutoff allows for marking 2–4 variables as significant.

RF was also used to demonstrate how the expected value of model prediction behaves as a function of a particular IV. This was achieved with the *ceteris paribus* analysis— "other things held constant" or "all else unchanged"[67]. The results for the 3 most important IVs for all analyzed measures are presented in Figs. 1, 2 and 3 (for the whole group, females and males, respectively). The results for all IVs are included in Supplementary Figs. S4–S8.

Next, the RF model was used to examine the interaction effects between the rs53576 SNP and the participant' gender (indicated by the additional independent variable SEX). This particular SNP was selected because it is expected to interact with the participants' gender in predicting dimensions of empathy[45]. This has been done by creating and training additional RF models, with rs53576 and SEX variables treated as (i) independent and (ii) interacting variables. The increase in the R^2 measure in these two models is due to the inclusion of interactions.

Variable	Genetic variations		
	Whole group	Women	Men
EC	rs1884051 RS1 rs6311 rs1042778	rs1884051 rs1042778 rs53576	rs1884051 rs4686302 rs6311 RS1
PD	RS1 rs6311	rs6311 RS1	RS1 rs6311 rs1042778
PT	rs6311 rs1884051 rs237887	RS3 rs1884051 rs53576 RS1	RS1 rs237887 rs1884051
DYADIC	rs6311 rs1884051 RS3	rs6311 rs13316193 rs53576	RS1 rs1884051 rs1042778 RS3
Dyadic congruence	rs1884051 rs6311 RS1 rs53576	rs6311 rs7632287 RS1 rs53576	rs1884051 rs4686302

Table 7. List of most important IVs as result of RF nonlinear modeling.

The highest R^2 increase (and thus stronger interactions) was observed for the DYADIC variable (130%) and EC (13%). Much lower increases (and thus weak interactions) were found for PD and PT (1% and 3%, respectively). Moreover, in Fig. 4, interaction plots for these variables are presented. These findings clearly confirm the above result.

Once again, H1 and H1a were partially confirmed. RF indicated that women with the AA rs53576 genotype scored lower in EC and dyadic empathy as well as in PT. Regarding H2, rs1042778 TT seemed to be linked to lower dyadic empathy, as expected (similarly to heterozygotes) but also to PD (as opposed to individuals with the AG and GG genotypes), but only in men. Other nonlinear effects occurred, especially for the *AVPR1a*, *ESR1* and *HTRA2* polymorphisms.

Discussion

Our findings confirmed the significance of certain genetic variations linked to oxytocin, vasopressin, serotonin and estrogen for dispositional and dyadic empathy in couples. Thus, the obtained results went beyond the frequently analyzed effects of *OXTR* polymorphisms, especially *OXTR* rs53576, on empathic tendencies. The results expanded upon earlier findings on genetic predictors of dimensions of empathy by adding its dyadic context, which was also predicted by several SNPs. It should be emphasized that while women scored higher in each dimension of dispositional empathy, there were no gender differences in dyadic empathy, which might reflect the quality of the relationship more than gender roles. The use of multiple statistical analyses, together with machine learning, allowed for a more accurate description of linear and nonlinear relationships between analyzed variables. There were many effects of the analyzed genetic variations on emotional and dyadic empathy shown in the regression predictive models (LME and RF), so conclusions on the linearity of the relationships between variables could be drawn. Overall, the weakest effects of genetic factors were obtained for perspective taking, which is in line with research emphasizing social influences on the development of cognitive empathy[73].

The obtained results shed new light on the effects of *OXTR* rs53576 on dimensions of empathy. When gender was added to the analysis, RF showed that women with AA genotypes reported lower EC and dyadic empathy (as confirmed with LME). The results confirmed and expanded the conclusions of Uzefovsky et al.[14] by pointing to EC, which is other-oriented as the dimension most strongly and nonlinearly associated with rs53576, similar to Huetter et al.[45], and the effects were much weaker for the PT dimension. We showed that not only empathizing emotionally but also caring for others, including intimate partners, by focusing on their situation might be more challenging for women with AA genotypes than for those with other genotypes. In this respect, our results indicate that the rs53576 polymorphism might be predictive of empathy expressed and experienced in a romantic relationship—empathic concern toward a partner and taking on their perspective—and that both aspects of empathy are closely linked to support, sensitive responsiveness and good communication in relationships[20,74–76]. However, we did not observe a significant increase in empathy scores in GG homozygotes compared with those in GA heterozygotes, which might serve as a limitation of the obtained results. Hence, these results go beyond a particular social domain and encompass both a more general social disposition and intimate relationships[44]. In this respect, H1 and H1a were partially confirmed. However, the data for rs53576 were unbalanced (only n = 18 women with AA homozygotes), so the conclusions should be drawn with caution. For the same reason, the effects of rs2254298 were not further discussed.

Two other frequently examined SNPs, rs1042778 and rs4686302, were associated with dyadic and emotional dispositional empathy. In accordance with earlier studies on the T allele as a risk factor for low empathy and disturbances in romantic relationships, GG homozygotes in rs1042778 reported more empathy expressed toward their romantic partners than did participants with at least one T allele. The more T alleles there are, the less that women and especially men view themselves as empathic in their intimate relationships. The stronger effects for

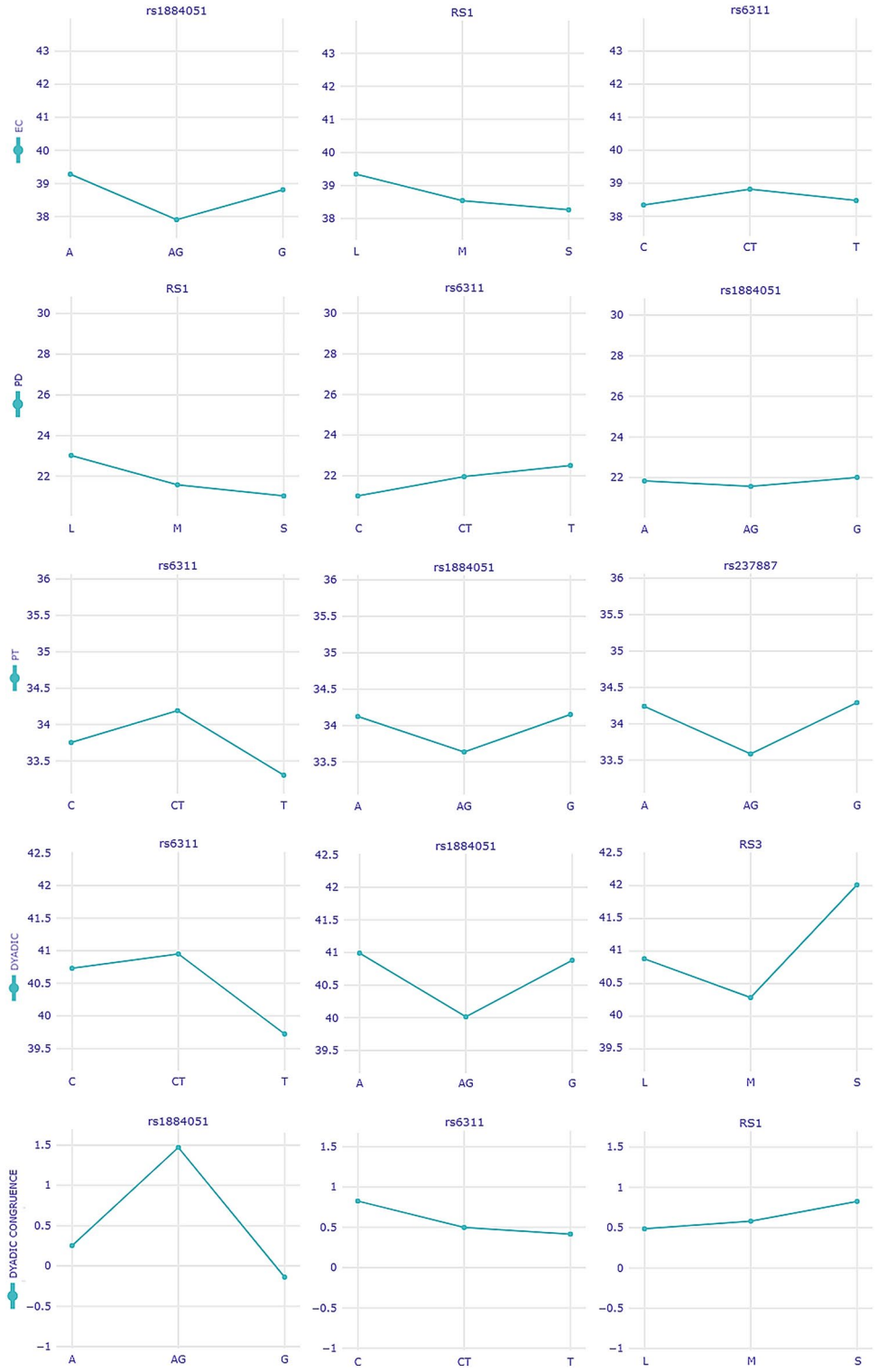

Fig. 1. Expected values of the RFs models predictions for all dependent variables (in rows), as function of three most important genetic variations (IVs). Results for the whole group.

men confirmed the earlier findings of Mattson et al.[57], but in the context of empathy and in a larger sample of couples. Other effects of this SNP seemed to be weaker. Therefore, H2 was partially confirmed.

Furthermore, the more C alleles there were in *OXTR* rs4686302, the greater the level of dyadic empathy and the lower the level of PD, as shown by linear analysis (LME); therefore, such participants displayed tendencies toward other-oriented empathy in intimate relationships and were more able to regulate their negative emotions when faced with others' distress. It seems that individuals with these alleles might function better in close relationships

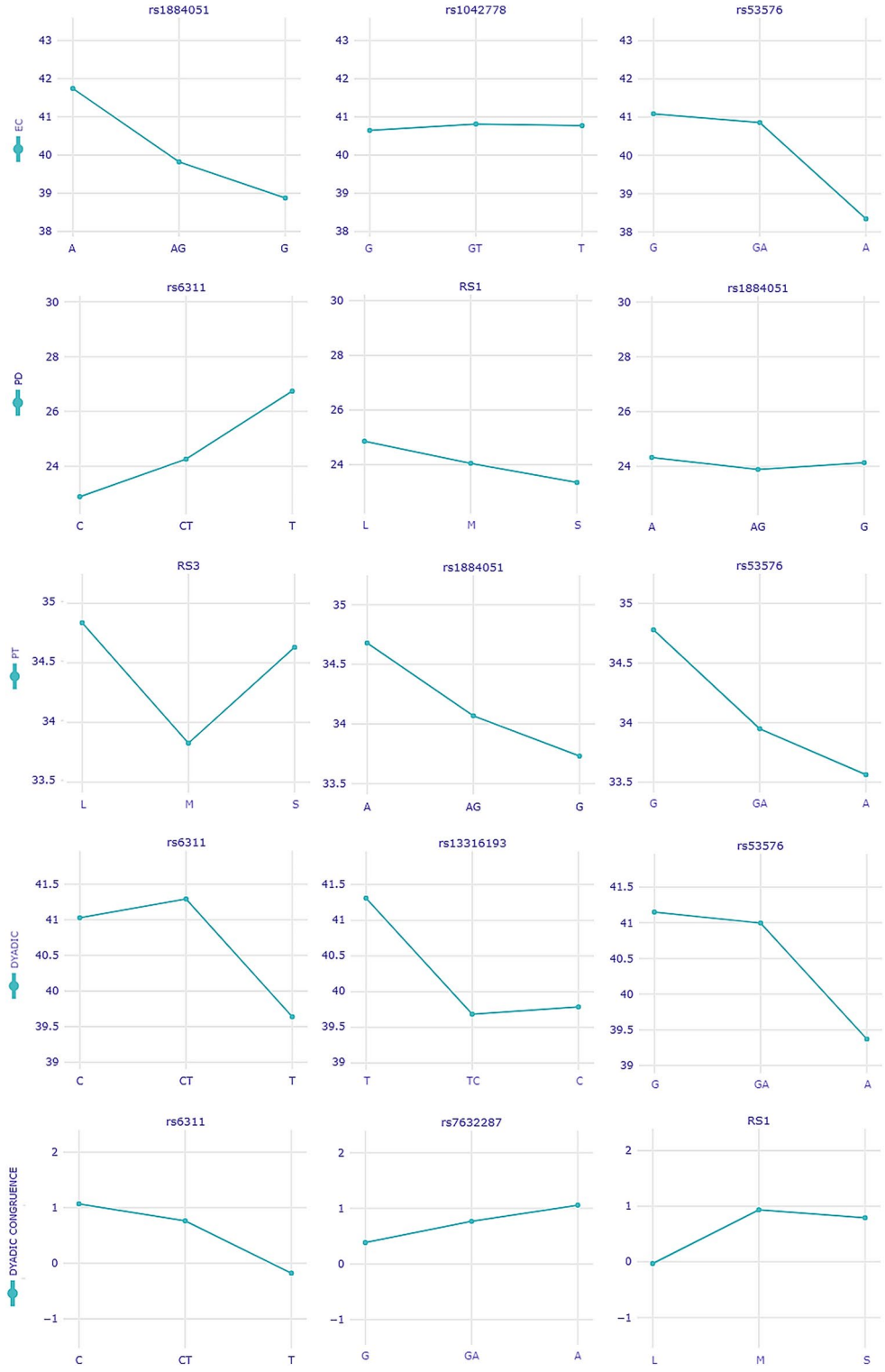

Fig. 2. Expected values of the RFs models predictions for all dependent variables (in rows), as function of three most important IVs. Results for women only.

than individuals with other alleles. As in Mattson et al.[57], we discovered that gender should be considered in the analysis of rs4686302 effects, but in our study, the effects were stronger for men, who additionally obtained lower scores in EC when T alleles were present (confirmation with LME and RF). However, in their study, there were no TT homozygotes, and in our study, there were only 13 participants with TT genotypes. Nevertheless, in our analyses, we detected the effects of rs4686302 on both LME and RF. Thus, these results provide new insights into the earlier, often contradictory, findings on the effects of rs4686302 on socioemotional functioning.

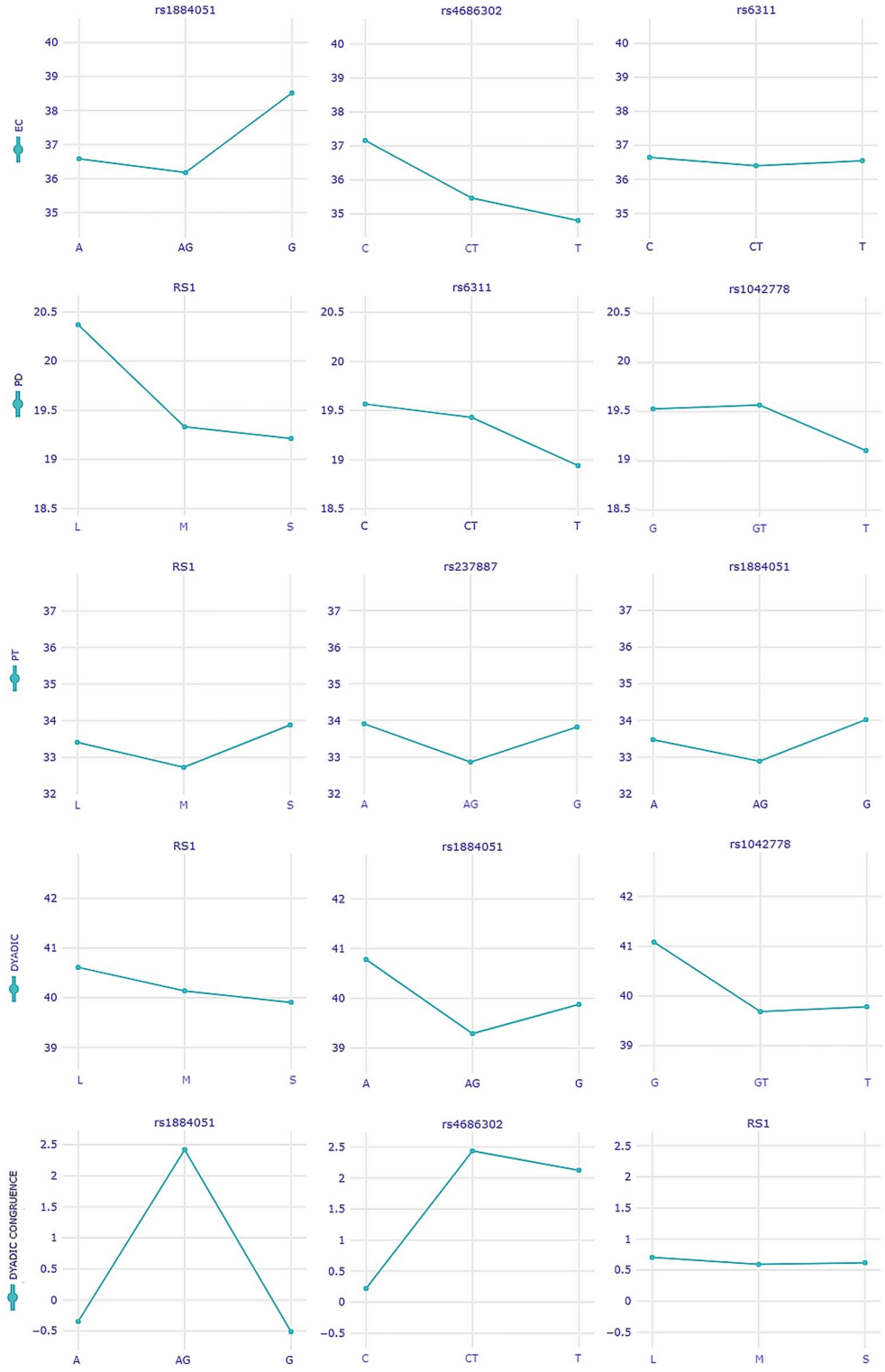

Fig. 3. Expected values of the RFs models predictions for all dependent variables (in rows), as function of three most important IVs. Results for men only.

Other effects of *OXTR* rs13316193 and rs237887 emerged. Earlier research regarding the effects of these two SNPs, as well as rs7632287, was inconclusive[40]. LME and RF indicated that women with the rs13316193 TT genotype obtained higher scores for dyadic empathy. LME showed that A alleles in rs237887 were linked to greater dyadic empathy. The effects of rs237887 on PT levels were relatively weak (the drop-out loss values were smaller than those for the other measures), again showing that this dimension is not predicted by SNPs to such an extent as emotional and dyadic empathy.

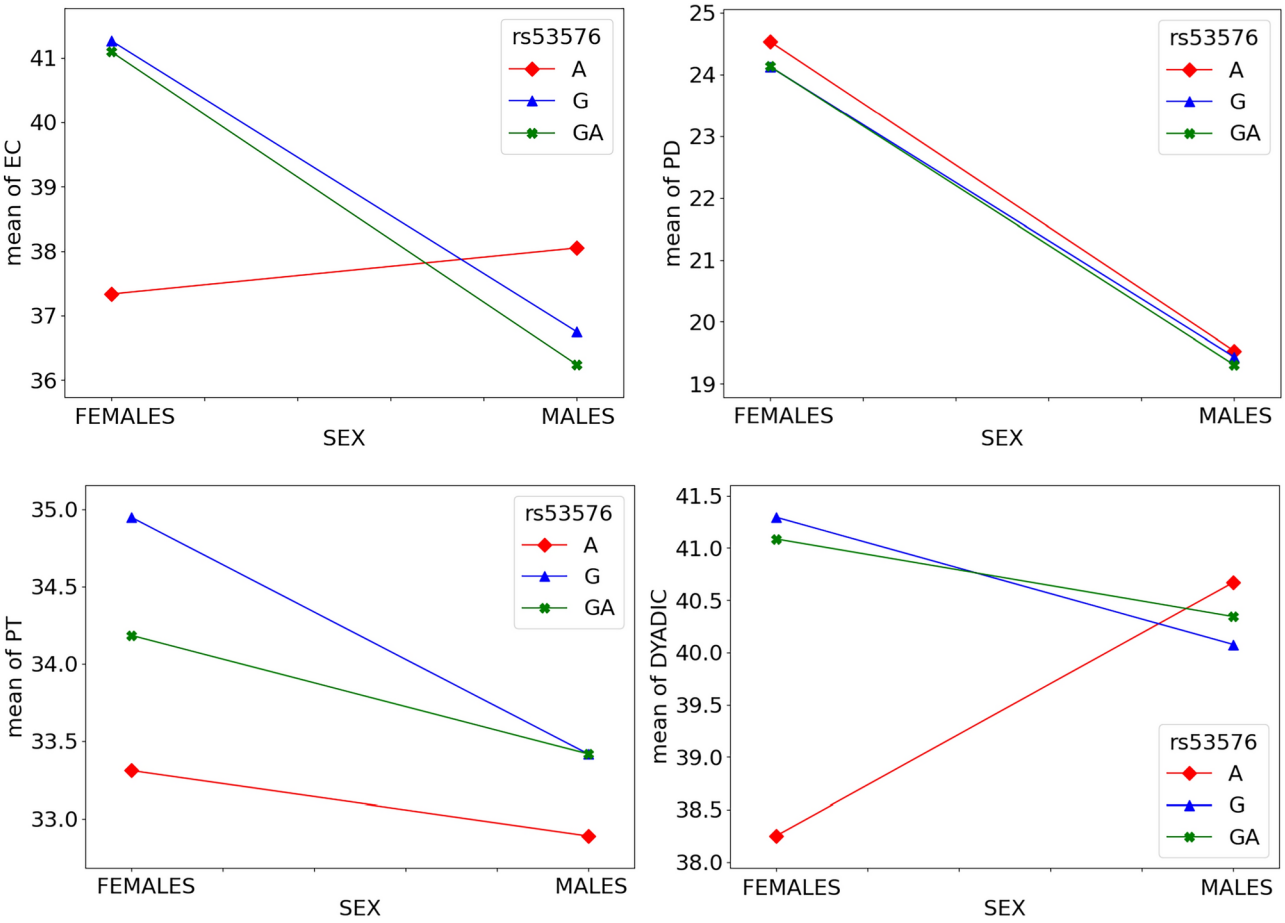

Fig. 4. Interaction plots for interactions between gender (SEX) and rs53576, resulting from the RF modeling, for EC, PD, PT and DYADIC, for the whole group. The crossed lines in the case of DYADIC and EC indicates the existence of interactions.

Exploration of the role of the RS1 and RS3 microsatellite repeats within the promoter region of *AVPRA1* in promoting empathy revealed its effect on empathic emotional contagion, and this effect was confirmed by all of the statistical analyses. Long repeats of RS1 were linked to higher scores in PD as compared with shorter variants (a strong effect obtained for both LME and RF). Moreover, individuals with short repeats of RS3 seemed to show more dyadic empathy (the effect was obtained only in RF). These results did not confirm the conclusion that individuals with shorter RS1-RS3 repeats are less prosocial[53], as they declared less self-oriented emotional contagion and experienced and expressed more partner-oriented empathy. However, earlier findings by Uzefovsky et al.[14] indicated that longer variants of RS3 were related to lower cognitive empathy, as measured by the IRI, and in our study, women with medium variants scored the lowest in the PT (confirmation with RF). These results, similar to our findings, might reflect subjective views on oneself and intimate relationships more than particular prosocial behaviors. In our earlier analyses, shorter variants of RS1 and RS3 were linked to lower levels of AVP after performing the task of simulated caregiving with an infant simulator[77], which could better reflect the impact of RS1 and RS3 microsatellite repeats on social behaviors. Thus, more research on *AVPRA1* alleles within the domain of intimate relationships with various types of empathy is needed.

Our analyses indicated that the inclusion of the *ESR1* SNP was particularly justified because it predicted EC and dyadic empathy scores, and these effects seemed to be nonlinear. AG heterozygotes scored the lowest in both dimensions. Additionally, AG heterozygotes obtained lower PT scores. These were the exploratory findings. These associations seem to reflect the role of neuroendocrine/hormonal factors in the effects of certain neural circuits (both empathy and ESR1 have been linked to, e.g., the amygdala or hypothalamus) on emotional and social reactions[8,78]. It might be concluded that some of the earlier findings indicated that AA genotypes were associated with lower harsh parenting[79], whereas G alleles were linked to selected personality traits, such as higher neuroticism, harm avoidance or imaginative thinking[80], but associations of the rs1884051 SNP with depressive symptoms are complex[81]. The ESR1 gene has been associated with the regulation of arousal[81], stress-related disorders, and social memory, especially in women[82,83]. As mentioned, the emotional regulation of responses to stressful stimuli and a focus on others' situations are crucial for the EC and PT[22]. One study revealed that the ESR1 promoter polymorphism rs2504063 was associated with increased social memory via voice recognition in women[83]. Another study showed that ESR1 and ESR2 gene polymorphisms modulated facial sadness recognition in healthy females[78]. Although not directly related to EC or dyadic empathy, these

findings suggest that ESR1 or ESR2 variants may influence certain social abilities. Thus, our results provide novel insight into the relationship between the *ESR1* SNP and empathy, which is understood as a personality trait, in the context of a particular, intimate relationship.

Finally, *HTR2A* SNPs could also be linked to empathy, as our results indicated that rs6311 TT homozygotes reported lower dyadic empathy than did the rest of the sample, especially in women, and the effect was nonlinear. Another effect emerged for PD, with rs6311 TT homozygotes scoring higher, again especially in women. Additionally, CT heterozygotes scored the highest in PT. These findings were also exploratory.

Interestingly, the polymorphisms described above predicted the congruence in dyadic empathy between partners. It is worth emphasizing that in the absence of observations of the partners' behaviors, as in[32], we managed to explore the subjective similarity in empathy experienced and expressed in dyads. Since higher dyadic empathy was previously linked to greater and relatively stable synchrony of positive reactions in a relationship and in various challenging situations that elicit the need for support, it might reflect more general tendencies of individuals (e.g., dispositional empathy, attachment[24]). It might also be predicted by genetic factors, as the presented results suggest. However, referring to the context of intimate couples, it has been concluded that partners focusing empathically on interactions in a romantic relationship to a greater extent perceive themselves and their partner as more similar, which might reflect a better emotional atmosphere in a relationship[84] and greater relationship satisfaction[1,85]. Thus, such congruence also reflects a high-quality intimate relationship, which, as we innovatively showed, partly depends on SNPs related to *OXTR* and *AVPR1a* polymorphisms and variation within genes encoding receptors that regulate oxytocin and vasopressin secretion (ESR1 and HTRA2). In couples, rs1884051 AG heterozygotes (particularly men) were the most incongruent in their experienced and mutually expressed empathy. The same seemed to be true for partners with short repeats of RS1 and for rs6311 CC homozygotes, who were not, relatively, the least empathic, but in these couples, women might not receive as much empathy as they provide to the relationship. Thus, in these couples, women and men differed relatively more in dyadic empathy than in the rest of the sample. This particular result indicates the need for further exploration of SNPs effects on dyadic empathy.

Strengths and limitations

Since all psychological variables were self-reported and measured simultaneously, it was impossible to draw conclusions on causal links between dispositional and dyadic empathy. Studies with longitudinal designs should be conducted to determine the causal relationships between genetic variations and empathy. Although self-reports might be biased, we should emphasize that we used one of the most popular and highly regarded measures of individual differences in empathy[86]. Additionally, the concept of dyadic empathy stemmed from Davis's model[25]. In earlier research, EC measured with the Empathic Sensitiveness Scale positively correlated with observed parental sensitivity[87,88]. Similarly, observations and psychological assessments of empathic dyadic interactions would allow us to analyze empathy in couples in a more objective way, which is worth including in future studies.

The study was conducted on a relatively large sample of couples, and we analyzed the same genetic variations of both partners. However, this study focused on cohabiting, childless couples. This limits the applicability of the findings to other life stages, such as couples with children or older couples, who might differ in empathy levels and expression. Nevertheless, the study design included both women and men, assessing their self-reported empathic responses toward their intimate partners with whom they share a household. As previously mentioned, dyadic empathy appears independent of relationship duration. However, prior research on couples suggests that the strength of the association between empathy and specific behaviors might change over time. Thus, another limitation of this study is that it did not focus on specific interactions between partners. Additionally, we tested whether the significant life-role transition of expecting a first child affects the relationship between genetic factors and empathy levels. However, we did not account for various life stressors that couples may face, which might impact their willingness to express empathy within their relationship. Expanding the research on more diversified samples, regarding the relationship status, duration or quality, would allow for more complex models with covariates.

Due to the implementation of machine learning techniques, we were able to explore not only linear but also nonlinear links between genetic predictors and empathic dimensions. The machine models used (linear mixed effects and random forests) are robust to overfitting and allow measurements of the significance of variables, allowing for insights into the most influential variables in the model. The choice of methods allowed for nonlinear modeling of the relationships within the data, with relatively little computational effort and high replicability. However, it is important to bear in mind that almost every machine model is subject to uncertainties, such as those associated with imperfect data, e.g., unbalanced data. In addition, the rigorous approach we used to account for compensation for multiple comparisons meant that, in some cases, the adjusted p-value rose (slightly) above the threshold for statistical significance, but with sufficient effect size. Therefore, the obtained results require independent confirmation. Additionally, there is no natural cutoff that can be used to discriminate between important and nonimportant variables when using variable importance measures in RF modeling, so the selection of the most important IVs, even based on the dropout loss metric, might be considered somewhat subjective.

Gender was the main factor impacting the results, which further expands the knowledge on biological (i.e., genetic) and sociocultural (i.e., gender roles) predictors of empathy that might interact. Furthermore, selected but multiple SNPs were analyzed, and certain effects were confirmed in more than one analysis, going beyond *OXTR* polymorphisms. Such findings might be important for developmental and relationship researchers and practitioners.

Taken together, our results expand the knowledge on the underpinnings of individual differences in empathy and indicate that not only dispositional but also dyadic empathy might be predicted by genetic variations linked

to oxytocin, vasopressin, serotonin and estrogen. The findings confirm the different effects of the analyzed SNPs on emotional and cognitive empathy and might be useful for understanding the mechanisms of psychopathology related to a lack of empathy. This study uniquely shows that inviting couples to participate in research on genetic predictors of empathy might open new venues for understanding the sources of partners' dyadic interactions.

Methods

Participants

The sample consisted of 442 adults from Northern Poland. They were partners in heterosexual relationships (N = 221). The intimate relationships had lasted on average about 6 years (M = 5.77, SD = 3.03), and all couples were cohabiting and childless. The majority of couples were in formal relationships (n = 122 married). 111 couples expected their first child. Female partners were 27 yrs. old on average: M = 26.66, SD = 3.24, as were their male partners: M = 27.71, SD = 3.53. Most participants worked in various occupations—78% of women and 92% of men, and were highly educated—84% and 73%, respectively.

DNA extraction

Cheek cells were collected from all participants by scraping the inside of the cheek with a cotton swab and stored at − 80 °C. Genomic DNA was isolated using the QIAamp DNA Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. NanoDrop® ND-1000 (Thermo Fisher Scientific, USA) was used to verify the quantity and quality of purified DNA. DNA was stored at − 20 °C.

Genotyping

Based on the literature review, we selected 9 single nucleotide polymorphisms (SNPs) within *OXTR*, (rs1042778, rs13316193, rs2228485, rs2254298, rs2268494, rs4686302, rs53576, rs237887, rs7632287), two in *HTRA2* (rs6311, rs6314) one in *ESR1* (rs1884051), and two *AVPRA1* microsatellites (RS1 and RS3). The MassARRAY ® 4 instrument, which combines mass spectrometry with endpoint PCR, was used to genotype all 12 SNPs. Microsatellite fragment analysis was performed by denaturing polyacrylamide gel electrophoresis. Both methods have been described in detail previously[77].

All 14 genetic variations analyzed were variable in the study group, but for 4 *OXTR* SNPs the genotype distribution was not in Hardy–Weinberg equilibrium (HWE): rs1042778, rs13316193, rs2268490, rs53576 ($p < 0.05$). Among the SNPs we investigated, *OXTR* rs237887 had the highest minor allele frequency (MAF) in our study population (42%), and *OXTR* rs2268494 and *HTRA2* rs6314 had the lowest MAF (6% each). Analysis of the LD between the variants using the Ensembl calculator showed that all SNPs were independent ($r2 < 0.4$). The RS1 microsatellite repeat within the promoter region of *AVPRA1* did not deviate from the HWE ($p = 0.66$), in contrast to the RS3 repeat ($p = 0.0001$). Based on the number of base pairs identified, the RS1 and RS3 repeats were grouped into 'long' (L) and 'short' (S) groups. For RS3, 'long' refers to repeats that are longer than 310 bp, and 'short' refers to all repeats that are shorter than 305 bp. For RS1, 'short' includes a fragment of 274 bp and all the shorter repeats, while 'long' includes all the repeats that are longer than 280 bp.

Measures

Dispositional empathic concern, personal distress and perspective taking

The Empathic Sensitiveness Scale: the widely used Polish version of the Interpersonal Reactivity Index by Davis[19] was used[62]. It consists of 28 items with a 5-point Likert response scale (1—does not describe me well, 5—I completely agree) and measures three empathic dimensions: emotional and other-oriented EC, emotional and self-oriented PD, and cognitive and other-oriented PT. The Cronbach's α for each of the subscales exceeded 0.7 in this study. Sample items: *'I often have tender, concerned feelings for people less fortunate than me'* (EC); *'When I see someone who badly needs help in an emergency, I go to pieces'* (PD); *'I try to look at everybody's side of a disagreement before I make a decision'* (PT).

Dyadic empathy (DYADIC)

The Interpersonal Reactivity Index for Couples[24]. The 10-item measure with a 5-point Likert response scale (1—do not agree at all/does not describe me well, 5—I completely agree/describes me very well) was used to measure empathy in the context of the intimate relationship and directed toward a partner (the concept of the scale was first introduced by Péloquin and Lafontaine[25]). The Polish version consists of two subscales: empathic concern (4 items) and perspective taking (6 items) of partners in the romantic relationship. However, in the Polish version, a total score of partner-oriented empathy has been recommended[24]. For the total score, Cronbach's α was 0.73 for women and 0.77 for men. Sample items: *'When I see my partner being taken advantage of, I feel kind of protective towards him/her'; 'Before criticizing my partner, I try to imagine how I would feel if I were in his/her place'*. Additionally, the variable resulting from the subtraction of the corresponding DYADIC values (for women and men) was created. Dyadic empathy congruence (DYADIC CONGRUENCE) assesses the similarity between partners in partner-oriented empathy. The higher the score, the greater the difference between partners, with higher scores for women.

Procedure

The presented research aim is a part of a larger project. The project explored the effects of empathy in couples who were not parents themselves on their responsiveness during simulated parental caregiving. In this paper, we uniquely examine selected genetic predictors of the main variable in the project, that is empathy (dispositional and dyadic). Referring to the Introduction, the inclusion of expectant couples allowed us to explore whether this significant life transition—marked by an increased need for empathy—alters the strength of associations between genetic factors and levels of EC, PD, PT, and dyadic empathy. Additionally, none of the participants

were in the early stage of dating, ensuring a focus on more established relationships. Invitations to participate in the study were distributed through social media and in institutions conducting antenatal classes. Inclusion criteria included being in emerging and young adulthood and in general good health. All participants had to be in a stable intimate relationship for at least 2 years, cohabiting, and with no children. The couples should not have significantly differed from each other on sociodemographic characteristics. Additionally, they were generally happy with their relationships (M = 8.46, SD = 1.27, score range, 1–10) and fairly satisfied with their economic life (M = 6.84, SD = 1.45, score range, 1–10).

Upon arrival at the laboratory, a salivary sample was collected from each participant to obtain the genetic material. Then, the participants completed the psychological questionnaires individually. The recruitment process and procedure have been described in detail by Kaźmierczak et al.[75]. Participants received 100 PLN (ca. 25 Euro) at the completion of the study. Informed consent was obtained from all participants. This study was performed in accordance with the conditions set out in the GDPR and the requirements of the Independent Bioethics Committee for Scientific Research at the Medical University of Gdańsk, Poland, and the Ethics Committee at the Institute of Psychology, University of Gdańsk, Poland. This study was approved by the Independent Bioethics Committee for Scientific Research at the Medical University of Gdansk, Poland (permission # NKBBN/154/2017) and the Ethics Committee at the Institute of Psychology, University of Gdansk, Poland (permission # 4/2016).

Data availability

The data that support the findings of this study are available in OSF https://osf.io/g9xwy/?view_only=89c6540489524a98922d775eae12c186 ; https://osf.io/2qxtj/?view_only=743bab4a38ea4cd493574d760e496879. In addition to random forests, artificial neural networks (ANNs) have been used as alternative nonlinear modeling methods. Five models were trained for each dependent variable, with a set of hyperparameters found during the tuning procedure. The final network architecture is presented in Supplementary Fig. S9. To prevent overfitting, each model was subjected to the k-fold cross-validation procedure (k = 10), yielding 10 independent submodels in each case. Then, the final models were composed of an ensemble (averaged) model made from the above submodels. Next, the significance of the individual IVs was calculated for each model using explanatory methods, as was the case for the RF models. The results are presented in Supplementary Table S6. The vast majority of the results overlap with those obtained by the RF method. Finally, RF was chosen as the main method presented and used in the article because of its lower computational cost and (in general) better explainability.

Received: 12 March 2024; Accepted: 4 November 2024
Published online: 09 November 2024

References

1. Busby, D. M. & Gardner, B. C. How do I analyze thee? Let me count the ways: considering empathy in couple relationships using self and partner ratings. *Fam. Process* **47**, 229–242 (2008).
2. Kaźmierczak, M. Couple empathy—the mediator of attachment styles for partners adjusting to parenthood. *J. Reprod. Infant Psychol.* **33**, 15–27 (2015).
3. Shalev, I., Eran, A. & Uzefovsky, F. Fluctuations and individual differences in empathy interact with stress to predict mental health, parenting, and relationship outcomes. *Front. Psychol.* **14**, 1237278 (2023).
4. Stern, J. A., Borelli, J. L. & Smiley, P. A. Assessing parental empathy: a role for empathy in child attachment. *Attach. Hum. Dev.* **17**, 1–22 (2015).
5. Carter, C. S., Harri, J. & Porges, S. W. Neural and evolutionary perspectives on empathy in *The social neuroscience of empathy* (eds. Decety, J. & Ickes, W.) 169–182 (MIT Press, 2011).
6. de Waal, F. B. M. & Preston, S. D. Mammalian empathy: behavioural manifestations and neural basis. *Nat. Rev. Neurosci.* **18**, 498–509 (2017).
7. Abramson, L., Uzefovsky, F. & Knafo-Noam, A. What do we (not) know about the genetics of empathy? in *The social brain: a developmental perspective* (ed. Decety, J.) 263–284 (The MIT Press, 2020).
8. Decety, J. The neural pathways, development and functions of empathy. *Curr. Opin. Behav. Sci.* **3**, 1–6 (2015).
9. Decety, J., Bartal, I. B. A., Uzefovsky, F. & Knafo-Noam, A. Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150077 (2016).
10. Toccaceli, V. et al. Adult empathy: possible gender differences in gene-environment architecture for cognitive and emotional components in a large Italian twin sample. *Twin Res. Hum. Genet.* **21**, 214–226 (2018).
11. Shamay-Tsoory, S. G. & Abu-Akel, A. The social salience hypothesis of oxytocin. *Biol. Psychiatry* **79**, 194–202 (2016).
12. Drolet, C. E. & Lucas, T. The neurogenetics of racial injustice: oxytocin receptor (OXTR) gene rs53576 is ASSOCIATED with perceived discrimination and other-oriented justice beliefs in African Americans. *Race Soc. Probl.* **13**, 102–109 (2021).
13. Kemp, A. H. & Guastella, A. J. The role of oxytocin in human affect: a novel hypothesis. *Curr. Dir. Psychol. Sci.* **20**, 222–231 (2011).
14. Uzefovsky, F. et al. Oxytocin receptor and vasopressin receptor 1a genes are respectively associated with emotional and cognitive empathy. *Horm. Behav.* **67**, 60–65 (2015).
15. Olofsdotter, S., Åslund, C., Furmark, T., Comasco, E. & Nilsson, K. W. Differential susceptibility effects of oxytocin gene (OXT) polymorphisms and perceived parenting on social anxiety among adolescents. *Dev. Psychopathol.* **30**, 449–459 (2018).
16. Monin, J. K., Goktas, S. O., Kershaw, T. & DeWan, A. Associations between spouses' oxytocin receptor gene polymorphism, attachment security, and marital satisfaction. *PLoS One* **14**, e0213083 (2019).
17. Ulloa, E. C., Hammett, J. F., Meda, N. A. & Rubalcaba, S. J. Empathy and romantic relationship quality among cohabitating couples. *Fam. J.* **25**, 208–214 (2017).
18. Davis, M. H. A multidimensional approach to individual differences in empathy. *J. Pers. Soc. Psychol.* **10**, 85 (1980).
19. Davis, M. H. Measuring individual differences in empathy: evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* **44**, 113–126 (1983).
20. Davis, M. H. Empathy in *Handbook of the sociology of emotions* (eds. Stets, J. E. & Turner, J. H.) 443–466 (Springer, 2006).
21. Batson, C. D. *These things called empathy: Eight related but distinct phenomena*. In *The social neuroscience of empathy* (eds. Decety, J. & Ickes, W.) 3–15 (MIT Press, 2009).
22. Eisenberg, N. & Eggum, N. D. Empathic responding: sympathy and personal distress in *The social neuroscience of empathy* (eds. Decety, J. & Ickes, W.) 71–83 (The MIT Press, 2009).
23. Davis, M. H. *Empathy: a social psychological approach* (Routledge, 2018).

24. Kaźmierczak, M. & Karasiewicz, K. Dyadic empathy in Polish samples: Validation of the interpersonal reactivity index for couples. *Curr. Issues Pers. Psychol.* **9**, 354–365 (2021).
25. Péloquin, K. & Lafontaine, M. F. Measuring empathy in couples: validity and reliability of the interpersonal reactivity index for couples. *J. Pers. Assess.* **92**, 146–157 (2010).
26. Rosen, N. O., Mooney, K. & Muise, A. Dyadic empathy predicts sexual and relationship well-being in couples transitioning to parenthood. *J. Sex Marital Ther.* **43**, 543–559 (2016).
27. Carasso, E. & Segel-Karpas, D. Marital strain and emotional intimacy in midlife couples: The moderating role of empathy. *Pers. Relatsh.* **31**(3), 648–663 (2024).
28. Devoldre, I., Davis, M. H., Verhofstadt, L. L. & Buysse, A. Empathy and social support provision in couples: Social support and the need to study the underlying processes. *J. Psychol.* **144**(3), 259–284 (2010).
29. Beadle, J. N. & de la Vega, C. E. Impact of aging on empathy: review of psychological and neural mechanisms. *Front. Psychiatry* **10**, 331 (2019).
30. Grühn, D., Rebucal, K., Diehl, M., Lumley, M., & Labouvie-Vief, G. Empathy across the adult lifespan: Longitudinal and experience-sampling findings. *Emotion* (Washington, D.C.) **8(6)**, 753–765 (2008).
31. Jarvis, A. L. et al. Emotional empathy across adulthood: A meta-analytic review. *Psychol. Aging.* **39**(2), 126 (2024).
32. Schneiderman, I., Kanat-Maymon, Y., Ebstein, R. P. & Feldman, R. Cumulative risk on the oxytocin receptor gene (OXTR) underpins empathic communication difficulties at the first stages of romantic love. *Soc. Cogn. Affect. Neurosci.* **9**, 1524–1529 (2014).
33. Kimmes, J. G., Edwards, A. B., Wetchler, J. L. & Bercik, J. Self and other ratings of dyadic empathy as predictors of relationship satisfaction. *Am. J. Fam. Ther.* **42**, 426–437 (2014).
34. Warrier, V. et al. Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Mol. Psychiatry* **23**, 1402–1409 (2018).
35. Knafo, A. & Uzefovsky, F. Variation in empathy: the interplay of genetic and environmental factors. In *The infant mind: origins of the social brain* (eds. Legerstee, M., Haley, D. W., & Bornstein, M. H.) 97–120 (The Guilford Press, 2013).
36. Melchers, M., Montag, C., Reuter, M., Spinath, F. M. & Hahn, E. How heritable is empathy? Differential effects of measurement and subcomponents. *Motiv. Emot.* **40**, 720–730 (2016).
37. Hatemi, P. K., Smith, K., Alford, J. R., Martin, N. G. & Hibbing, J. R. The genetic and environmental foundations of political, psychological, social, and economic behaviors: a panel study of twins and families. *Twin Res. Hum. Genet.* **18**, 243–255 (2015).
38. Warrier, V. et al. Genome-wide analyses of self-reported empathy: Correlations with autism, schizophrenia, and anorexia nervosa. *Transl. Psychiatry* **8**, 35 (2018).
39. Gong, P. et al. Revisiting the impact of OXTR rs53576 on empathy: a population-based study and a meta-analysis. *Psychoneuroendocrinology* **80**, 131–136 (2017).
40. Barchi-Ferreira, A. M. & Osório, F. L. Associations between oxytocin and empathy in humans: a systematic literature review. *Psychoneuroendocrinology* **129**, 105268 (2021).
41. Chander, R. J. et al. The influence of rs53576 polymorphism in the oxytocin receptor (OXTR) gene on empathy in healthy adults by subtype and ethnicity: a systematic review and meta-analysis. *Rev. Neurosci.* **33**, 43–57 (2022).
42. Bakermans-Kranenburg, M. J. & Van Ijzendoorn, M. H. A sociability gene? Meta-analysis of oxytocin receptor genotype effects in humans. *Psychiatr. Genet.* **24**, 45–51 (2014).
43. Li, J. et al. Association of oxytocin receptor gene (OXTR) rs53576 polymorphism with sociality: a meta-analysis. *PLoS One* **10**, e0131820 (2015).
44. Pearce, E., Wlodarski, R., Machin, A. & Dunbar, R. I. M. Variation in the β-endorphin, oxytocin, and dopamine receptor genes is associated with different dimensions of human sociality. *Proc. Natl. Acad. Sci. USA.* **114**, 5300–5305 (2017).
45. Huetter, F. K. et al. Association of a common oxytocin receptor gene polymorphism with self-reported "empathic concern" in a large population of healthy volunteers. *PLoS One* **11**, e0160059 (2016).
46. Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P. & Keltner, D. Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21437–21441 (2009).
47. Israel, S. et al. The oxytocin receptor (OXTR) contributes to prosocial fund allocations in the dictator game and the social value orientations task. *PLoS One* **4**, e5535 (2009).
48. Christ, C. C., Carlo, G. & Stoltenberg, S. F. Oxytocin receptor (OXTR) single nucleotide polymorphisms indirectly predict prosocial behavior through perspective taking and empathic concern. *J. Pers.* **84**, 204–213 (2016).
49. Wu, N., Li, Z. & Su, Y. The association between oxytocin receptor gene polymorphism (OXTR) and trait empathy. *J. Affect. Disord.* **138**, 468–472 (2012).
50. Schöner, J. *On empathy, memory and genetics: what role does human age play? (Unpublished Master's Thesis)* (Stockholm University, 2012).
51. Kalyoncu, T., Özbaran, B., Köse, S. & Onay, H. Variation in the oxytocin receptor gene is associated with social cognition and ADHD. *J. Atten. Disord.* **23**, 702–711 (2019).
52. Montag, C. et al. Association between oxytocin receptor gene polymorphisms and self-rated "empathic concern" in schizophrenia. *PLoS One* **7**, e51882 (2012).
53. Zhang, Y. et al. Neural mechanisms of AVPR1A RS3-RS1 haplotypes that impact verbal learning and memory. *NeuroImage* **222**, 117283 (2020).
54. Murakami, G., Hunter, R. G., Fontaine, C., Ribeiro, A. & Pfaff, D. Relationships among estrogen receptor, oxytocin and vasopressin gene expression and social interaction in male mice. *Eur. J. Neurosci.* **34**, 469–477 (2011).
55. Markham, A., Bains, R., Franklin, P. & Spedding, M. Changes in mitochondrial function are pivotal in neurodegenerative and psychiatric disorders: how important is BDNF?. *Br. J. Pharmacol.* **171**, 2206–2229 (2014).
56. Toyama, Y., Harkness, R. W. & Kay, L. E. Structural basis of protein substrate processing by human mitochondrial high-temperature requirement A2 protease. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2203172119 (2022).
57. Mattson, R. E. et al. Oxytocin receptor gene (OXTR) links to marital quality via social support behavior and perceived partner responsiveness. *J. Fam. Psychol.* **33**, 44–53 (2019).
58. Eisenberg, N. & Lennon, R. Sex differences in empathy and related capacities. *Psychol. Bull.* **94**, 100–131 (1983).
59. Joel, S. et al. Machine learning uncovers the most robust self-report predictors of relationship quality across 43 longitudinal couples studies. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 19061–19071 (2020).
60. Hey, T., Butler, K., Jackson, S. & Thiyagalingam, J. Machine learning and big scientific data. *Philos. Trans., Math. Phys. Eng. Sci.* **378**, 20190054 (2020).
61. Mertz, D. *Cleaning data for effective data science* (Packt Publishing, 2021).
62. Kaźmierczak, M., Plopa, M. & Retowski, S. Empathic sensitiveness scale. *Prz. Psychol.* **50**, 9–24 (2007).
63. Mittag, F., Römer, M. & Zell, A. Influence of feature encoding and choice of classifier on disease risk prediction in genome-wide association studies. *PLoS One* **10**, e0135832 (2015).
64. Gelman, A. & Hill, J. *Data analysis using regression and multilevel/hierarchical models* (Cambridge University Press, 2006).
65. Raudenbush, S. W. & Bryk, A. S. *Hierarchical linear models: applications and data analysis methods (advanced quantitative techniques in the social sciences)* (SAGE Publications, 2001).
66. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
67. Biecek, P. & Burzykowski, T. *Explanatory model analysis* (Chapman and Hall/CRC, 2021).

68. McKinney, B. A., Reif, D. M., Ritchie, M. D. & Moore, J. H. Machine learning for detecting gene-gene interactions: A review. *Appl. Bioinform.* **5**, 77–88 (2006).
69. Guo, W., Romano, J. A. Generalized Sidak-Holm procedure and control of generalized error rates under independence. *Stat. Appl. Genet. Mol. Biol.* **6** (2007).
70. Hastie, T., Tibshirani, R. & Friedman, J. *The elements of statistical learning* (Springer, 2009).
71. Fisher, A., Rudin, C. & Dominici, F. All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *J. Mach. Learn. Res.* **20**, 1–81 (2019).
72. Janitza, S., Celik, E., Boulesteix, A. L. A computationally fast variable importance test for random forests for high-dimensional data. Adv. *Data Anal. Classif.* **12** (2016).
73. Hoffman, M. L. *Empathy and moral development: implications for caring and justice* (Cambridge University Press, 2001).
74. Davis, M. H. Empathy: negotiating the border between self and other in *The social life of emotions* (eds. Leach, C. W. & Tiedens, L. Z.) 19–42 (Cambridge University Press, 2004).
75. Kaźmierczak, M. et al. Empathy and hormonal changes as predictors of sensitive responsiveness towards infant crying: a study protocol. *Int. J. Environ. Res. Public Health* **18**, 4815 (2021).
76. Kaźmierczak, M., van Ijzendoorn, M. H. & Bakermans-Kranenburg, M. Do empathy and oxytocin predict responsiveness to a crying infant simulator in expecting and non-expecting couples? A multilevel study. *Attach. Hum. Dev.* **24**, 624–644 (2022).
77. Rybicka, M. et al. (Re-)activity in the caregiving situation: genetic diversity within Oxytocin-Vasopressin Pathway is associated with salivary oxytocin and vasopressin concentrations in response to contact with a crying infant-simulator. *Psychoneuroendocrinology* **131**, 105294 (2021).
78. Gutiérrez-Muñoz, M., Fajardo-Araujo, M. E., González-Pérez, E. G., Aguirre-Arzola, V. E. & Solís-Ortiz, S. Facial sadness recognition is modulated by estrogen receptor gene polymorphisms in healthy females. *Brain Sci.* **8**(12), 219 (2018).
79. Lahey, B. B. et al. Preliminary genetic imaging study of the association between estrogen receptor-α gene polymorphisms and harsh human maternal parenting. *Neurosci. Lett.* **525**, 17–22 (2012).
80. Miller, A. et al. Estrogen receptor alpha (ESR-1) associations with psychological traits in women with PMDD and controls. *J. Psychiatr. Res.* **44**, 788–794 (2010).
81. Huo, L. et al. Risk for premenstrual dysphoric disorder is associated with genetic variation in ESR1, the estrogen receptor alpha gene. *Biol. Psychiatry* **62**, 925–933 (2007).
82. Mikhailova, S. V. et al. Assessment of the genetic characteristics of a generation born during a long-term socioeconomic crisis. *Genes (Basel)* **14**(11), 2064 (2023).
83. Karlsson, S. et al. Social memory associated with estrogen receptor polymorphisms in women. *Soc. Cogn. Affect. Neurosci.* **11**(6), 877–883 (2016).
84. Andrews, D. W. Perspective taking as a predictor of marital adjustment. *J. Pers. Soc. Psychol.* **59**(1), 126–131 (1990).
85. Cramer, D. & Jowett, S. Perceived empathy, accurate empathy and relationship satisfaction in heterosexual couples. *J. Soc. Pers. Relation.* **27**, 327–349 (2010).
86. Ilgunaite, G., Giromini, L. & Di Girolamo, M. Measuring empathy: A literature review of available tools. *BPA-Appl. Psychol. Bull. (Bollettino di Psicologia Applicata)* **65**(280), 2–28 (2017).
87. Anikiej-Wiczenbach, P., Kaźmierczak, M. & Czerwiński, S. Exploring the key drivers of responsive parenting in mothers and fathers—observed and self-reported responsiveness. *Health Psychol. Rep.* **12**(1), 39 (2024).
88. Kaźmierczak, M. et al., Couples' empathy and sensitive responsiveness to a crying baby simulator. *J. Soc. Pers. Relat.* 02654075241240078 (2024).

Acknowledgements

This work was supported by the National Science Centre, Poland (2016/22/E/HS6/00237). The funding source had no involvement in research and/or preparation of the article.

Author contributions

MK—Study design, Data collection, Statistical analysis, Data interpretation, Manuscript preparation, Literature search, Funds collection. MR-M—Data collection, Statistical analysis, Data interpretation, Manuscript preparation. PS—Data preparation and preprocessing, Training machine learning models, Visualization and interpretation of prediction results, Manuscript preparation.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-78857-7.

Correspondence and requests for materials should be addressed to P.S.

Reprints and permissions information is available at www.nature.com/reprints.

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