The Oxytocin Receptor Gene (OXTR) and Face Recognition

Roeland J. Verhallen1, Jenny M. Bosten1,2, Patrick T. Goodbourn1,3, Adam J. Lawrance-Owen1, Gary Bargary1,4, and J. D. Mollon1

1Department of Psychology, University of Cambridge; 2School of Psychology, University of Sussex; 3School of Psychological Sciences, University of Melbourne; and 4Division of Optometry and Visual Science, City University London

Abstract
A recent study has linked individual differences in face recognition to rs237887, a single-nucleotide polymorphism (SNP) of the oxytocin receptor gene (OXTR; Skuse et al., 2014). In that study, participants were assessed using the Warrington Recognition Memory Test for Faces, but performance on Warrington’s test has been shown not to rely purely on face recognition processes. We administered the widely used Cambridge Face Memory Test—a purer test of face recognition—to 370 participants. Performance was not significantly associated with rs237887, with 16 other SNPs of OXTR that we genotyped, or with a further 75 imputed SNPs. We also administered three other tests of face processing (the Mooney Face Test, the Glasgow Face Matching Test, and the Composite Face Test), but performance was never significantly associated with rs237887 or with any of the other genotyped or imputed SNPs, after corrections for multiple testing. In addition, we found no associations between OXTR and Autism-Spectrum Quotient scores.

Keywords
OXTR, oxytocin receptor gene, rs237887, Cambridge Face Memory Test, face recognition, Autism-Spectrum Quotient

Received 2/2/16; Revision accepted 9/12/16

Polymorphisms of the oxytocin receptor gene (OXTR) have been linked to individual variations in an array of human social behaviors, including maternal sensitivity, empathy, prosocial behavior, and recognition of affect (Ebstein, Knafo, Mankuta, Chew, & Lai, 2012; Feldman, Monakhov, Pratt, & Ebstein, 2016; Kogan et al., 2011; Poulin, Holman, & Buffone, 2012). Genetic variations in OXTR have also been linked to autism spectrum disorder (Di Napoli, Warrier, Baron-Cohen, & Chakrabarti, 2014; LoParo & Waldman, 2015). However, although OXTR has been a favorite candidate gene for influencing social behaviors, the evidence is not consistent (Cornelis et al., 2012; Kiy, Wilhelm, Hildebrandt, Reuter, & Sommer, 2013). Thus a recent meta-analysis of rs53576 and rs2254298, two commonly cited single-nucleotide polymorphisms (SNPs) of OXTR, found that the average effect size did not differ from zero for either empathetic aspects of personality or for social behaviors (Bakermans-Kranenburg & van IJzendoorn, 2014). In addition, an analysis of OXTR polymorphisms in a large sample found no overall association with a measure of social integration that was based on marital status, contact with close friends, and participation in community groups (Chang et al., 2014).

A recent article added face recognition to the list of behaviors positively associated with OXTR. Skuse et al. (2014) reported a significant association between performance on the Warrington Test of Recognition Memory for Faces (WPS, Torrance, CA; referred to hereafter as “the Warrington test”) and the SNP rs237887, which lies in the third and last intron of OXTR. The subjects were a subset (n = 333) of a sample of high-functioning children with autism and their first-degree relatives (i.e., parents and siblings). Of the 18 SNPs of OXTR genotyped by Skuse et al., rs237887 was the only one found to be
significantly associated with performance on the Warrington test. This effect was mostly driven by the neurotypical first-degree relatives, despite the study's focus on children with autism. Skuse et al. genotyped an additional 42 SNPs from the region surrounding this gene, but none were significantly associated with performance on the Warrington test.

However, the validity of the Warrington test as a pure index of face recognition ability has previously been questioned: Duchaine and Weidenfeld (2003) observed that some participants could perform normally on this test even when the internal features of the face stimuli had been removed. Duchaine and Nakayama (2006) went on to develop the Cambridge Face Memory Test, which has since been used in a wide range of studies to investigate both neurotypical subjects and subjects with prosopagnosia (e.g., Busigny, Joubert, Felician, Ceccaldi, & Rossion, 2010; Germaine et al., 2012; Hedley, Brewer, & Young, 2011; Wilmer et al., 2010).

In the course of a genome-wide association study, we genotyped 17 SNPs of OXTR, and we investigated another 75 SNPs of OXTR using imputation methods. We used four different tests of face perception and recognition to index several aspects of face processing: the Cambridge Face Memory Test, the Mooney Face Test, the Glasgow Face Matching Test, and the Composite Face Test. In addition, we measured performance on the Autism-Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). Given its sample size, our study was well powered (99%) to detect an association between face recognition and rs237887 of a size similar to that reported by Skuse et al. (2014). It is crucial that replication studies be adequately powered so that failure to replicate can be convincingly demonstrated.

The Cambridge Face Memory Test has been administered to thousands of participants in both online and lab-based studies (Germaine et al., 2012; Wilmer, Germaine, & Nakayama, 2014). Performance on the Cambridge Face Memory Test is highly heritable, as shown by a twin study (Wilmer et al., 2010), and has little or no correlation with general intelligence (Shakeshaft & Plomin, 2015; Wilmer et al., 2014). Bate et al. (2014) reported a significant increase in performance on the Cambridge Face Memory Test after intranasal administration of oxytocin, but only for subjects with prosopagnosia; they found a similar pattern in a face-matching test (the Cambridge Face Perception Test; Duchaine, Yovel, & Nakayama, 2007).

The original Mooney Face Test is a test of face perception that has been widely used in clinical testing. It comprises 40 images that each depict a face. The images consist solely of pure black and pure white elements without any shading (Mooney, 1957). This renders the perception of the faces into an all-or-nothing question: Either the black and white elements coalesce meaningfully into a face, or they remain seemingly unrelated. In the present study, we used an online three-alternative forced-choice version of the test that incorporated the 40 original images (Verhallen et al., 2014).

In the Glasgow Face Matching Test, the participant is shown two photographs simultaneously; the photographs depict either the same person or two different people (Burton, White, & McNeill, 2010). However, photographs depicting the same person are not physically identical: The two images were obtained using different cameras, different angles, and different lighting. Results of this test show strong correlations with results from tests of face recognition (Burton et al., 2010).

In the Composite Face Test, two faces are presented, one after the other; each face is composed of the top half of one source face and the bottom half of another source face (Richler, Cheung, & Gauthier, 2011; Young, Hellawell, & Hay, 1987). On a given trial, the top half of the second face may or may not be the same as the top half of the first face. Likewise, the bottom halves may or may not differ. The participant is asked only whether the top half remains the same from one face to the next. Changes in the bottom half are thought to interfere with participants’ judgment of the top half when the two halves are aligned to form one face; when the two halves are misaligned, participants experience no interference. This difference in performance is thought to reflect holistic processing (Richler & Gauthier, 2013; Rossion, 2013), a measure that has been found to correlate with face recognition (Richler et al., 2011).

**Method**

**Participants**

The sample for analyses with the four tests of face processing consisted of 370 participants (235 women; for more detailed descriptive statistics for the four tests, see Verhallen et al., 2016). The sample for analyses with the AQ questionnaire (Baron-Cohen et al., 2001) consisted of 521 participants (333 women). Both samples are subsets of a cohort of 1,060 participants who had previously completed a battery of perceptual tests in our laboratory as part of the PERGENIC (PERceptual GENetics In Cambridge) project (Bosten et al., 2015; Goodbourn et al., 2014; Lawrance-Owen et al., 2013). Ethics permission for the study was given by the Cambridge University Psychology Research Ethics Committee. Our participants were healthy young adults between the ages of 18 and 42 (mean age = 24 years), all White. The majority were students at the University of Cambridge.
Materials

The Mooney Face Test that we used is described in Verhallen et al. (2014); the 40 original Mooney (1957) faces were used in an online three-alternative forced-choice paradigm. We used the shortened Glasgow Face Matching Test as described in Burton et al. (2010) and the Cambridge Face Memory Test as described in Duchaine and Nakayama (2006). We used the version of the Composite Face Test developed by Richler et al. (2011), which incorporates stimuli from the Max Planck Institute Face Database (Troje & Bülthoff, 1996). We administered this test according to the procedure described by Richler et al. (2011), following the complete design. Two performance variables were investigated for this test: the bolistic index—calculated by regressing participants’ performance on the misaligned congruent trials from their performance on the aligned congruent trials (see Verhallen et al., 2016)—and a more straightforward calculation of percentage correct across all trials (referred to hereafter as raw score).

Procedure

The performance data were collected online. All participants completed the four tests of face processing in the following sequence: the Mooney Face Test, the Glasgow Face Matching Test, the Cambridge Face Memory Test, and finally the Composite Face Test. The procedure for the Mooney Face Test is described in Verhallen et al. (2014); the Glasgow Face Matching Test, the Cambridge Face Memory Test, and the Composite Face Test were all administered according to their original procedures as described in their respective sources (see introduction section). Because the distribution of scores was nonnormal for all four tests, we converted raw performance scores to ranks; if two or more participants had the same score, they were all assigned the average rank of that score. Full details of the four tests, their intercorrelations, and their phenotypic correlates are given in Verhallen et al. (2016).

The 50 questions that made up the AQ questionnaire (Baron-Cohen et al., 2001) were the first items in a questionnaire that previously had been administered online as part of the PERGENIC project.

Genetics

Genetic data were collected during the original PERGENIC project. All genetic analyses were performed in PLINK (Purcell et al., 2007); imputation was performed using IMPUTE2 (Howie, Donnelly, & Marchini, 2009; Howie, Marchini, & Stephens, 2011) and 1,000 phased haplotypes from the 1000 Genomes Project Consortium (2010). In all genetic analyses reported in this article, sex was entered as a covariate, as well as the top three principal components of genetic variation, to control for population stratification (for details on genotyping and quality control, see Lawrance-Owen et al., 2013, and Goodbourn et al., 2014).

Results

Performance on the Cambridge Face Memory Test ranged from 26 to 72 trials correct ($M = 54.15$, $SD = 9.04$). If the data from Duchaine and Nakayama (2006) are taken as norms, the range of performance we observed was very wide: The $z$ scores ranged from $−4.04$ to $+1.78$; mean performance corresponded to a $z$ score of $−0.47$.

Our genetic analysis of 370 healthy White participants (235 women) revealed no significant association between performance on the Cambridge Face Memory Test and SNP rs237887—the SNP located in OXTR for which Skuse et al. (2014) found a significant association. When we entered performance on the Cambridge Face Memory Test as ranked data, the uncorrected $p$ value was $.88$ ($r^2 = 5.89 \times 10^{-3}$); for the raw performance data, the uncorrected $p$ value was $.90$ ($r^2 = 4.47 \times 10^{-3}$). For the single variant rs237887, we had greater than 99% power to detect an association (at an $\alpha$ level of .05) if the polymorphism accounted for 10% of phenotypic variance, as estimated by Skuse et al.; even a polymorphism accounting for only 2% of variance would have been detected with 86% power. The minor-allele frequency of rs237887 was 42% in our sample, which is similar to the minor allele frequency of 45% in the sample of Skuse et al.

For the 16 other SNPs within OXTR that we genotyped, uncorrected $p$ values for ranked performance on the Cambridge Face Memory Test ranged from $.06$ to $.98$ (see Table 1), whereas $r^2$ ranged from $0.0098$ to $1.70 \times 10^{-6}$. Sex was entered as a covariate in all genetic analyses presented in this article. However, even separate genetic analyses by sex of ranked performance data for the Cambridge Face Memory Test did not yield a significant association with rs237887 ($r^2 = 10^{-6}$). For the single variant rs237887, we had greater than 99% power to detect an association with $r^2 = 0.0095$ or with any of the other genotyped SNPs.

To further investigate OXTR, we imputed a 60-kilobase-pair region centered on the gene. This procedure yielded an additional 75 SNPs within OXTR. However, among these imputed SNPs, the lowest uncorrected $p$ value for an association with performance on the Cambridge Face Memory Test was only $.03$ (see Fig. 1), and no association remained significant after correction for multiple testing.

As noted, we administered three tests of face processing in addition to the Cambridge Face Memory Test: the Mooney Face Test ($M = 34.9$, $SD = 2.8$, range = 25–39),
Table 1. Uncorrected *p* Values for Associations Between Each of 17 Genotyped Single-Nucleotide Polymorphisms (SNPs) Within the Oxytocin Receptor Gene (*OXTR*) and Scores for Face Processing and Autism-Spectrum Quotient (AQ)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Cambridge Face Memory Test</th>
<th>Mooney Face Test</th>
<th>Glasgow Face Matching Test</th>
<th>Composite Face Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uncorrected <em>p</em></td>
<td>Corrected <em>p</em></td>
<td>Raw score</td>
<td>Holistic index</td>
</tr>
<tr>
<td>rs2324728</td>
<td>.014</td>
<td>.28</td>
<td>.03</td>
<td>.12</td>
</tr>
<tr>
<td>rs237884</td>
<td>.07</td>
<td>.50</td>
<td>.05</td>
<td>.30</td>
</tr>
<tr>
<td>rs1042778</td>
<td>.82</td>
<td>.97</td>
<td>.56</td>
<td>.82</td>
</tr>
<tr>
<td>rs237885</td>
<td>.38</td>
<td>.20</td>
<td>.83</td>
<td>.20</td>
</tr>
<tr>
<td>rs11706648</td>
<td>.58</td>
<td>.43</td>
<td>.99</td>
<td>.27</td>
</tr>
<tr>
<td>rs237887</td>
<td><strong>.88</strong></td>
<td><strong>.85</strong></td>
<td><strong>.75</strong></td>
<td><strong>.64</strong></td>
</tr>
<tr>
<td>rs2268490</td>
<td>.46</td>
<td>.96</td>
<td>.24</td>
<td>.14</td>
</tr>
<tr>
<td>rs237888</td>
<td>.35</td>
<td>.08</td>
<td>.18</td>
<td>.03</td>
</tr>
<tr>
<td>rs918316</td>
<td>.25</td>
<td>.96</td>
<td>.78</td>
<td>.97</td>
</tr>
<tr>
<td>rs4086301</td>
<td>.59</td>
<td>.53</td>
<td>.57</td>
<td>.09</td>
</tr>
<tr>
<td>rs2268491</td>
<td>.46</td>
<td>.90</td>
<td>.44</td>
<td>.15</td>
</tr>
<tr>
<td>rs237889</td>
<td>.36</td>
<td>.41</td>
<td>.43</td>
<td>.68</td>
</tr>
<tr>
<td>rs11131149</td>
<td>.65</td>
<td>.65</td>
<td>.80</td>
<td>.21</td>
</tr>
<tr>
<td>rs2268495</td>
<td>.14</td>
<td>.03</td>
<td>.60</td>
<td>.99</td>
</tr>
<tr>
<td>rs237897</td>
<td>.35</td>
<td>.29</td>
<td>.38</td>
<td>.90</td>
</tr>
<tr>
<td>rs237899</td>
<td>.98</td>
<td>.31</td>
<td>.84</td>
<td>.86</td>
</tr>
<tr>
<td>rs2301261</td>
<td>.61</td>
<td>.95</td>
<td>.77</td>
<td>.08</td>
</tr>
</tbody>
</table>

Note: For the Cambridge Face Memory Test (Duchaine & Nakayama, 2006), the Mooney Face Test (Mooney, 1957), and the Glasgow Face Matching Test (Burton, White, & McNeill, 2010), performance data were entered into the analysis as ranks, and ties were assigned the average rank of the tied values. For the Composite Face Test (Richler, Cheung, & Gauthier, 2011), holistic index was entered into the analysis unranked. For the AQ (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), scores were entered into the analysis unranked. All SNPs listed had a minor-allele frequency of .05 or greater in our sample. The boldface type highlights the results for the SNP rs237887, for which Skuse et al. (2014) found a significant association with face recognition (see the main text).

*This association between rs2324728 and ranked performance on the Mooney Face Test had the lowest uncorrected *p* value of all the associations of the performance measures with the 17 genotyped SNPs of *OXTR*. It did not survive a conventional Bonferroni correction (.004 × 17 SNPs = .068) or even a more moderate effective correction that took genetic linkage into account (.004 × 13.19 effective corrections = .053). The number of effective corrections was determined using the Genetic Type 1 Error Calculator (http://grass.cgs.hku.hk/gec/; Li, Yeung, Cherny, & Sham, 2012).
One SNP of OXTR that repeatedly has been identified with prosocial behaviors is rs53576 (Bakermans-Kranenburg & van IJzendoorn, 2014; Ebstein et al., 2012; Kogan et al., 2011). This SNP was one of the 75 imputed SNPs that we tested, and its imputation quality was high (squared correlation between proximal imputed and genotyped SNPs, or RSQR = .84). However, it was not significantly associated with any of our measures of face processing. For the association with performance on the Cambridge Face Memory Test in particular, the uncorrected $p$ value was .32.

The association between OXTR and face recognition observed by Skuse et al. (2014) was largely driven by the parents and siblings in their sample, none of whom “had significant autistic traits” (p. 1988), as measured by the AQ. Skuse et al. concluded that all of the parents and siblings “could be considered neurotypical in that respect” (p. 1988). We can confirm that our sample, too, was neurotypical in this respect. Of the participants who completed our tests of face processing, 316 (203 women) had also completed the AQ questionnaire earlier (Baron-Cohen et al., 2001): Their mean score was 17.85 ($SD = 7.94$, range $= 3–39$); the maximum possible score is 50, and a score of 32 or higher is suggestive of autism spectrum disorder (only 21 participants in the current study reached this score).

There are several reports of an association between OXTR and autism spectrum disorder (Di Napoli et al., 2014; LoParo & Waldman, 2015). We therefore checked whether there was a link between AQ score and polymorphisms in OXTR in the present population of young adults. For this analysis, we used all 521 participants (333 women) who had served in the original cohort and who had completed the AQ questionnaire ($M = 17.32$, $SD = 7.58$, range $= 3–39$; 25 participants scored at or above 32).

As in all other genetic analyses in this study, we entered sex as a covariate, especially because we observed a significant sex difference in AQ score (Mann-Whitney $U = 25,252$, $p = .00024$; mean AQ score for women = 16.48, mean AQ score for men = 18.82).

We found no significant association between AQ score and rs237887 or any other genotyped SNPs (see Table 1; the lowest uncorrected $p$ value was .17). Moreover, none of the imputed SNPs, including rs53576, was significantly associated with AQ score (see Fig. 2; the lowest uncorrected $p$ value was .07). Even when we followed the
method of Rhodes, Jeffery, Taylor, and Ewing (2013) in calculating a total AQ score—that is, totaling the raw scores of all items rather than labeling responses to items in a binary fashion, which is the usual approach—we observed no significant associations (uncorrected \( p \) values ranged from .43 to .99). For genetic correlates of AQ score, we had 99% power (at an \( \alpha \) level of .05) to detect associations with an effect size as small as \( r^2 = .05 \) (74% power for \( r^2 > .01 \); 49% power for \( r^2 > .005 \)). Any association of OXTR with autism spectrum disorder may be confined to people who explicitly exhibit the condition.

**Discussion**

Our primary findings are at odds with those of Skuse et al. (2014), who reported a significant association between rs237887 and recognition memory for faces. Despite our large sample size \( (n = 370) \) and ample statistical power, we observed no association between rs237887 and the Cambridge Face Memory Test, which is widely accepted as a pure measure of face recognition ability. Nor did we find an association between rs237887 and the Mooney Face Test or the Glasgow Face Matching Test. A marginal association between rs237887 and raw score on the Composite Face Test was not significant after correction for multiple testing.

What, then, might be the critical difference between our study and that of Skuse et al. (2014)? Could it be the difference in population? Skuse et al. investigated children with autism as well as the parents and siblings of those children. However, the association they observed with rs237887 was weakest for the autistic probands, whereas it was stronger for the parents and siblings; the association became significant (after Bonferroni correction) only for the combined sample, two thirds of which were nominally healthy participants. Skuse et al. mentioned that they specifically selected a sample of autistic probands and their immediate family members “to maximize the range of social cognitive abilities under investigation” (p. 1991). Although we ourselves did not select a restricted sample, our results show a very wide range of face recognition ability (see Results). However, it remains possible that rs237887 is associated with face recognition within the special population—relatives of children with autism—that was studied by Skuse et al.

The fact that we tested our participants online instead of in the lab should not have substantially influenced the reliability of our results: Germine et al. (2012) found that online testing \( (n = 4,080) \) yielded high-quality data that were as reliable as data gathered from three lab-based samples (combined \( N = 327 \)). At the request of a reviewer, we reran our genetic association, omitting the 10 out of
370 (2.7%) participants whose scores were 2 or more standard deviations below the mean score on the Cambridge Face Memory Test. The results were little changed. As for a possible difference in age range between our sample and that of Skuse et al. (2014), it is not possible to make a comparison, because only a subset of their total sample completed the Warrington test, and the age range for that subset was not reported.

What is possible is that the variance accounted for by rs237887 in Skuse et al. (2014) is not specific to face recognition, but rather reflects some other ability required for performance on the Warrington test. One candidate might be general intelligence, because Skuse et al. observed a significant association between performance on the Warrington test and IQ ($r = .30$), whereas, in contrast, the Cambridge Face Memory Test exhibits little or no correlation with general intelligence (Davis et al., 2011; Shakeshaft & Plomin, 2015; Wilmer et al., 2014). In addition, Skuse et al. did not use the raw performance data from the Warrington test; they first standardized the performance for age. (Oddly, they used performance data from a test of affect recognition—the Ekman and Friesen, 1976, Pictures of Facial Affect—to do so. The extent to which their observed association of OXTR with face recognition might be the effect of this unconventional standardization method is unclear.) When we repeated our genetic analysis using age as a covariate, the uncorrected $p$ values for rs237887 were little changed from those in Table 1.

Our negative finding for rs237887 is limited, of course, to the recognition of facial identity. It may well be the case that polymorphisms of OXTR are associated with individual differences in the ability to infer emotional states from facial expressions—individual differences that may derive from, or contribute to, prosocial behaviors and to social anxieties.

In sum, the conclusion of Skuse et al. (2014) may still apply to a particular test, or to a particular population of relatives of children with autism, and it remains possible that polymorphisms of OXTR are related to individual differences in empathy or social anxiety. What we can firmly conclude, however, is that in a population of healthy young adults, there was no strong association between SNP rs237887 of OXTR and the ability to recognize previously seen faces. This is the conclusion that we emphasize.

**Action Editor**

Alice J. O’Toole served as action editor for this article.

**Author Contributions**

All the authors developed the study concept. All the authors contributed to the study design and performed testing and data collection. R. J. Verhallen performed the data analysis and interpretation under the supervision of J. D. Mollon. R. J. Verhallen drafted the manuscript, and J. M. Bosten, P. T. Goodbourn, and J. D. Mollon provided critical revisions. All the authors approved the final version of the manuscript for submission.

**Acknowledgments**

We thank Mike Burton and his colleagues for making available the Glasgow Face Matching Test, Brad Duchaine and his colleagues for making available the Cambridge Face Memory Test, and Jennifer Richler and her colleagues for making available their version of the Compositional Face Test. We are grateful to Emily Clemente, Julien Bauer, and Kerry Cliffe of Cambridge Genomic Services for their valuable help. We thank Richard Bethlehem for helpful conversations.

**Declaration of Conflicting Interests**

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

**Funding**

This work was supported by Gatsby Charitable Foundation Grant GAT2903.

**References**


Verhallen et al.


