Mind-reading receptors: The role of the vasopressin receptor 1a gene (*AVPR1A*) in human empathy, social interaction and health

Alyssa Rhiannon Morse

June 2019

A thesis submitted for the degree of

Doctor of Philosophy of The Australian National University.

This page left intentionally blank.

Student Statement

This statement certifies that the following work entitled *Mind-reading receptors: The role of the vasopressin receptor 1a gene (*AVPR1A*) in human empathy, social interaction and health* is the author's own original work, except where otherwise specified, complies with The Australian National University Award Rules and has not previously been accepted for award of a degree or diploma to any other university or institution of higher learning. This thesis was supported by an Australian Government Research Training Program (AGRTP) Scholarship.

According to the Procedure: Higher degree research – submission and examination of theses, the word count of this thesis without appendices is 43733.

Student signature:..... Date: 11/06/2019

Acknowledgements

The literature I've read over the years of this project has taught me about the importance of social support for health and wellbeing. My experience working on this project has certainly taught me the personal value of social support and of my social networks – it takes a village to raise a PhD Scholar!

To my supervisory panel, thank you for your guidance and advice over the course of my candidature, I am grateful for all I have learned. To my Primary Supervisor, Simon Easteal, it has been a long journey to get here since my candidature commenced in 2012. Thank you for staying the course with me, through all its twists and turns. To Richard Burns, thank you for joining my panel in the later years of my candidature. Your support and the knowledge you imparted have been invaluable. I am grateful for the friendly working environment created by the staff and students at the John Curtin School of Medical Research who have shared parts of this journey with me. Particular thanks to Susan Tan, for always being happy to talk through a challenge, exchange knowledge, or just stop for a chat.

To my wonderful colleagues at the Centre for Mental Health Research, since I joined the team in 2016 you have been a constant source of encouragement and inspiration. Working at CMHR has shown me how much I have learned through my PhD candidature and supported me to continue to grow my skills and confidence. A special thank you to Michelle Banfield, for that first opportunity you gave to me and all your support since then.

To my friends, thank you for reminding me to engage in social interaction, not just study it. To my dear friend Liz Huxley, I will always be grateful for the support, understanding and much needed coffee breaks and crafternoons. Music has been another important part of this journey; this thesis was written to an instrumental soundtrack ranging from JS Bach to Dexter Gordon to Darren Korb. To my singing teacher, Dennis Brain, thank

you for encouraging me to continue pursuing my passion for music, it has been such an important creative outlet for me.

Last, but certainly not least, I am so very grateful to my amazing family. Mum and Dad, I would not have made it here without your constant support, thank you for your patience and for creating the nurturing space I needed to complete this project. To my brothers, Callum and Elliott, thank you for taking my mind off work when things were tough and for always knowing how to make me smile. To the newest (and furriest) members of our family, Lily and Jasper, thank you for the schnauzer hugs, listening ears and reminders to get outside and take a walk. And to Nana and Papa Wilson, thank you for your love and prayers. Papa, although you couldn't be here to see the end of this journey, I have kept you with me. Your fountain pen has helped me draft many of my chapters.

Thank you to the village of people who have supported me throughout this process. I can't wait to see what comes next.

Abstract

Social support is beneficial for human mental and physical health, whereas experiencing social strain can be detrimental. The impact of social interaction varies between individuals and this variation may partly be driven by biological factors. Evidence suggests that variation in the 5' promoter region of the vasopressin 1a receptor gene (*AVPR1A)* influences individual differences in human social behaviours, social cognition, and how a person responds to their social environment. The research described in this thesis aimed to determine whether variation in the short tandem repeat polymorphism *AVPR1A RS3* was directly associated with individual differences in mental and physical health, and whether *AVPR1A RS3* genotype moderated the impact of social support and social strain on health. The project also aimed to explore the relationship between *AVPR1A RS3* variation and empathy, a trait that could provide a link between *AVPR1A* and health outcomes.

Relationships between genetic variation and social and health phenotypes were explored. Two series of nested multi-level linear models found a significant association between *AVPR1A RS3* genotype and women's self-reported physical and mental health. Results indicate that women carrying one or two copies of the Short *RS3* allele report better mental and physical health than women homozygous for the Long *RS3* allele. There was some evidence of an interaction between genotype and social support and strain variables, but the overall pattern of interaction effects was unclear. No direct relationship between *AVPR1A* and health was found for men. Due to the limitations of the data set used, a population-based analysis and a twin-modelling analysis could not draw strong conclusions about the relationship between *AVPR1A RS3* genotype and individual differences in empathy. However, within the context of existing literature, the results of this project support the hypothesis that genetic variation at the *RS3* locus is associated with individual differences in tendency to empathise and cognitive empathy skill.

In this project, empathy was measured using three short-form versions of two common empathy measures, the Empathy Quotient and the Reading the Eyes in the Mind Test. The psychometric properties of these measures were explored to determine if they were appropriate for use in future research. A short-form of the Empathy Quotient, the EQ-17, was judged to be appropriate for future use. Two short-forms of the Reading the Mind in the Eyes test, the RMET-17 and RMET-14, were found to have similar psychometric properties to the full-length test, however, results indicated that they were inappropriate for future research use.

The results of this project align with previous research linking *AVPR1A RS3* variation with social traits and add new evidence indicating that this genetic variant is also associated with individual differences in women's mental and physical health. The relationship between *AVPR1A RS3* genotype and both mental and physical health emphasises the importance of continuing to improve our understanding of the vasopressin system in human social behaviour.

Table of Contents

Figures and Tables

List of Figures

List of Tables

Chapter 1: Introduction

Introduction

Social interaction is an essential part of human society, permeating almost every domain of everyday life, from the home to the classroom or workplace. The importance of social interaction is demonstrated by the well-established relationships between social support, social strain, and physical and mental health (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016; Yang, Schorpp, & Harris, 2014). The impact of both social support and strain varies between individuals (Ditzen & Heinrichs, 2014). Understanding the biological basis of this variation could help elucidate the mechanisms underlying the influence of social support and strain on health. In future, such knowledge could be applied to tailoring psychosocial interventions to individual needs.

There has been a rapidly growing interest in the role of the neuropeptide vasopressin in human social behaviour, and its potential therapeutic applications for mental disorders involving social deficits (Bartz & Hollander, 2008; Finger, 2011; McGregor & Bowen, 2012; Modi & Young, 2012; Striepens, Kendrick, Maier, & Hurlemann, 2011). Vasopressin and the vasopressin 1a (V1a) receptor underlie inter- and intra-species differences in a variety of mammalian social behaviours. Vasopressin and the genes that encodes it are highly conserved across species (Donaldson & Young, 2008). Interspecies differences in the effects of this neuropeptide are driven by variation in the neural distribution of V1a cell-surface receptors (Albers, 2012; Caldwell, Lee, Macbeth, & Young Iii, 2008; Donaldson & Young, 2008; McCall & Singer, 2012). Genetic variation in the 5' promoter region of the gene encoding the V1a receptor (*AVPR1A*) may influence the impact of the receptor on human behaviour (Hammock & Young, 2005).

Evidence suggests that variation in the 5' promoter region of the *AVPR1A* gene influences individual differences in social behaviours, social cognition and in how a person responds to their social environment (Bisceglia et al., 2012; Knafo, Israel, et al., 2008; Levin

et al., 2009; Uzefovsky et al., 2015). Thus, determining the role of *AVPR1A* in human sociality could improve our understanding of the biological mechanisms underlying the connection between social interaction and physical and mental health. Social support is beneficial for physical and mental health, whereas experiencing social strain can be detrimental (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016; Yang et al., 2014). The impact of social interaction on health varies between individuals and some of this variation may be driven by biological factors (Ditzen & Heinrichs, 2014).

The project described in this thesis aimed to determine whether variation in the *AVPR1A* gene was directly associated with individual differences in mental and physical health, and whether *AVPR1A* genotype moderated the impact of social support and social strain on health. The project also aimed to explore the relationship between *AVPR1A* variation and empathy, a trait that could provide a link between *AVPR1A* and health outcomes. Aims and hypotheses were drawn from a critical review of literature presenting convergent evidence for the role of vasopressin and the *AVPR1A* gene in human social behaviour and other social traits.

The investigation was conducted using a candidate gene analysis approach. The genetic variant of interest, *AVPR1A RS3*, was selected based on evidence for the biologically functional consequences of short tandem repeat variation at this locus (Hammock & Young, 2002; Hammock & Young, 2005; Knafo, Israel, et al., 2008; Tansey et al., 2011). Evidence suggests that the repeat length of the *RS3* microsatellite has functional consequences for gene expression, which are subsequently associated neuroanatomical differences in receptor distribution, and differences in social behaviour (Hammock & Young, 2005; Knafo, Israel, et al., 2008; Tansey et al., 2011).

A literature review, and the research questions and aims drawn from it are presented in Chapter 2 alongside an explanation of the candidate gene approach. Chapter 3 describes the

first study of the project which determined whether *AVPR1A RS3* genotype was associated with individual differences in mental and physical health in a large population-based longitudinal sample using multi-level linear modelling techniques. The study also investigated whether genotype moderates the relationship between positive social interaction and health, and/or the relationship between negative social interaction and health. The second study, described in Chapter 4, explored the relationship between *AVPR1A RS3* genotype and individual differences in empathy, a trait hypothesised to provide a link between *AVPR1A* variation and health outcomes. This relationship was tested in a population-based regression modelling analysis and a twin-modelling analysis. In Chapter 4, empathy was measured using three new short-forms of two common empathy measures, the Empathy Quotient (Baron-Cohen & Wheelwright, 2004) and the Reading the Mind in the Eyes test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). These short-form scales had not been psychometrically assessed. Chapter 5 describes the psychometric properties of the new short-form measures, comparing them to their full-length equivalents. The aim of this investigation was to determine if the new measures were appropriate for use in future research, including future candidate gene research. The thesis concludes in Chapter 6 with a general discussion of the results of the three projects, the collective contribution they make to current literature, and a consideration of future directions for research.

The results of this project indicate that *AVPR1A RS3* genotype is directly associated with individual differences in women's self-reported mental and physical health and suggest that *AVPR1A RS3* genotype may moderate the relationship between positive and negative social interaction and health for both sexes. Strong conclusions about the relationship between *AVPR1A* and empathy could not be drawn but the limitations of the project suggest recommendations for future research in this area. The relationship between *AVPR1A RS3*

genotype and both mental and physical health emphasises the importance of continuing to improve our understanding of the vasopressin system in human social behaviour.

Chapter 2: Literature Review and Aims

Chapter Summary

Social support, particularly perceived social support, is beneficial for our mental and physical health. Conversely, social strain is detrimental to mental and physical health (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016; Seeman, Gruenewald, Cohen, Williams, & Matthews, 2014). The impact of social support and strain varies between individuals, and this variation may be driven by biological factors (Ditzen & Heinrichs, 2014). One potential driving factor is the neuropeptide vasopressin. Vasopressin is implicated in a range of social behaviours and cognitive processes in humans and non-human mammals (Albers, 2012; Caldwell et al., 2008; McCall & Singer, 2012). Vasopressin and the gene that encodes it are highly conserved across species, however, the distribution of its neural receptors varies both within and between species (Donaldson & Young, 2008; Hammock & Young, 2002; Insel & Shapiro, 1992; Insel, Wang, & Ferris, 1994). Variation in V1a receptor distribution, driven by variation in the 5' promoter region of the *AVPR1A* gene, is thought to underlie differences in vasopressin-related parental, pair-bonding, and other social behaviours between species, and between individuals within a species (Hammock & Young, 2005).

Intranasal administration of exogenous vasopressin can influence human responses to facial emotional expressions, social interaction, and social stressors (Rilling et al., 2012; Shalev et al., 2011; Thompson, George, Walton, Orr, & Benson, 2006). Variants in the *AVPR1A* gene are also associated with behaviours and traits that influence how people respond to their social environment, including empathy and empathy-related traits (Bisceglia et al., 2012; Knafo, Israel, et al., 2008; Uzefovsky et al., 2015). By influencing how people respond to social support and social strain, variation in the *AVPR1A* gene may modulate the health effects of social interaction. *AVPR1A* may also indirectly impact on mental and physical health by influencing the quality of close social relationships, including spousal relationships and mother-infant interactions (Bisceglia et al., 2012; Walum et al., 2008). The

project described in this thesis aimed to determine whether variation in the *AVPR1A* gene was directly associated with individual differences in mental and physical health, and whether *AVPR1A* genotype moderated the impact of social support and social strain on health. The project also aimed to further explore the relationship between *AVPR1A* variation and empathy, a trait that could function as a link between *AVPR1A* and health outcomes.

Social Support and Wellbeing

Social support is an important component of health and wellbeing. The practical and emotional support a person receives from their friends, family members, spouse, and other social relationships is beneficial for their mental and physical health (Caltabiano, Sarafino, & Byrne, 2008). Social support can be measured as the number of close social connections a person has (social network or social integration) or as the amount of support a person perceives themselves as receiving. Perceived social support is particularly important in relation to health and wellbeing; people who perceive themselves as more socially integrated are happier and healthier than people who feel lonely or isolated (Ditzen & Heinrichs, 2014). The association between support and health has been well-established (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016). Historically, epidemiological studies have found that people with at least one close confidant have a higher quality of life and widowed persons are found to have increased death rates (Ditzen & Heinrichs, 2014). More recently, a meta-analysis found that people with stronger social relationships have a 50% increased likelihood of survival (Holt-Lunstad, Smith, & Layton, 2010). The effect of social relationships on mortality was comparable to other well-established risk factors and the results were consistent across age, sex, initial health status, follow-up period and cause of death (Holt-Lunstad et al., 2010). Higher levels of perceived social support are also associated with higher life satisfaction, better reported mental health and reduced risk of developing mental illnesses such as major depressive disorder (Lee & Szinovacz, 2016; Santini et al., 2016; Schuster, Kessler, & Aseltine, 1990; Teo, Choi, & Valenstein, 2013; Wong, Wu, Gregorich, & Perez-Stable, 2014).

Conversely, social strain can have negative effects on mental and physical health and is often found to have a stronger effect on physical health than social support (Seeman et al., 2014; Yang et al., 2014). Social strain encompasses the negative interactions within

relationships, including conflict, tension, criticism, irritation and disappointment (Lee & Szinovacz, 2016; Schuster et al., 1990). Social strain has been associated with increases in inflammation, inflammatory responses (Miller, Rohleder, & Cole, 2009; Seeman et al., 2014; Yang et al., 2014) and allostatic load, a multi-system measure of biological risk (Seeman et al., 2014). For example, chronic interpersonal stress has been associated with increased expression of pro- and anti-inflammatory signalling molecules and an increased immune response to microbial challenge in young women (Miller et al., 2009). High levels of social strain are also associated with increased levels of psychological distress (Schuster et al., 1990), lower reported mental health (Lyu & Agrigoroaei, 2017), higher levels of depressive symptoms (Lee & Szinovacz, 2016; Santini et al., 2016) and a higher risk of developing a major depressive disorder (Teo et al., 2013). The strength of these associations varies depending on the source of strain and characteristics of the individual (Lee & Szinovacz, 2016; Lyu & Agrigoroaei, 2017; Santini et al., 2016; Schuster et al., 1990; Teo et al., 2013).

The impact of both social support and strain varies between individuals (Ditzen & Heinrichs, 2014). Understanding the basis of this variation could help elucidate the mechanisms underlying the influence of social support and strain on health. Such knowledge could be applied to tailoring psychosocial interventions to individual needs. There is evidence to suggest that individual differences in the response to social interaction may be partly driven by biological factors (Ditzen & Heinrichs, 2014). A promising stream of research in this area has focused on the role of two neuropeptides, oxytocin and vasopressin, in social behaviour. These neuropeptides have been associated with a variety of complex social behaviours in non-human mammals such as pair-bonding, maternal behaviour, aggression, social recognition and social communication (Albers, 2012; Caldwell et al., 2008; McCall & Singer, 2012).

Social Neuropeptides: Oxytocin and Vasopressin

There has been a rapidly growing interest in the role of oxytocin and vasopressin in human social behaviour, and their potential therapeutic applications. These neuropeptides have been have been implicated in a range of social behaviours and cognitive processes, from facial emotion perception and social learning (Hurlemann et al., 2010; Thompson et al., 2006) to parenting behaviour and altruism (Avinun, Ebstein, & Knafo, 2012; Israel et al., 2009; Knafo, Israel, et al., 2008). Oxytocin, vasopressin, and the genes that encode them are highly conserved across species (Donaldson & Young, 2008), however, the distribution of their receptors in the brain varies widely both between and within species (Hammock & Young, 2002; Insel & Shapiro, 1992; Insel et al., 1994). This variation in receptor distribution is thought to underlie the variation in parental, pair-bonding and other social behaviours between species, as well as the individual variation in these behaviours within a species. Except for post-mortem studies, it is not possible to directly study the distribution of oxytocin and vasopressin receptors in the human brain. However, we can gain insight into the effects of receptor distribution on human behaviour by studying genetic variation in the oxytocin and vasopressin receptor genes and their relationship with social phenotypes.

Oxytocin and vasopressin are nonapeptides, a family of neuropeptides characterised by their nine amino acid structure, and differ from each other at only two amino acid positions (Donaldson & Young, 2008; Insel, 2010). In mammals, they are produced in the hypothalamus and are released from the posterior pituitary into the periphery where they are involved in a diverse range of functions. Oxytocin- and vasopressin-expressing neurons in the hypothalamus also project centrally and the neuropeptides affect various brain regions where their receptors are present (Donaldson & Young, 2008; Insel, 2010). The genetic structure and expression of oxytocin and vasopressin have been remarkably well-conserved and these neuropeptides play a general role in social and reproductive behaviour across a wide

range of organisms (Donaldson & Young, 2008). They also appear to have sex specific roles in social behaviour, with oxytocin playing a larger role in female behaviour and vasopressin in male behaviour (Albers, 2012; Dumais & Veenema, 2016; McCall & Singer, 2012).

The evolution of these neuropeptide systems is primarily mediated by changes in the oxytocin and vasopressin receptors (Pare et al., 2016). Oxytocin has one receptor, OXTR, which is expressed both in the periphery and the brain (Inoue et al., 1994). Vasopressin has three receptors, two of which, V1a and V1b, are expressed in the brain (Insel, 2010). The neural distribution of these receptors varies widely between species and it is this variation that drives interspecies differences in neuropeptide modulated social behaviours (Hammock & Young, 2002; Insel & Shapiro, 1992; Insel et al., 1994). Social behaviour can also be modulated by variation in oxytocin and vasopressin receptor distributions within a species (Hammock & Young, 2005) and even by changes in receptor distribution in the same individual over time (Insel & Shapiro, 1992). Thus, the neural distribution of oxytocin and vasopressin receptors drives differences in mammalian social behaviour.

Except for post-mortem studies (Knafo, Israel, et al., 2008; Loup, Tribollet, Dubois-Dauphin, & Dreifuss, 1991) we cannot directly observe oxytocin and vasopressin receptor distributions in the human brain. Thus, to determine how oxytocin and vasopressin receptors affect individual differences in human social behaviour we need an indirect measure that allows us to approximate receptor distribution. One way to achieve this is to use variation in the genes that encode these receptors as an indication of receptor expression and/or function. Variation in the prairie vole (*Microtus ochrogaster*) V1a receptor gene has been shown to result in different gene expression levels, receptor distributions in the brain and social behaviour profiles (Hammock & Young, 2005). Currently, similar evidence of function does not exist for variation of the oxytocin receptor gene.

Much research has focused on the role of oxytocin and the oxytocin receptor gene (*OXTR*) in human social behaviour (Ebstein, Knafo, Mankuta, Chew, & Lai, 2012; Zink & Meyer-Lindenberg, 2012). While there is substantial interest in the potential clinical applications of the oxytocin neuropeptide (Bartz & Hollander, 2008; Finger, 2011; McGregor & Bowen, 2012; Modi & Young, 2012; Striepens et al., 2011), a recent meta-analysis suggests that the two most commonly investigated *OXTR* single nucleotide polymorphisms (SNPs) are not associated with individual differences in human behaviour (Bakermans-Kranenburg & van IJzendoorn, 2014). Comparatively little research has focused on the role of vasopressin and the vasopressin receptor genes, despite the close relationship of the two neuropeptide systems (Pare et al., 2016). Existing research suggests that vasopressin plays an important role in social behaviour (Albers, 2012; Ebstein et al., 2012) and there is evidence for the functional role of genetic variation in the promoter region of the gene encoding the V1a receptor, *AVPR1A* (Hammock & Young, 2005; Tansey et al., 2011). Additionally, vasopressin can modulate human responses to social cues and social stress (Gozzi, Dashow, Thurm, Swedo, & Zink; Shalev et al., 2011; Thompson et al., 2006). Thus, vasopressin and the *AVPR1A* gene may form part of the biological link between social interaction and health outcomes. This thesis describes an investigation of the role of variation in the *AVPR1A* gene in human social interaction and health.

Before discussing the current evidence for the role of the *AVPR1A* gene in human social behaviour, it is important to understand the methods commonly used to explore this association. Associations between a single gene, or small group of genes, and a complex phenotypic trait are often studied using candidate gene analysis techniques. The next section will describe candidate gene analysis and discuss why this is the most appropriate technique to explore the association between *AVPR1A* variation and complex human traits.

Candidate Gene Analysis

The candidate gene approach is a valuable, effective and economical tool for studying the inheritance and development of complex traits (Parens, Chapman, & Press, 2006; Zhu & Zhao, 2007). A complex, or quantitative, trait is a phenotype that does not segregate in simple Mendelian fashion (the genotype cannot be directly inferred from the phenotype). Instead, it is influenced by multiple genes and environmental factors (Griffiths, Wessler, Carroll, & Doebley, 2012). The trait must display measurable variation that can be either continuous or discrete (Members of the Complex Trait Consortium, 2003). Most traits and common diseases of interest to psychology, evolutionary biology and medicine show complex patterns of inheritance (Griffiths et al., 2012; Parens et al., 2006).

A candidate gene is a gene with a known biological function that, on the basis of existing evidence, is hypothesised to directly or indirectly influence the trait of interest (Zhu & Zhao, 2007). The candidate gene approach examines genetic influences on a complex trait by (1) identifying genes that may influence the complex trait, (2) developing *a priori* hypotheses about the effect of the selected gene(s) on the trait, (3) identifying variants in or near the gene(s) that are likely to cause functional changes in protein structure or gene expression, (4) genotyping the selected variant(s) in a sample, and (5) using statistical methods to determine whether the genetic variant is associated with the phenotype (Tabor, Risch, & Myers, 2002). Amassing evidence about associations between a gene and related phenotypes, or a phenotype and related genetic variants, can build understanding about the biological mechanisms underlying a trait (Tabor et al., 2002).

The candidate gene approach is particularly well-suited to detecting small-to-moderate effects of a single gene or small group of genes (Tabor et al., 2002) and can be applied to examine potential gene-environment interactions (Parens et al., 2006). An important advantage of the candidate gene approach is that a range of molecular markers can be used as

candidates. Another common method, the genome wide association (GWA) study, detects associations between a phenotype and single-nucleotide differences in a genetic sequence called single nucleotide polymorphisms (SNPs; (Griffiths et al., 2012)). GWA studies are usually only able to explain a small portion of the heritability of a trait or disease (Griffiths et al., 2012; Hannan, 2009). The "missing heritability" is thought to be partly explained by non-SNP polymorphisms in the genome, including tandem repeat polymorphisms, which GWA studies may be unable to detect (Bakhtiari, Shleizer-Burko, Gymrek, Bansal, & Bafna, 2018; Griffiths et al., 2012; Hannan, 2009). The candidate gene approach can be used to explore associations between quantitative traits and tandem repeat polymorphisms, which may otherwise be missed in GWA studies.

Tandem repeat polymorphisms are an important source of variation in the human genome. A tandem repeat polymorphism is a repetitive sequence of DNA for which the number of repeats varies between individuals (Griffiths et al., 2012). They are primarily located in introns and intergenic regions but are also located in exons (Hannan, 2009). This repeat number variation is a potential mechanism for continuous phenotypic variation, allowing for a large number of alleles at a single locus, where SNP variation only allows for two alleles (Hammock & Young, 2005). Evidence indicates that tandem repeat polymorphisms affect a wide range of biological processes and can modulate gene expression and the structure and function of RNAs and proteins (Hannan, 2009). The role of these polymorphisms in gene expression is thought to underlie their contribution to quantitative traits. The candidate gene approach can be used to explore the role of tandem repeat variation in determining those traits.

The candidate gene approach is limited by its reliance on previous evidence of gene function, variant function and the underlying biology of the trait of interest (Tabor et al., 2002). This evidence is required to select candidate genes and genetic variants that could

have a biologically plausible link to the phenotype of interest. In the case of *AVPR1A,* evidence supports the biological function of commonly studied genetic variants. Vasopressin and its receptors are implicated in biological pathways associated with social behaviour, and there is accumulating evidence linking variation in the *AVPR1A* gene with a variety of human social behaviours. The following section provides an overview of the evidence for the biological function of *AVPR1A* variants.

Biological Function of *AVPR1A* Variation

Two vasopressin receptors are expressed in the brain, V1a and V1b (Insel, 2010). V1a has been most strongly implicated in mammalian social behaviour (Albers, 2012; Caldwell et al., 2008) and is focused on in this project. V1a is encoded by the *AVPR1A* gene which has two exons separated by a 2.2 kilobase (kb) intron (Thibonnier et al., 1996). The human genome contains a single copy of *AVPR1A* located on chromosome 12q14-q15 (Thibonnier et al., 1996). Microsatellite variation (a type of tandem repeat polymorphism) is most commonly studied in this gene. Four microsatellite motifs were identified in *AVPR1A* by Thibonnier and colleagues (2000) two of which are commonly studied in relation to human social behaviour; *RS1* and *RS3* (Ebstein et al., 2012). These tandem repeat polymorphisms are located in the 5' promoter region of *AVPR1A*. Both are associated with human social behaviour (Ebstein et al., 2012) and there is evidence to suggest that promoter region microsatellite repeat length has functional consequences for gene expression (Hammock & Young, 2005; Knafo, Israel, et al., 2008; Tansey et al., 2011).

Evidence from Animal Studies

Genetic variation in the *AVPR1A* gene is thought to influence behaviour by determining the distribution of V1a receptors in the brain. The classic example of this

association are the observed differences in *Avpr1a* genotype, V1a receptor distribution and pair-bonding behaviour between the socially monogamous prairie vole (*Microtus ochrogaster*) and the promiscuous montane vole (*Microtus montanus)*. Prairie voles are highly social and monogamous, whereas the montane vole is relatively asocial and non-monogamous (Hammock & Young, 2002). These species display similar levels of neural vasopressin expression and distribution but display different neural distributions of the V1a receptor. For example, prairie voles have a high density of V1a expression in the ventral palladium, as do several other socially monogamous species such as the pine vole, California mouse and common marmoset (Young, Nilsen, Waymire, MacGregor, & Insel, 1999). In contrast, montane voles have a lower density of V1a expression in this region, as do other non-monogamous species like the meadow vole (Young et al., 1999). In male prairie voles, intracerebroventricular injection of vasopressin results in an increase in affiliative behaviour towards a stimulus female, this change in behaviour does not occur when vasopressin is administered to montane voles (Young et al., 1999). Thus, evidence suggests that vasopressin acts through the V1a receptors to promote affiliative behaviour in prairie voles, but not in montane voles. It is the expression pattern of the V1a receptor that drives these differences in social behaviour (Hammock & Young, 2002).

Interspecies and intraspecies genetic variation. Differences in receptor expression appear to be at least partly driven by variation in the promoter region of the *Avpr1a* gene. Most of the coding and surrounding non-coding sequences of this gene are highly conserved, however, there is an approximately 400 base pair (bp) sequence rich in microsatellite DNA that is present in the prairie and pine vole *AVPR1A* promoter region, but not in the montane or meadow vole region (Hammock & Young, 2002; Young et al., 1999). This region may be an expansion of a short imperfect repetitive sequence present in the montane vole gene and contains a series of tandem repeat sequences, interspersed with non-repetitive sequences

(Hammock & Young, 2002). While the montane vole shows little inter-individual variation in the *Avpr1a* promoter region, the expanded prairie vole region allows for inter-individual genotype variation and alleles have been found to differ by up to 140bp (Hammock & Young, 2002). Similarly, montane voles display little individual variation in V1a binding in the brain whereas prairie voles display individual differences in both neural receptor binding and behaviour (Hammock & Young, 2005). While intraspecific differences are smaller than interspecific differences, intraspecific prairie vole genetic variation is sufficient to modify gene expression. Two prairie vole *Avpr1a* regulatory region alleles differing by 19bp were compared for their ability to drive the luciferase reporter in A7r5 cells (Hammock & Young, 2005). The long allele displayed significantly increased luciferase activity compared to the short allele in three independent cell culture experiments (Hammock & Young, 2005).

Gene-driven differences in receptor distribution. Genetic variation in the *Avpr1a* gene is associated with neuroanatomical differences in receptor distribution. Two transgenic mouse studies have demonstrated the ability of the promoter region of the *Avpr1a* gene to alter V1a receptor expression. Compared to non-transgenic mice, mice transgenic for the prairie vole *Avrp1a* gene display a V1a receptor distribution that is more similar to the prairie vole distribution (Young et al., 1999). A more recent study examined the impact of replacing 3.4kb of the 5' flanking region of the mouse *Avrpr1a* gene with prairie vole and meadow vole homologues (Donaldson & Young, 2013). Meadow vole, prairie vole long, and prairie vole short variants conferred differences in receptor binding in the thalamus, amygdala and dentate gyrus. Variation in the proximal 5' flanking region was not found to be sufficient to confer species-specific patterns of receptor binding in the transgenic mice but it was sufficient to change expression levels in a direction consistent with species-specific binding patterns (Donaldson & Young, 2013). The genetic mechanisms that influence diversity in brain receptor patters are likely to be complex and the effect of *Avpr1a* variation may be cell-type

dependent, allowing for differential expression across brain regions, rather than global differences in expression (Donaldson & Young, 2013; Hammock & Young, 2005). Overall, this evidence suggests that changes in the promoter region of *Avpr1a* are capable of at least moderately driving differences in V1a receptor expression.

Gene associated differences in social behaviour. Inter-individual variation in *Avpr1a* genotype also correlates with differences in prairie vole social and reproductive behaviour. Shorter promoter region alleles are associated with higher pup mortality in first-time parents, compared to longer promoter region alleles (Hammock & Young, 2005). Male prairie voles with longer alleles display more frequent pup grooming, reduced latency for investigating a juvenile animal, reduced latency to approach social odours and a higher frequency and duration of investigating social odours, compared to shorter allele males (Hammock & Young, 2005). Longer allele males also display partner preferences, whereas shorter allele males do not; this may be due to better social discrimination in the long allele males (Hammock & Young, 2005). A recent transgenic study in mice found that different alleles of the prairie vole *Avpr1a* promoter region were associated with differences in coping strategy during a forced swim test (Donaldson & Young, 2013). Overall, the evidence from these classic vole studies suggests that variation in the 5' flanking region of the *Avpr1a* gene results in differential patterns of V1a receptor binding in the brain, which in turn result in inter- and intraspecies differences in social behaviours (Donaldson & Young, 2013; Hammock & Young, 2002; Hammock & Young, 2005; Young et al., 1999).

Similar associations have been found in a variety of species (Donaldson & Young, 2008; McCall & Singer, 2012). For example, in rodents, vasopressin and the V1a receptor are implicated in maternal responsiveness and defence, social bond formation, social memory and stress neurobiology (Freeman, 2016; Kormos & Gaszner, 2013). In non-human primates, correlations have been found between *Avpr1a* genotype and sociality and personality traits

(Hopkins, Donaldson, & Young, 2012; Mahovetz, Young, & Hopkins, 2016; Staes et al., 2015; Staes et al., 2016). While an association between promoter region variation, neuroanatomical differences, and behavioural differences is observed in multiple species, the specific brain regions and behaviours affected by *Avpr1a* variation vary between species (Donaldson & Young, 2008; McCall & Singer, 2012). Thus, to understand the role of *AVPR1A* in humans, we need to determine the function of variation in the human *AVPR1A* promoter region.

Evidence from Human Studies

The association between *AVPR1A* genotype and V1a receptor distribution in humans can be investigated using post-mortem studies. At least two studies have investigated the correlation between *AVPR1A* genotype and receptor distribution in specific areas of the human brain. Knafo, Israel and colleagues (2008) examined the association between *AVPR1A RS3* genotype and hippocampal mRNA levels. Higher *AVPR1A* mRNA levels were found in post-mortem hippocampal samples from individuals with long *AVPR1A RS3* repeats, compared to samples from individuals with short *RS3* repeats (Knafo, Israel, et al., 2008). These repeat lengths had previously been associated with altruistic behaviour; participants homozygous for the short *RS3* allele displayed less altruistic behaviour and self-reported prosocial behaviour, compared to participants homozygous for the long *RS3* allele (Knafo, Israel, et al., 2008). Additionally, a SNP in the *AVPR1A* gene has been associated with gene expression in post-mortem human prefrontal cortex samples (Maher et al., 2011). The rs11174811 variant is thought to disrupt a microRNA binding site. *AVPR1A* expression levels were higher in samples from individuals homozygous or heterozygous for the minor T allele of this variant, compared to samples from individuals homozygous for the major A allele (Maher et al., 2011). In the same study, the rs11174811 SNP was also associated with

spousal satisfaction (measured using the Family Assessment Measure of the Dyadic Adjustment Scale; Spanier, 1976) and substance use in men (Maher et al., 2011).

Human *AVPR1A* promoter region variants have also been associated with promoter activity *in vitro*. Using the neuroblastoma cell line SH-SY5Y, shorter *AVPR1A RS3* and *RS1* allele lengths have been associated with decreased promoter activity (Tansey et al., 2011). Cell lines homozygous for a long *RS1* allele displayed 2.7 times higher promoter activity than cell lines homozygous for a short *RS1* allele. For *RS3*, the long homozygote was 1.4 times more active than the short homozygote (Tansey et al., 2011). This demonstrates that *RS3* and *RS1* variant length is capable of modulating promoter activity.

A limited number of studies have investigated the association between *AVPR1A* genotype and human brain structure and function (Zink & Meyer-Lindenberg, 2012). Grey matter volume in the right fusiform face area (FFA) is associated with *AVPR1A RS3* genotype in men; men homozygous for long alleles have higher grey matter volume in this area than males carrying one or two copies of short alleles (Wang et al., 2016). Allele length and grey matter volume were also positively correlated with altruistic behaviour (Wang et al., 2016). *AVPR1A* genotype has also been associated with amygdala activation in response to images of angry and fearful facial expressions. Longer *RS3* variants and shorter *RS1* variants are associated with stronger activation in response to these emotional stimuli (Meyer-Lindenberg et al., 2009).

Together, the evidence from human and non-human mammal studies supports the functional role of tandem repeat variants in the 5' flanking region of the *AVPR1A* gene. Thus, *AVPR1A* promoter region variants, particularly the *RS3* variant, are good candidates for investigating the relationship between the vasopressin V1a receptor and human behaviour. To develop hypotheses about the role of *AVPR1A* promoter region variation in human social behaviour, cognition and health, it is important to examine the evidence for the role of the

vasopressin system in human behaviour. Exogenous vasopressin administration studies support the role of vasopressin in responses to social cues and social stressors, and in cooperative social behaviours. These associations suggest that vasopressin could moderate the relationship between social interaction and health outcomes.

Social Role of Vasopressin: Evidence from Exogenous Vasopressin Administration Studies

Vasopressin can alter how an individual perceives and responds to social cues. Existing evidence indicates that exogenous vasopressin can alter responses to facial emotional expressions, social interaction and social stressors. By influencing how a person perceives and responds to social cues, vasopressin could modulate the impact of social support and social strain on health.

Facial Emotional Expression Recognition

Intranasal vasopressin administration influences the response to and recognition of facial emotion expression stimuli. The neuropeptide abolishes the reduced activity in the medial prefrontal cortex that typically occurs in response to observing fearful and angry faces (Zink, Stein, Kempf, Hakimi, & Meyer-Lindenberg, 2010). This change is localised to the subgenual cingulate cortex, an area involved in amygdala regulation and fear extinction (Zink et al., 2010). Administering vasopressin also increases agonistic responses in men and affiliative responses in women to same-sex facial emotion stimuli (Thompson et al., 2006). In men, vasopressin increases the corrugator supercilii (frowning) facial electromyography response (EMG) to male neutral faces and reduces participants' approachability ratings of male happy faces (Thompson et al., 2006). Compared to placebo, the neuropeptide also improves men's recognition of the identity of happy and angry faces, but not neutral faces

(Guastella, Kenyon, Alvares, Carson, & Hickie, 2010). In women, vasopressin increases the zygomaticus major (smiling) EMG response to neutral female faces and increases participants' approachability ratings of those faces (Thompson et al., 2006). Together this evidence suggests that vasopressin influences how facial emotional expressions are perceived and responded to, it may also influence identity recognition.

Cooperative Social Interaction

Vasopressin can also alter people's responses during social interactions. In men, administering vasopressin can increase rates of reciprocated cooperation in the Prisoner's Dilemma economic game (Rilling et al., 2012). Vasopressin also increases male participants' willingness to engage in risky mutually-beneficial cooperation in another economic game, the Stag Hunt (Brunnlieb et al., 2016). For women, administering vasopressin increases the probability that a Prisoner's Dilemma participant will make a conciliatory gesture after the other player has defected (Rilling et al., 2014). This suggests that exogenous vasopressin increases cooperative behaviours in men and women. However, vasopressin administration also reduces women's cooperation with a computer partner after mutual defection, indicating that vasopressin may increase participants' anthropomorphisation of their computer partner and subsequently increase aversion to betrayal (Rilling et al., 2014).

Neural responses to social interaction. Changes in behaviour following vasopressin administration are associated with changes in neural activity during social interactions. Compared to placebo, vasopressin augments the male neural response to cooperation in the bilateral insula, the left supramarginal gyrus (Feng, Hackett, et al., 2015), and in a region spanning areas rich in vasopressin receptors in other mammals, including the bed nucleus of the stria terminalis, the stria terminalis and the lateral septum (Rilling et al., 2012). For women, vasopressin attenuates the neural response to cooperation in the bilateral insula and the left supramarginal gyrus (Feng, Hackett, et al., 2015), rendering women's neural

responses more similar to the typical male response to reciprocal cooperation (Feng, Hackett, et al., 2015; Rilling et al., 2014). The consistent role of the insula in these results suggests that vasopressin may modulate the salience of social stimuli, increasing the salience of positive social interactions for men and decreasing it for women (Feng, DeMarco, Haroon, & Rilling, 2015; Feng, Hackett, et al., 2015; Rilling et al., 2012). Additionally, administering vasopressin to male participants down regulates activity in the left dorsolateral prefrontal cortex in response to risky mutually-beneficial cooperation (Brunnlieb et al., 2016). This brain region is associated with risk integration (Brunnlieb et al., 2016).

Vasopressin can also modulate male neural responses to negative social interactions. For example, after unreciprocated cooperation, vasopressin administration increases activation in the cingulate cortex, lateral temporal lobe, and the medial, ventromedial and ventrolateral prefrontal cortex (Feng, DeMarco, et al., 2015; Rilling et al., 2012). The augmented activity in the lateral temporal lobe suggests that vasopressin may increase emotional regulation in response to negative social interaction (Feng, DeMarco, et al., 2015). The effects of the neuropeptide in response to both negative and cooperative interaction are more pronounced in men who self-report high levels of neuroticism (Feng, DeMarco, et al., 2015).

Compared to administering a placebo, administering vasopressin also increases amygdala connectivity with the bilateral ventral anterior insula, subgenual anterior cingulate cortex, and inferior lateral temporal cortex, and decreases amygdala connectivity with the brainstem (Rilling et al., 2012). Increased connectivity between the amygdala and insula may facilitate subjective emotional responses to salient social stimuli, while decreased connectivity between the amygdala and brainstem may indicate that the neuropeptide can modulate the automatic fear response (Rilling et al., 2012). Functional connectivity between the left dorsolateral prefrontal cortex and the ventral palladium is also increased by

vasopressin administration (Brunnlieb et al., 2016). The ventral palladium is associated with vasopressin-mediated social reward processing in other mammals (Brunnlieb et al., 2016). Together, the evidence for the behavioural and neural effects of vasopressin suggest that the neuropeptide can increase the salience of social stimuli and modulate how social interactions are responded to emotionally and behaviourally.

Response to Social Stressors

Exogenous vasopressin also modulates how men respond to social stressors. In male participants, compared to placebo, intranasally administering vasopressin increases autonomic responsiveness to threatening social stimuli, including angry faces, fearful faces and social evaluative threat, i.e. a situation in which one's performance could be negatively judged by others (Shalev et al., 2011; Thompson et al., 2006). It has also been found to increase self-reported state-anxiety after viewing facial emotion expression stimuli, including happy, neutral, fearful and angry expressions (Thompson et al., 2006). However, vasopressin does not increase the physiological response to non-evaluative threat or create a stress response in the absence of a threat or a social stimulus (Shalev et al., 2011). Vasopressin has also been found to modulate male participants' neural response to negative social feedback after task performance. Under placebo, brain regions including the temporoparietal junction, anterior insula, supplementary motor area and fusiform face area preferentially respond to negative social feedback compared to positive social feedback (Gozzi et al., 2017). These areas are associated with Theory of Mind, pain processing and facial emotion processing. After vasopressin administration, this response is attenuated, abolishing the significant difference between the neural response to positive and negative feedback (Gozzi et al., 2017). Evidence for the role of vasopressin in female stress responses is limited, but one study reports that women who have received vasopressin also display increased autonomic responsiveness to

threatening social stimuli and report higher levels of state anxiety after viewing facial emotion expression stimuli (Thompson et al., 2006).

Exogenous vasopressin administration can alter neural, emotional, physiological, and behavioural responses to social stimuli, including positive and negative social interaction. The effects of exogenous vasopressin are not always consistent and may be influenced by drug dosage and timing (Chen et al., 2016). However, current evidence consistently supports a role for vasopressin in modulating the salience of social stimuli and emotional regulation in response to those stimuli, particularly the regulation of fear responses and stress responses (Feng, DeMarco, et al., 2015; Feng, Hackett, et al., 2015; Rilling et al., 2012; Shalev et al., 2011; Thompson et al., 2006). Thus, vasopressin may play a role in moderating the impact of social support and social strain on an individual's physical and mental health. While fewer studies have investigated the effects of vasopressin administration in women, it is important to note that the available evidence suggests that they differ from the effects found in men (Feng, Hackett, et al., 2015; Rilling et al., 2014). The effects of the neuropeptide may also vary with personality traits, such as neuroticism (Feng, DeMarco, et al., 2015).

Social Role of Vasopressin: Evidence from Candidate Gene Research

AVPR1A and Responses to Social Stimuli

As previously discussed, individual differences in vasopressin-related behaviours can be driven by variation in the *AVPR1A* gene, particularly by variants in the 5' promoter region (Hammock & Young, 2002; Hammock & Young, 2005; Insel & Shapiro, 1992; Insel et al., 1994). Variants in this gene have been associated with individual differences in behaviours and traits that influence how a person responds to their social environment. *AVPR1A* may also play a role in how people respond to stressors, particularly social stressors. By

influencing an individual's response to social support and social strain, genetic variation in the *AVPR1A* gene may moderate the impact of social interaction on health.

The *AVPR1A* gene has been linked with personality and behavioural traits that may influence an individual's response to social interaction, including approach behaviour, avoidance behaviour and related personality traits. *AVPR1A RS3* allele length is associated with self-reported impulsive aggression in people with Borderline Personality Disorder (Vogel et al., 2012). Differences in self-reported aggression were tested against multiple, systematic splits of allele length into short and long categories. Participants carrying short alleles reported significantly higher levels of impulsive aggression, compared to participants carrying long alleles. This relationship held across all modelled allele categories (Vogel et al., 2012). *RS1* allele length was not associated with impulsive aggression (Vogel et al., 2012), but is associated with novelty seeking and harm avoidance (Meyer-Lindenberg et al., 2009). Adult carriers of the 320bp *AVPR1A RS1* allele report increased novelty seeking and decreased harm avoidance compared to non-carriers (Meyer-Lindenberg et al., 2009). Together, these studies suggest that *AVPR1A* genotype may influence approach behaviours. Evidence also suggests that *AVPR1A* genotype is associated with the Behavioural Inhibition System (BIS). Behavioural inhibition is a tendency to respond to unfamiliar people, environments or objects with cautious, avoidant and restrained behaviours (Bisceglia, Jenkins, Barr, Wigg, & Schmidt, 2014). In adult men and women, a functional SNP in the *AVPR1A* gene, rs11174811, is associated with a self-report measure of BIS sensitivity (Reuter, Cooper, Smillie, Markett, & Montag, 2015). In children, SNPs in the *AVPR1A* and *AVPR1B* genes were not found to be associated with observed behavioural inhibition (Bisceglia et al., 2014). However, a SNP in the *AVP* (arginine vasopressin) gene, rs3761249, was associated with individual differences in BIS sensitivity (Bisceglia et al., 2014). Short tandem repeat variation was not examined in these studies.

Variation at the *AVPR1A RS1* locus has also been associated with sexual behaviour in adolescent women (Prichard, Mackinnon, Jorm, & Easteal, 2007). Women homozygous for longer *RS1* alleles have been found to be more likely to have sexual intercourse before the age of fifteen (reported retrospectively), compared to women carrying medium or shorter length alleles (Prichard et al., 2007). In the same study, *RS3* allele length was nominally associated with precocious age of first intercourse in men; men homozygous for longer *RS3* alleles were more likely to report having intercourse before the age of fifteen (Prichard et al., 2007).

In addition to personality and behavioural traits, *AVPR1A* variation may influence how threatening stimuli are processed and responded to. Prepulse inhibition is an automatic response that is used to measure sensorimotor gating, i.e. the ability to ignore or inhibit perception of a sensory stimulus (Levin et al., 2009). This response is hereditary and may be associated with social cognition in humans. Adults carrying longer *AVPR1A RS3* alleles have been found to display greater prepulse inhibition (inhibition of the startle response) compared to adults with shorter *RS3* alleles (Levin et al., 2009). Carriers of the *RS1* 320bp allele report less positive emotions following exposure to social stress in the Trier Social Stress Test, compared to non-carriers (Moons, Way, & Taylor, 2014). Additionally, male carriers of this allele with relatively higher levels of post-stress plasma vasopressin respond to social stress with more anger, compared to non-carriers (Moons et al., 2014). Together, these studies suggest that variation in the 5' promoter region of the *AVPR1A* gene can influence how startling or threatening stimuli are responded to.

As part of a multigene risk score, the *AVPR1A RS3* variant has been associated with mental health outcomes following war exposure in children (Feldman, Vengrober, & Ebstein, 2014). The *RS3* 334bp allele was included in a risk score with 4 other risk markers from the *OXTR* and *CD38* genes. Risk was calculated by adding the total number of risk markers carried by an individual (Feldman et al., 2014). In war exposed children, a higher genetic risk

score was positively associated with a child's risk of developing a mental illness. This association was not observed in children who had not been exposed to war (Feldman et al., 2014). Child, mother and father genetic risk scores were also higher in families where war exposed children developed chronic Post Traumatic Stress Disorder (PTSD), compared to families where the child recovered from PTSD during the study period (Feldman et al., 2014). This evidence, though indirect, supports the hypothesis that *AVPR1A* may play a role in threat processing or response.

AVPR1A and Empathy

Empathy is a key part of healthy social, emotional and moral development (Allemand, Steiger, & Fend, 2015; Tone & Tully, 2014). Being over- or under-sensitive to the feelings and thoughts of other people can be detrimental to social functioning and mental health. Candidate gene research has linked *AVPR1A* variation with individual differences in empathy and empathy-related traits. This may be another pathway via which *AVPR1A* can influence the impact of social support and social strain on health.

Empathy is defined as the ability to recognise and understand the emotions, thoughts and intentions of other people (Decety & Jackson, 2004). It is generally split into two related but different domains; emotional empathy and cognitive empathy (Decety & Jackson, 2004; Singer, 2006). Emotional empathy refers to the ability of an individual to share or experience the emotions of another person upon observing their emotional state, to recognise that emotion state, and to be aware that it originated within the other, not within the self (de Vignemont & Singer, 2006; Decety & Jackson, 2004; Fan, Duncan, de Greck, & Northoff, 2011; Singer, 2006). It is important to note that emotional empathy is different to sympathy and personal distress. Sympathy is a feeling of sadness or concern in response to the emotional state of another person (Decety & Jackson, 2004) and thus lacks the emotional mirroring that is central to the concept of emotional empathy (de Vignemont & Singer, 2006;

Maibom, 2012). Personal distress describes feelings of anxiety or distress triggered by observing the emotional state of another person (Davis, 1980). This also lacks the emotional mirroring of true emotional empathy and is a self-oriented rather than an other-oriented response (Davis, 1980; Maibom, 2012). Cognitive empathy, also referred to as mentalising or Theory of Mind, encompasses the ability to represent and understand the mental states of other people, including their emotions, thoughts, beliefs, desires and intentions (Decety & Jackson, 2004; Singer, 2006). It encompasses the ability to take the perspective of another person (Decety & Jackson, 2004), i.e. the ability to "step into another person's shoes".

Successful empathy development contributes to healthy social, emotional and moral development (Allemand et al., 2015; Tone & Tully, 2014). However, being overly sensitive to the negative emotions of other people can also be detrimental. While empathy is usually protective for mental health, it has been theorised that high empathic sensitivity can be a 'risky skill' (Tone & Tully, 2014). For example, while better perspective-taking skills are associated with positive friendship quality in adolescence, they can also result in higher levels of empathic distress (Smith & Rose, 2011). Empathic distress involves sharing a relationship partner's distress as if it were one's own (Smith & Rose, 2011). Thus, there appears to be a trade-off between the relationship-improving qualities of empathy and the emotional consequences of shared distress. High sensitivity to others' distress could lead to maladaptive self-focussed responses, including increased personal distress and excessive interpersonal guilt (Tone & Tully, 2014). These cognitive and affective responses have been linked to internalising disorders, social isolation, neuroticism and burnout (Smith & Rose, 2011). Thus, if *AVPR1A* is associated with individual differences in empathy, *AVPR1A RS3* genotype may moderate the impact of social interactions on health.

A recent study found an association between *AVPR1A* variation and self-reported empathy, particularly cognitive empathy (Uzefovsky et al., 2015). Carrying one or two

copies of an allele length previously found to be nominally over-transmitted in Autism Spectrum Disorders, 327bp (or 334bp depending on the genotyping method used; Kim et al., 2002), predicted lower total empathy and lower cognitive empathy, compared to individuals who did not carry the allele. The 327bp allele was not associated with emotional empathy (Uzefovsky et al., 2015). The study included both male and female participants and sex was found to predict self-reported empathy (Uzefovsky et al., 2015). The interaction between genotype and sex was not examined. Given the commonly found sex differences in the effects of vasopressin and *AVPR1A* variation (Albers, 2012; Dumais & Veenema, 2016; McCall & Singer, 2012) and the association between sex and self-reported empathy (Maibom, 2012), it will be important for future research to determine if there are sex differences in the impact of *AVPR1A* variation on empathy.

Total empathy was measured using a composite score from three measures: (1) the Interpersonal Reactivity Index (IRI; Davis, 1983), (2) the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004), and (3) the Questionnaire Measure of Emotional Empathy (QMEE; Mehrabian & Epstein, 1972). The cognitive and emotional domains of empathy were measured using the subscales of the IRI (Uzefovsky et al., 2015). The IRI measures four factors (Davis, 1980):

1. Perspective-Taking: the tendency to spontaneously adopt the psychological viewpoint of others
2. Fantasy: the tendency to transpose oneself into the feelings and actions of fictional characters in books, movies and plays
3. Empathic Concern: other-oriented feelings of sympathy and concern for unfortunate others
4. Personal Distress: self-oriented feelings of personal anxiety and unease in tense interpersonal settings

As is commonly done, perspective-taking and fantasy were combined to form a measure of cognitive empathy, and empathic concern and personal distress were combined into a measure of emotional empathy (Uzefovsky et al., 2015). The factor descriptions indicate that while the IRI provides a self-report measure of cognitive empathy, it does not measure emotional empathy. Instead it measures sympathy and personal distress, constructs which are not included in the definition of emotional empathy utilised in this thesis (de Vignemont & Singer, 2006; Decety & Jackson, 2004; Maibom, 2012; Singer, 2006). Thus, while the results of Uzefovsky and colleagues (2015) provide evidence for a link between *AVPR1A* and cognitive empathy, they do not rule out a potential association between *AVPR1A* and emotional empathy. A relationship between *AVPR1A RS3* may have health consequences; having too much or too little empathy can be detrimental for mental health (Smith & Rose, 2011; Tone & Tully, 2014).

***AVPR1A*, Autism Spectrum Disorders and Autistic Traits.** Deficits in cognitive empathy, or Theory of Mind, are a key feature of Autism Spectrum Disorders (ASDs; World Health Organization, 1993). ASD risk has repeatedly been associated with variation in the *AVPR1A* gene, though the implicated variants and strength of the association differs between studies. ASD risk is most commonly associated with transmission disequilibrium for *RS1* alleles (Kantojärvi et al., 2015; Tansey et al., 2011; Wassink et al., 2004; Yang et al., 2010). Shorter *RS1* alleles were found to be over-transmitted in ASD in two studies (Kantojärvi et al., 2015; Tansey et al., 2011) and a specific allele has been found to be under-transmitted in a non-language impaired group of people with ASD (Wassink et al., 2004). However, other studies find no direct association between *RS1* and ASD (Kim et al., 2002; Yirmiya et al., 2006). ASD risk has also been linked to the *RS3* and *AVR* (an intronic microsatellite) variants, a SNP (rs7307997) in the 5' promoter region of *AVPR1A,* and a haplotype of *RS1*, *RS3* and *AVR* (Kantojärvi et al., 2015; Kim et al., 2002; Wassink et al., 2004; Yang et al.,

2010; Yirmiya et al., 2006). Again, these results are not consistent across studies (Kantojärvi et al., 2015; Tansey et al., 2011) and specific non-randomly transmitted alleles differ between studies (Wassink et al., 2004). The inconsistencies in the associations that have been found may be due to differences between studies in methodology, diagnostic criteria or population. Together the results provide some support for a role of *AVPR1A* in the aetiology of ASD.

At the general population level, autistic traits, as measured by the Autism Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), are associated with *AVPR1A* variation. The AQ is designed to measure the extent to which an adult with normal intelligence has five key traits associated with the autism spectrum: social skills, attention switching, attention to detail, communication and imagination (Baron-Cohen, Wheelwright, Skinner, et al., 2001). Women who are homozygous for longer *RS3* alleles score higher on the AQ than those with at least one copy of a shorter allele (Procyshyn, Hurd, & Crespi, 2017). The same genetic association was not found for men but this may have been due to the smaller sample size and higher variability of this group (Procyshyn et al., 2017). The commonly investigated 334bp target allele was also associated with AQ scores; carriers of this allele scored higher on the AQ than non-carriers and a linear trend between allele copy number and AQ score suggest that this effect may be dose dependent (Procyshyn et al., 2017). Again, this relationship was primarily driven by female participants. The evidence linking *AVPR1A*, ASD and autistic traits provides support for the hypothesis that variation in this gene is associated with individual differences in social cognition, including empathy.

AVPR1A and Empathy-related Traits

The association between *AVPR1A* and empathy has not been replicated, however, there is considerable evidence linking variation in *AVPR1A* to individual differences in a range of social behaviours that are likely to be associated with empathy, including altruism (Avinun et al., 2012; Knafo, Israel, et al., 2008) and maternal sensitivity (Bisceglia et al.,

2012; Leerkes, Su, Calkins, Henrich, & Smolen, 2017). In a biological sense, altruism is defined as a behaviour that increases the fitness of the recipient at the cost of the actor (de Waal, 2008). It has been argued that empathy motivates altruism (de Waal, 2008). By reflecting the emotions of another individual within ourselves, empathy makes the distress of that individual our own, therefore, we are motivated to engage in altruistic behaviour in order to relieve both their distress and our own (de Waal, 2008). Thus, evidence for an association between *AVPR1A* variation and altruistic behaviour suggests that *AVPR1A* may play a role in empathy. Associations between *AVPR1A* and maternal sensitivity also suggest that that *AVPR1A* may play a role in empathy. Maternal sensitivity is a mother's ability to accurately perceive and respond to her infant's cues promptly and appropriately (Ainsworth, Blehar, Waters, & Wall, 1978). This concept is similar to empathy and, as such, associations between maternal sensitivity and *AVPR1A* variation provide evidence for an association between this gene and empathy.

***AVPR1A* and altruism.** The relationship between *AVPR1A* variation and human altruism has been investigated using an economic game called the Dictator Game (Avinun et al., 2011; Knafo, Israel, et al., 2008; Wang et al., 2016). In this game, the player is asked to divide a set amount of money between themselves and a second player. Player one can give any amount of money to player two, from $0 to the total amount of money available and the second player must accept whatever division the first player decides on. Players are anonymous so there is no risk of retribution for player one's decision. Despite there being no negative consequences if player one allocates all the money to themselves, most players will allocate some money to player two (Forsythe, Horowitz, Savin, & Sefton, 1994). Allocating money to player two represents an act of altruism because there is no benefit to player one in giving away money and no risk to player one in keeping all the money for themselves (Forsythe et al., 1994). The amount of money allocated to player two can be used as a

measure of individual differences in altruism (Knafo, Israel, et al., 2008; Wang et al., 2016). The relationship between altruism and variation in the length of the *AVPR1A RS3* microsatellite was first investigated in a sample of 203 male and female university students who participated in the Dictator Game, and their parents who were also genotyped for the purposes of family-based analysis. In both a population- and family-based analysis, individuals homozygous for longer *RS3* alleles were found to allocate significantly more money to player two than those homozygous for shorter *RS3* alleles (Knafo, Israel, et al., 2008). Money allocation did not differ by sex and there was no interaction between participant sex and *AVPR1A RS3* genotype (Knafo, Israel, et al., 2008).

The relationship between *AVPR1A RS3* length and Dictator Game allocation has been replicated in a sample of Han Chinese men (Wang et al., 2016). In this sample, men carrying two copies of a long *RS3* allele allocated more money to player two than men with one or two copies of a short *RS3* allele. However, a relationship between allele length and money allocation was not found for women (Wang et al., 2016). In men, allele length was also associated with grey matter volume in the right fusiform face area (FFA) in a whole-brain analysis, and grey matter volume in the FFA was found to mediate the relationship between genotype and money allocation. Long *RS3* repeats predicted greater grey matter volume, which in turn was associated with higher money allocation in the Dictator Game (Wang et al., 2016). Together, these studies suggest that individuals homozygous for long *AVPR1A RS3* alleles are more altruistic than individuals carrying one or two copies of a short *RS3* allele (Knafo, Israel, et al., 2008). It is unclear whether this relationship differs between men and women.

The relationship between *AVPR1A RS3* genotype and altruism is further supported by a study conducted in a sample of pre-schoolers. The Dictator Game was modified so that it could be easily understood by a pre-school student and monetary rewards were replaced with

sticker charts (Avinun et al., 2011). Rather than comparing short and long alleles of *AVPR1A RS3*, this study compared the commonly studied *RS3* target allele (a 334bp or 327bp repeat depending on the PCR primers used) with all other alleles found in the sample (Avinun et al., 2011). Ninety-eight three-and-a-half-year-old twin pairs (34 monozygotic and 64 dizygotic) participated in the study (Avinun et al., 2011). Children carrying the target allele (homozygotes and heterozygotes) were significantly less likely to donate at least one sticker chart to another unknown child than children who did not carry the target allele (Avinun et al., 2011). This relationship was found in both a population-based and a family-based analysis. The potential interaction between genotype and sex was not investigated (Avinun et al., 2011).

These results support the relationship between *AVPR1A* variation and individual differences in altruism found in adults (Knafo, Israel, et al., 2008; Wang et al., 2016), however, the precise nature of this relationship differs between the age groups. The target *RS3* allele associated with lower altruistic behaviour in children, 334bp, was included in the long *RS3* allele category associated with more altruistic behaviour in adults (Avinun et al., 2011; Knafo, Israel, et al., 2008; Wang et al., 2016). While these results appear to contradict each other, the authors suggest that this discrepancy could be due to the age difference between the two samples; the development of cognitive and emotional abilities related to altruism are likely to differ between pre-schoolers and adults (Avinun et al., 2011). For example, the development of both cognitive and emotional empathy differs with age (Knafo et al., 2009; Knafo, Zahn-Waxler, Van Hulle, Robinson, & Rhee, 2008).

Additionally, although long *AVPR1A RS3* alleles were associated with greater grey matter volume in the right FFA, a post-hoc analysis indicated that the target *RS3* allele was associated with smaller grey matter volume in this area (Wang et al., 2016). This discrepancy in results suggests the relationship between *AVPR1A* and behaviour may be more complex

than a short/long division of alleles would suggest. However, current evidence from animal (Hammock & Young, 2005), *in vitro* (Tansey et al., 2011) and human research (Knafo, Israel, et al., 2008; Vogel et al., 2012) suggests that relative allele length is an appropriate categorisation strategy. As altruism is motivated by empathy (de Waal, 2008), the results from both childhood and adult studies support a possible role for *AVPR1A* genotype in empathy.

***AVPR1A* and maternal sensitivity.** An association between maternal sensitivity and *AVPR1A* variation provides further support for the role of this gene in empathy. Maternal sensitivity is a mother's ability to accurately perceive and respond to her infant's cues promptly and appropriately (Ainsworth et al., 1978). Bisceglia and colleagues (2012) investigated whether there was a relationship between *AVPR1A RS3* allele length and maternal sensitivity in 337 mothers. Interactions between mothers and their children were videotaped and rated for maternal sensitivity (Bisceglia et al., 2012). *RS3* allele length was categorised into short and long allele groups and mothers homozygous for the long allele were rated as significantly less sensitive than heterozygous mothers or mothers homozygous for the short allele. This result suggests there is a relationship between *AVPR1A* variation and individual differences in maternal sensitivity (Bisceglia et al., 2012).

This relationship is supported by a recent study which found an indirect association between *AVPR1A RS3* genotype and maternal sensitivity. Leerkes and colleagues (2017) found an association between *AVPR1A RS3* allele length and self-reported mother-oriented cry processing. Mothers carrying longer *RS3* alleles were more likely to focus on their own needs and endorse negative cognitions about their infant's distress, compared to mothers not carrying the longer allele (Leerkes et al., 2017). *RS3* allele length did not have a direct effect on maternal sensitivity, but the indirect effect of genotype on maternal sensitivity via infant cry processing was significant. Mothers carrying longer alleles were more likely to report

self-focussed cry processing and display lower maternal sensitivity (Leerkes et al., 2017). These results align with the findings of Bisceglia and colleagues (2012). Maternal sensitivity and empathy are similar concepts; thus, together these result support the hypothesis that there is an association between *AVPR1A* and empathy, and between *AVPR1A* and the quality of important social relationships.

AVPR1A and Relationship Quality

AVPR1A variation may also influence the quality of close social relationships, including mother-infant interactions (discussed above) and spousal relationships. In addition to the evidence linking the *AVPR1A* gene to empathy, a trait that contributes to the quality of interpersonal relationships, *AVPR1A* is also associated with relationship satisfaction. Thus, *AVPR1A* genotype may affect the exposure an individual has to positive and negative social interactions.

Beyond mother-infant interactions, empathy contributes to the quality of other interpersonal relationships, including friendships (Chow, Ruhl, & Buhrmester, 2013; Smith & Rose, 2011), and romantic relationships (Cramer & Jowett, 2010). Social perspective-taking ability is positively associated with adolescent friendship quality (Smith & Rose, 2011). In this age group, higher empathy has been associated with more intimacy and better conflict management in dyadic same-sex friendships, resulting in more friendship closeness and less discord (Chow et al., 2013). In romantic relationships, self-reported empathy is positively associated with positive relationship behaviours (warmth, even temper, positive outlook, good communication) and partner satisfaction, and negatively associated with negative relationship behaviours (insensitivity, untrustworthiness, possessiveness; Davis & Oathout, 1987). Perceived empathy has also been found to be positively associated with relationship satisfaction in men and women in heterosexual couples (Cramer & Jowett, 2010).

One study has found a link between *AVPR1A RS3* variation and spousal relationship satisfaction. In a sample of 552 male twin-pairs the commonly studied *AVPR1A RS3* target allele, 334bp, was found to be associated with self-rated partner bonding (Walum et al., 2008). Partner bonding was measured using the Partner Bonding Scale, a specifically developed measure that is a composite of existing behavioural items corresponding to the features that are observed when studying pair-bonding in non-human primates (Walum et al., 2008). Men carrying the 334bp allele scored lower on the Partner Bonding Scale than men not carrying the allele. This effect was dose dependent with men homozygous for the 334bp allele reporting the lowest levels of partner bonding (Walum et al., 2008). Additionally, 334bp homozygotes were more likely to be in non-marriage relationships and more likely to report experiencing a marital crisis in the last 12 months (Walum et al., 2008).

The impact of *AVPR1A RS3* genotype on partner bonding and relationship quality was supported by spouse ratings of marital satisfaction, measured using the Dyadic Adjustment Scale (DAS; Spanier, 1976). Women who were married to 334bp carriers reported lower Affection Expression, Dyadic Consensus and Dyadic Cohesion in their relationships, compared to women married to non-carriers (Walum et al., 2008). No genotype effect was found for women's self-reported partner-bonding, suggesting that this is a male-specific effect (Walum et al., 2008). The association between the *AVPR1A RS3* locus and partner-bonding has not been replicated, though spousal satisfaction in men has been found to be significantly associated with a functional SNP, rs11174811, in the *AVPR1A* gene (Maher et al., 2011). Thus, *AVPR1A* variation may influence the quality of important interpersonal relationships. Subsequently, the gene may impact on physical and mental health.

Conclusion

Social support, particularly perceived social support, is beneficial for our mental and physical health, whereas social strain can be detrimental (Ditzen & Heinrichs, 2014; Seeman

et al., 2014). Genetic variation in the 5' promoter region of the *AVPR1A* gene may drive individual differences in responses to social interaction, or the quality of a person's social interactions and relationships. Subsequently, this may influence the impact of social support and social strain on health. Exogenous vasopressin administration can alter behavioural and neural responses to social interaction. Evidence suggests that the neuropeptide moderates the salience of social interaction and may play a role in the regulation of emotional responses to fear- and stress-provoking stimuli (Feng, DeMarco, et al., 2015; Feng, Hackett, et al., 2015; Rilling et al., 2014; Rilling et al., 2012; Shalev et al., 2011). Variation in the *AVPR1A* gene is also associated with individual differences in behaviours and traits that influence how a person responds to their social environment. These traits include impulsive aggression, novelty seeking, and harm avoidance (Meyer-Lindenberg et al., 2009; Vogel et al., 2012). Variation in this gene also appears to influence how stressful or threatening stimuli are responded to (Feldman et al., 2014; Levin et al., 2009; Moons et al., 2014). Variation in the *RS3* variant, located in the 5' promoter region of the *AVPR1A* gene, is associated with cognitive empathy and empathy-related behaviours, including altruism and maternal sensitivity (Bisceglia et al., 2012; Knafo, Israel, et al., 2008; Uzefovsky et al., 2015). Through its potential role in empathy, *AVPR1A* genotype may influence how sensitive a person is to the feelings and distress of other people, and the quality of their relationships (Cramer & Jowett, 2010; Smith & Rose, 2011; Tone & Tully, 2014). For example, the gene has been linked to pair-bonding and satisfaction in spousal relationships (Maher et al., 2011; Walum et al., 2008). By modulating individual responses to social interactions, or by influencing the quality of a person's close social relationships, *AVPR1A* genotype may underlie individual differences in the impacts of social support and social strain on physical and mental health.

Aims

The available evidence suggests that vasopressin and the *AVPR1A* gene play a role in human social behaviour, cognition and responses to social interaction. Given the established effects of social support and social strain on mental and physical health (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016), it was hypothesised that genetic variation in the *AVPR1A* gene would have consequences for individual health outcomes. Tandem repeat variation at the *RS3* locus, located in the 5'promoter region of the gene, was selected as the candidate variant for this project according to current evidence for its biological function and accumulating evidence for its role in social phenotypes. *AVPR1A RS3* genotype may underlie some of the individual differences in the impact of social support and strain on health by modulating how individuals respond to social stimuli. The project described in this thesis aimed to determine whether variation at the *AVPR1A RS3* locus was directly associated with individual differences in mental and physical health, and whether *AVPR1A RS3* genotype moderated the impact of social support and social strain on health. Converging evidence also suggested that *AVPR1A RS3* genotype is related to individual differences in empathy (Uzefovsky et al., 2015). This relationship may have consequences for health outcomes as having too much or too little empathy can be detrimental for mental health (Tone & Tully, 2014). Individual differences in empathy are also associated with individual differences in the quality of close social relationships (Chow et al., 2013; Davis & Oathout, 1987), a factor that *AVPR1A RS3* variation is also associated with (Walum et al., 2008). Thus, this project also aimed to explore the relationship between *AVPR1A* variation and empathy, seeking to replicate the relationship between *AVPR1A RS3* variation and self-reported empathy (Uzefovsky et al., 2015), and determine whether genotype at this locus was associated with cognitive empathy skills.

Chapter 3: The vasopressin receptor 1a gene (*AVPR1A*) and health: Investigating relationships between genotype, social interaction and mental and physical health

Chapter Summary

Positive social interactions and support from social networks are beneficial for our health, whereas the strain created by negative social interactions is detrimental (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016). The impact of social interaction varies between individuals and evidence suggests this variation is partly driven by biological factors (Ditzen & Heinrichs, 2014). The neuropeptide vasopressin is associated with human social behaviours and stress responses (Ebstein et al., 2012; Shalev et al., 2011). This study aimed to determine whether short tandem repeat variation at the *AVPR1A RS3* locus: (1) had a direct effect on mental and physical health outcomes or (2) moderated the effect of social support and social strain on mental and physical health.

Participants were from the 20+ cohort of the Personality and Total Health (PATH) Through Life project (Anstey et al., 2012). Baseline genotype data (*AVPR1A RS3*) and longitudinal phenotype data (positive social interaction, negative social interaction, mental health, physical health) from three time-periods collected at four-year intervals was used. Two series of nested multi-level linear models were analysed using maximum likelihood estimation with a Gaussian distribution. A 10000-sample stratified bootstrap resampling of individuals was conducted for each model to produce 95% confidence intervals around fixed parameter estimates. Analyses were conducted separately for men and women, and for the physical and mental health dependent variables.

AVPR1A RS3 genotype was significantly associated with women's mental health and women's physical health; women with at least one copy of the 'Short' allele reported better mental and physical health than women homozygous for the 'Long' allele. There was some evidence of an interaction between genotype and social interaction (both positive and negative), suggesting that genotype may moderate the impact of social support and strain on health. However, the overall pattern of interaction effects was unclear. The effects of social

interaction over time were consistent across genotype groups and no direct genotype effects were observed for men. *AVPR1A RS3* genotype is significantly associated with mental and physical health outcomes in women. Genotype may moderate the impact of social interaction on mental health and future research should investigate this mechanism.

The vasopressin receptor 1a gene (*AVPR1A*) and health: Investigating relationships between genotype, social interaction and mental and physical health

Social support and social strain have an important impact on our mental and physical health (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016). The impact of support and strain varies between individuals, and this variation is likely partly due to biological factors (Ditzen & Heinrichs, 2014). Vasopressin, the vasopressin V1a receptor and the gene encoding this receptor (*AVPR1A*) are associated with a range of social behaviours and traits in humans and non-human mammals (Albers, 2012; Hammock & Young, 2005; Insel, 2010). Variation at the *AVPR1A RS3* locus may moderate the influence of social support and social strain on mental and physical health in humans. The study described in this chapter examines the relationship between *AVPR1A RS3* genotype, social interaction and health.

Social Support and Social Strain

Social support. Social support can be defined as the practical and emotional support a person receives from their friends, family members, spouse and other social relationships (Caltabiano et al., 2008). Perceived social support, rather than the actual amount of support an individual receives, is particularly important for health and wellbeing (Caltabiano et al., 2008; Ditzen & Heinrichs, 2014). Individuals can vary in the amount of support they require (Caltabiano et al., 2008) and the impact that social support has on their health (Ditzen & Heinrichs, 2014). As discussed in Chapter 2, social support has a well-established positive influence on mental and physical health (Ditzen & Heinrichs, 2014; Holt-Lunstad et al., 2010; Lee & Szinovacz, 2016). For example, higher levels of perceived social support are associated with higher life satisfaction, better reported mental health and a reduced risk of developing mental illnesses such as major depressive disorder (Lee & Szinovacz, 2016; Santini et al., 2016; Schuster et al., 1990; Teo et al., 2013; Wong et al., 2014).

Social support can attenuate the mental and physiological effects of stressors (Ditzen et al., 2007; Heinrichs, Baumgartner, Kirschbaum, & Ehlert, 2003). This attenuation is one mechanism by which social support is thought to influence health (Caltabiano et al., 2008; Ditzen & Heinrichs, 2014). Experimental studies have demonstrated the impact that support from a friend or spouse can have on responses to psychosocial stressors. The Trier Social Stress Test (TSST) is an experimental protocol that is used to create a naturalistic exposure to a psychosocial stressor in a controlled environment (Kirschbaum, Pirke, & Hellhammer, 1993). It typically features a public speaking task and a mental arithmetic task, presented in front of an unknown audience and preceded by a short preparation time (Ditzen et al., 2007; Heinrichs et al., 2003; Kirschbaum et al., 1993). This protocol has been used to determine the effect of social support on psychological stress reactivity (Ditzen et al., 2007; Heinrichs et al., 2003). For men, instrumental and emotional social support provided by a close friend reduces salivary cortisol and state anxiety responses to the TSST, and increases self-reported calmness during the stressor (Heinrichs et al., 2003). Women who receive positive physical contact from their partner prior to the stressor display significantly lower heart-rate and salivary cortisol increases after the stressor (Ditzen et al., 2007). Verbal support from a partner did not significantly reduce women's physiological stress reactions compared to a no support condition (Heinrichs et al., 2003). Thus, certain kinds of social support can moderate the impact of stressors, and the most effective kinds of support appear to differ by sex (Ditzen et al., 2007; Heinrichs et al., 2003).

Social strain. Social relationships can also involve negative interactions, including conflict, tension, criticism, irritation and disappointment (Lee & Szinovacz, 2016; Schuster et al., 1990). As discussed in Chapter 2, negative social interactions have been consistently associated with poorer health outcomes and elevated levels of psychological distress (Schuster et al., 1990; Seeman et al., 2014; Yang et al., 2014). For example, two prospective

studies have linked the quality of an individual's social relationships with their risk of experiencing depression (Santini et al., 2016; Teo et al., 2013). Poor quality relationships with friends, family and spouse, and high levels of perceived loneliness have been associated with a higher risk of experiencing depression (Santini et al., 2016; Teo et al., 2013). Teo and colleagues (2013) found that the subjective perception of the quality of social relationships was more important than their objective qualities. A person's relationship status and their frequency of social contact were not associated with depression risk and did not interact with relationship quality to predict depression risk (Teo et al., 2013). In regards to physical health, experiencing higher levels of social strain has been linked to increases in inflammatory responses (Miller et al., 2009; Seeman et al., 2014; Yang et al., 2014) and allostatic load, a multi-system measure of biological risk (Seeman et al., 2014).

The biological mechanisms underlying the relationship between social strain, social support and health are not well understood (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016). Vasopressin and *AVPR1A* have been associated with reactivity to social stress and other social stimuli. Thus, *AVPR1A* genotype may play a role in the pathway between social strain, social support and health. This potential relationship is discussed in detail below.

Vasopressin and Responses to Social Stimuli

As discussed in Chapter 2, exogenous vasopressin administration can influence how an individual perceives and responds to social cues. For example, compared to placebo, intranasally administering vasopressin increases agonistic responses in men and affiliative responses in women to same-sex facial emotion stimuli (Thompson et al., 2006). After neuropeptide administration, men and women also display increased autonomic responsiveness to angry and fearful faces, and report increased state anxiety after viewing facial emotional expression stimuli including happy, neutral, fearful and angry expressions (Thompson et al., 2006). Exogenous vasopressin also influences neural responses to facial

stimuli, abolishing the reduced activity in the medial prefrontal cortex that typically occurs in response to observing fearful and angry faces (Zink et al., 2010). Thus, vasopressin may influence how a person responds to social interaction.

Vasopressin can also alter physiological responses to psychosocial stressors. This has been demonstrated in a study examining the responses of 157 male participants to the TSST and other, non-evaluative stressors. Under conditions of evaluative social threat, intranasal administration of vasopressin results in an increased pulse rate and salivary cortisol response, compared to placebo (Shalev et al., 2011). Under conditions of no stressor and non-evaluative stressors (a modified TSST with no audience and a physical exercise condition), vasopressin was not found to have the same effect (Shalev et al., 2011). Vasopressin has also been found to modulate male participants' neural response to negative social feedback after task performance (Gozzi et al., 2017). Vasopressin administration attenuates the neural response to negative feedback in brain regions associated with Theory of Mind, pain processing and facial emotion processing, rendering the response more similar to that stimulated by positive social feedback (Gozzi et al., 2017). While the direction of the relationship is unclear, this indicates that vasopressin can alter an individual's response to psychosocial threat and may influence the impact of social stress on physical and mental health.

AVPR1A Genotype and Responses to Social Stimuli

Variation in the *AVPR1A* gene has been associated with several behaviours and traits that can affect an individual's response to their social environment. Three variants in the gene have been associated with approach and avoidance traits. *RS3* allele length is associated with self-reported impulsive aggression in people with Borderline Personality Disorder (Vogel et al., 2012). Though not associated with aggression, the 320bp allele of the *RS1* locus is associated with increased novelty seeking and decreased harm avoidance in adults (Meyer-

Lindenberg et al., 2009; Vogel et al., 2012). Lastly, a functional SNP in the *AVPR1A* gene, rs11174811, is associated with self-reported behavioural inhibition (the tendency to respond to unfamiliar people, environments or objects with cautious, avoidant and restrained behaviours) in adults (Bisceglia et al., 2014; Reuter et al., 2015). However, this SNP was not found to be associated with observed behavioural inhibition in children (Bisceglia et al., 2014).

In addition to approach and avoidance traits, there is evidence linking *AVPR1A* variation with differences in how individuals process social stimuli, including stressful stimuli. Adults carrying longer *RS3* alleles have been found to display greater prepulse inhibition (an automatic response associated with social cognition) compared to adults carrying shorter alleles (Levin et al., 2009). In relation to stressful social stimuli, adult carriers of the 320bp *RS1* allele report less positive emotions following exposure to social stress in the TSST, compared to non-carriers (Moons et al., 2014). Male carriers of this allele who have relatively high levels of post-stress plasma vasopressin also respond to social stress with more anger, compared to non-carriers (Moons et al., 2014). Together, these studies suggest that *AVPR1A* genotype may influence the response of individuals to social support and social strain. Thus, *AVPR1A* genotype may moderate the impact of support and strain on health, explaining some of the between-individual variation in the health effects of social stimuli.

Given the evidence supporting the role of the *AVPR1A* gene in social interaction and the role of social support and strain in physical and mental health, it was of interest to determine whether *AVPR1A* variation has consequences for health and wellbeing. The study described in this chapter examined whether *AVPR1A* genotype is associated with physical and mental health outcomes, and whether genotype moderates the impact of social support and social strain on health. The study focuses on the *AVPR1A RS3* short tandem repeat (STR)

locus because, as demonstrated in Chapter 2, there is evidence to support the biological function of this region (Donaldson & Young, 2013; Hammock & Young, 2005; Knafo, Israel, et al., 2008; Tansey et al., 2011) and its role in human social behaviour (Ebstein et al., 2012).

Aim

The aim of the analyses described in this chapter was to determine whether variation in the *AVPR1A RS3* genotype (1) resulted in a direct effect on physical and mental health outcomes or (2) whether *AVPR1A RS3* genotype moderated the effect of social support and social strain on physical and mental health. This moderation effect was hypothesised to only occur for men, or to occur more strongly in men than in women.

Methods

Participants

The participants for this project were from the 20+ cohort of the Personality and Total Health (PATH) Through Life project; a description of this project and its aims and purpose can be found in Anstey and colleagues (2012). The participants in the 20+ cohort were born between 1975 and 1979. Participants in the PATH study are residents of the Australian Capital Territory and the neighbouring town of Queanbeyan, New South Wales, who were drawn from the electoral rolls in these areas (Anstey et al., 2012)**.** This recruitment method provides a representative sample because Australian citizens aged 18 years and over are legally required to be enrolled to vote. The present study uses a subset of the data from the first three waves of data collection, obtained through interviews at four-yearly intervals (Anstey et al., 2012). The PATH project includes a range of genetic and phenotypic measures, the measures used in the present study are described in detail below. The ethical

aspects of the PATH study were approved by The Australian National University Human Research Ethics Committee.

Time

As there were only three waves of observations spaced four years apart, time was treated as a discrete variable (category labels: Wave 1, Wave 2 and Wave 3). Baseline observations (Wave 1) were treated as the reference category.

Missing Data

Participants with missing data were retained in the data set and missing values were not imputed. Parameter estimates were calculated using maximum likelihood estimation, which can account for missing data. However, for bootstrapped confidence intervals, participants with missing data were excluded list-wise from the analysis.

Questionnaires

Social interaction scales. Perceived positive and negative social interaction with friends, family and partners were measured using scales developed by Schuster and colleagues (1990). These scales were administered at each wave of data collection. Positive interactions with friends and family were measured in separate scales using two parallel items (e.g., "How often to friends/family make you feel cared for?"), and negative interactions with friends and family were measured using three parallel items (e.g., "How often do friends/family make too many demands on you?"; Schuster et al., 1990). Participants rated how often these kinds of interactions occurred on a 4-point Likert scale ranging from "Often" to "Never". Positive and negative interactions with a partner were measured using two five-item scales (Schuster et al., 1990). Items on the Positive Interaction scale described a range of behaviours engaged in by the partner and the respondent (e.g. "How much does your partner understand the way you feel about things?" and "How much can you open up to your

partner about things that are really important to you?") and participants rated how much they, or their partner, engaged in these behaviours using a four-point Likert scale ranging from "A lot" to "Not at all" (Schuster et al., 1990). Items on the Negative Interaction scale described a different range of interactions, one of which was rated on the "A lot" to "Not at all" scale ("How much tension is there between you and your partner?") and the other four rated on the "Often" to "Never" scale (e.g. "How often do you have an unpleasant disagreement with your partner?").

A total score was calculated for each Positive and Negative Interaction scale. Individuals with a partner had six Interaction scores, and individuals without a partner had four Interaction scores. All scales were scored so that a rating of "Often" or "A lot" was scored highest and "Never" or "Not at all" was scored lowest. Thus, a high score on a Positive Interaction scale indicated frequent positive social interactions, and a high score on a Negative Interaction scale indicated frequent negative social interactions. Total scores for each scale were linearly transformed to have a theoretical range between 0 and 1, using the technique described in the original paper (Schuster et al., 1990). This allowed scores to be compared across different scales and for composite Positive and Negative Interaction scores to be calculated. The following equation was used to transform the Interaction Scale total scores for each individual:

$$I_t = {}^{(I - i)}/_{(4i - i)}$$

Where i is the number of items in the scale, I is the original total score for the scale and I_t is the transformed score.

Scores on the Positive and Negative Interactions with Friends, Family and Partner scales were transformed into z-scores to improve the interpretability of their effects in the models. To preserve information about change, z-scores were calculated using a Grand Mean

and Grand Standard Deviation across the three waves of data collection. T-scores (Mean = 50; SD = 10) were then calculated from the z-scores so that the Interaction variables were on a similar scale to the SF-12 variables. Composite Positive and Negative Social Interaction scores were then calculated from the T-scores.

Due to the age range of the participants (20 - 26 years at Wave 1), a minority of participants had data for the Partner Interaction scales at Wave 1. Less than 5% of participants were missing data for Positive and Negative Interactions with Family and Friends. In contrast, as only 558 participants were married or living with de facto partners at Wave 1, 76.7% of participants were missing Partner Interaction data. Additionally, as the data was collected over a long period of time, some participants' relationship status changed between waves. To make use of the full range of Interaction Scales with minimum missing data, a composite Positive Social Interaction and Negative Social Interaction score was calculated for each individual at each wave. As scores from scales of the same valence were moderately positively correlated (see Table 3.1), Positive and Negative composite scores were created by taking the average of the Partner Interaction, Friends Interaction, and Family Interaction scores of the same valence. A composite score was created for all participants with a score for at least one of the component scales. The average was calculated from all available data for a participant at each wave, allowing Partner Interaction data to be included in the composite score despite few participants having partners. Composite scores were moderately-to-strongly correlated with Partner Interaction, Friends Interaction, and Family Interaction scores of the same valence. Positive Social Interaction and Negative Social Interaction were moderately negatively correlated (Table 3.1).

Table 3.1. Bivariate correlations between Positive and Negative Interaction with Friends, Family and Partner scales, and Composite Social Interaction scores at Wave 1.

Variable	Friend Positive	Family Positive	Partner Positive	Friend Negative	Family Negative	Partner Negative	Positive Interaction
Family Positive	*r* = 0.330 *n* = 2382						
Partner Positive	*r* = 0.289 *n* = 559	*r* = 0.210 *n* = 559					
Friend Negative	*r* = -0.157 *n* = 2379	*r* = -0.110 *n* = 2379	*r* = -0.112 *n* = 557				
Family Negative	*r* = -0.139 *n* = 2382	*r* = -0.333 *n* = 2383	*r* = -0.145 *n* = 558	*r* = 0.393 *n* = 2380			
Partner Negative	*r* = -0.308 *n* = 559	*r* = -0.166 *n* = 559	*r* = -0.540 *n* = 559	*r* = 0.188 *n* = 557	*r* = 0.254 *n* = 558		
Positive Interaction	*r* = 0.797 *n* = 2383	*r* = 0.796 *n* = 2384	*r* = 0.693 *n* = 559	*r* = -0.167 *n* = 2380	*r* = -0.284 *n* = 2384	*r* = -0.472 *n* = 559	
Negative Interaction	*r* = -0.200 *n* = 2383	*r* = -0.266 *n* = 2384	*r* = -0.371 *n* = 559	*r* = 0.814 *n* = 2380	*r* = 0.816 *n* = 2384	*r* = 0.672 *n* = 559	*r* = -0.300 *n* = 2385

Note: Coefficients are Pearson correlation coefficients. All correlations significant at $p < 0.01$.

Short-form health survey. Estimates of self-reported physical and mental health were measured using the Physical Component Summary (PCS) and Mental Component Summary (MCS) from the 12-Item Short-Form Health Survey (SF-12; Ware Jr, Kosinski, & Keller, 1996). The SF-12 is a brief overall measure of health designed for large-scale research and has been shown to be a reliable and valid measurement of physical and mental health in the general population (Ware Jr et al., 1996). The PCS measures self-reported physical functioning, the effect of physical functioning on role-based activities (e.g. work activities), bodily pain and general health. The MCS measures self-reported mental health, the effect of mental health (or emotional functioning) on role-based activities (e.g. work activities), the impact of mental health (or emotional functioning) on social activities and vitality (perceived energy levels; Ware Jr et al., 1996). A validation of the MCS has been conducted using data from the PATH study (Kiely & Butterworth, 2015). The study demonstrated that the MCS is a valid instrument for detecting depression and anxiety, as classified by the World Health Organisation Composite International Diagnostic Interview, at the population level and for measuring the severity of these disorders (Kiely & Butterworth, 2015).

Genotyping

Cheek swab samples were collected from the 20+ PATH cohort and genotyped; the genotyping method is described in Prichard and Easteal (2006). Briefly, the *AVPR1A RS3* polymorphism was genotyped as part of a multiplex Polymerase Chain Reaction (PCR)-based assay which simultaneously genotyped individuals at 15 loci across 10 candidate genes (Prichard & Easteal, 2006). The *RS3* locus was amplified with primers: 5'-gctcaaaggcacactgttctc-3' (forward) and 5'-gtcttgggaatctggtcagg-3' (reverse). The 15 loci were amplified by PCR and separated by capillary electrophoresis using an ABI 3730 gene analyzer (Biomolecular Resource Facility, John Curtin School of Medical Research, The

Australian National University) using a multiplex design. The PCR program used was adapted for multiplex PCR of microsatellite loci: "denaturation: 94°C, 30s, annealing: 57°C, 90s, extension: 72°C, 60s for 28 cycles with an additional final step of 60°C, 30 minutes to minimize presence of artifact fragments" (Prichard & Easteal, 2006, p. 3). The *AVPR1A RS3* locus was in Hardy-Weinberg Equilibrium (Prichard & Easteal, 2006) and the allele frequencies for this polymorphism are reported in Table 3.2. The allele labels are those suggested by Prichard and Easteal (2006) and approximately correspond to the number of repeats present. *AVPR1A RS3* contains two dinucleotide repeats and it was not possible to exactly determine genotype using fragment analysis to determine length (Prichard & Easteal, 2006). Thus, alleles for this locus were named by relative fragment length in base pairs, not by absolute number of repeats present. However, a recent validation has found that the fragment length estimation does correspond to total allele repeat length in base pairs[1].

[1] Unpublished data, validation conducted by Susan Tan and Simon Easteal, John Curtin School of Medical Research, The Australian National University.

Table 3.2. *AVPR1A RS3* allele frequencies for the 20+ PATH cohort.

Allele	PCR fragment length (bp)	Number of repeats	PATH frequency
12	254	25	0.013
13	256	26	0.001
14	258	27	0
15	260	28	0
16	262	29	0.005
17	264	30	0.057
18	266	31	0.084
19	268	32	0.242
20[a]	270[a]	33	0.218
21	272	34	0.096
22	274	35	0.132
23	276	36	0.03
24	278	38	0.012
25	280	39	0.070
26	282	40	0.031
27	284	41	0.006
28	286	42	0.003
29	288	43	0

Note: Allele names represent the number of di-nucleotide repeats present (Prichard & Easteal, 2006).

a. Allele 20 (270bp) is equivalent to the 334bp target allele previously found to be nominally over-transmitted in Autism Spectrum Disorders (Kim et al., 2002). Fragment length is proportional to repeat number, but it includes non-repetitive flanking sequence that may vary depending on the genotyping method used.

There is a wide range of *AVPR1A RS3* allele lengths present in the population, and alleles at the extreme ends of this distribution are present at low frequencies (Table 3.2). Consequently, there many genotypes present at low frequencies in the population. In previous literature, short tandem repeat loci with similar allele distributions are commonly analysed by binning allele lengths into 'Short' and 'Long' groups (e.g. Knafo, Israel, et al., 2008). This improves the power of statistical analyses by creating a smaller number of genotype categories with larger cell sizes. Binning *AVPR1A RS3* alleles into Short and Long categories is also supported by evidence from animal and human studies suggesting that allele length is linearly associated with differences in gene transcription (Hammock & Young, 2005; Knafo, Israel, et al., 2008). In the present study, alleles were binned into Short and Long categories based on the median allele length. *RS3* allele lengths of 19 repeats or less were classified as Short and lengths of 20 repeats or more were classified as Long. Based on the Short and Long categories, participants were classified into three genotypes: Short/Short, Short/Long and Long/Long. Genotype sample sizes for each gender are reported in Table 3.3. The Long/Long genotype was used as the reference category, so that analyses compared Long/Long homozygotes to individuals who had one copy of the Short allele (Short/Long) and individuals who had two copies of the Short allele (Short/Short).

Table 3.3. Number of participants in each genotype group at Wave 1.

Genotype	Male	Female
Short/Short	181	178
Short/Long	538	570
Long/Long	389	431
Missing Genotype data	54	63

Statistical Analysis

All analyses were conducted using SPSS Statistics 23. Before analyses were conducted, the distribution of each variable was examined, and all variables were screened for outliers and multicollinearity. Variables were not transformed. Model building analyses were conducted separately for men and women. Vasopressin and *AVPR1A RS3* genotype have previously been found to have different effects in men and women (Albers, 2012; Dumais & Veenema, 2016; Feng, Hackett, et al., 2015; McCall & Singer, 2012), as has social support (Ditzen & Heinrichs, 2014). Separate analyses were also conducted for the Mental Health and Physical Health dependent variables.

Two series of nested multi-level linear models were analysed using maximum likelihood estimation with a Gaussian distribution. Starting from an intercepts-only model, predictors were individually introduced and their effects analysed in subsequent models (Singer & Willet, 2003). All models included a fixed and random intercept. A 10000-sample stratified bootstrap resampling of individuals was conducted for each model to produce 95% percentile confidence intervals (CIs) around fixed parameter estimates. A Wald Z test was conducted to produce 95% percentile CIs around random parameter estimates. The first series of nested models examined the cross-sectional association between the independent variable (IV) and the dependent variable (DV) across multiple periods of time. The second series of models examined whether within-person change in Social Interaction over time was associated with differences in Mental and Physical Health. Both series of models investigated whether there were interactions between *AVPR1A RS3* genotype and Social Interaction variables. The procedures for the analyses are described in detail below. To compare the nested models, three information criteria statistics are reported: -2 Log Likelihood (-2LL), Akaike's Information Criterion (AIC), and Schwarz's Bayesian Criterion (BIC). These

abbreviations are used in the results section and information criteria are displayed in a smaller-is-better form.

Screening for drop-out effect. All variables considered for use in the modelling analyses were screened to determine whether they were associated with participant drop-out using binary logistic regression. Screening was done separately for men and women, as all subsequent analyses were split by gender. Variables included *AVPR1A RS3* genotype, Positive Social Interaction, Negative Social Interaction, Physical Health and Mental Health. These variables were regressed onto the DV, Drop-Out, which was a dichotomous variable indicating whether a participant had completed all three waves of data collection (0, reference category) or had completed less than three waves (1).

Repeated cross-sectional models. Repeated cross-sectional models were built and tested to examine the effect of genotype, social support, social stress and time on physical and mental health at each measurement occasion. Separate analyses were conducted for each DV using the following procedure.

Model 1 was a random intercepts (variance components) model with no predictors used to calculate the intra-class correlation coefficient, which represented the proportion of variance in the DV (Mental Health or Physical Health) due to variation between individuals. Model 2 included the fixed effect of *AVPR1A RS3* genotype. As the Long/Long genotype was the reference category, the intercept of Model 2 represents average health of individuals with the Long/Long *RS3* genotype. Model 3 included the fixed effects of *AVPR1A RS3* and Time. Model 4 included the fixed effects of *AVPR1A RS3*, Time, Positive Social Interaction and Negative Social Interaction. First order interaction effects between *RS3* genotype and Positive and Negative Social Interaction were also tested. Due to the small sample size of the Short/Short genotype group, interaction effects were entered into the model in small groups and tested for significance and impact on model fit.

Analysis of within-person change in Social Interaction. After completing the cross-sectional analysis, a second series of models were constructed to examine whether the associations between Positive Social Interaction, Negative Social Interaction and Physical and Mental Health were due to within-person change in social variables, or between person differences in baseline Negative and Positive Social Interaction. Interactions between *AVPR1A RS3* genotype and the social interaction variables were also examined.

Between-person differences were measured using each individual's baseline score (Wave 1) for Positive Social Interaction and Negative Social Interaction. Within-person change over time was measured by subtracting an individual's baseline score from their score at each wave of data collection (Singer & Willet, 2003; Twisk, 2003). This resulted in every individual receiving a score of 0 for Wave 1, a score representing the difference between Wave 2 and Wave 1, and a score representing the difference between Wave 3 and Wave 1.

As for the repeated cross-sectional models, separate analyses were conducted for the Mental and Physical Health DVs and all models included a fixed and random intercept. Models 1 to 3 were the same as those described for the repeated cross-sectional analysis and so these steps were not repeated. This analysis tested one new model, Model 5, which included the fixed effects of *AVPR1A RS3* genotype, Time, between- and within-person differences in Positive Social Interaction and between- and within-person differences in Negative Social Interaction.

The number of interactions that could be tested simultaneously was limited by the sample size of the Short/Short genotype group. Interactions were entered into the model and tested in small groups. Interactions between *AVPR1A RS3* genotype and between-person differences in Positive and Negative Social Interaction were tested in one model, and interactions between *RS3* genotype and within-person change in Positive and Negative Social Interaction were tested in a separate model.

Results

Participants

Data was collected from the same group of participants on three occasions with approximately four years between data collection periods. The average age of participants, in years, was 22.60 at Wave 1 (*SD* = 1.51, Range: 20 - 26), 26.72 at Wave 2 (*SD* = 1.50, Range: 24 - 30) and 30.69 at Wave 3 (*SD* = 1.50, Range: 28 - 34). The total number of participants in the 20s age-cohort was 2404 (1162 men, 1242 women). Of those, 1917 participants (889 men, 1028 women) were interviewed at all three waves of data collection, 204 participants (118 men, 86 women) were only interviewed at Wave 1, 222 participants (124 men, 98 women) were interviewed at Wave 1 and 2, and 61 participants (31 men, 30 women) were interviewed at Wave 1 and 3. Sample sizes for the continuous variables at each Wave are reported in Table 3.4.

Descriptive Statistics

Descriptive statistics for the continuous variables are reported in Table 3.4. The distributions of the continuous variables were examined at each Wave. Negative Social Interaction was approximately normally distributed. Positive Social Interaction, Mental Health and Physical Health were all negatively skewed. These variables were not transformed as a negatively skewed distribution was expected for these phenotypes and the effect of untransformed variables is easier to interpret. Allele frequencies and genotype category sample size are reported in Table 3.2 and Table 3.3 in the Methods section.

Table 3.4. Descriptive statistics for continuous variables.

	Male					Female				
Variable	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Physical Health										
Wave 1	1155	52.31	6.52	20	59	1229	50.81	7.28	17	59
Wave 2	1000	52.36	6.20	19	59	1108	50.67	7.57	17	59
Wave 3	918	51.89	6.54	17	59	1058	50.27	8.20	17	59
Mental Health										
Wave 1	1155	49.17	9.04	21	65	1230	45.62	9.82	19	65
Wave 2	1004	48.61	9.30	20	64	1119	46.02	9.65	20	64
Wave 3	919	48.73	9.19	18	64	1057	46.43	9.90	17	64
Negative Social Interaction										
Wave 1	1155	51.71	8.35	32.16	79.21	1230	51.72	7.91	32.16	79.21
Wave 2	1012	49.89	7.91	32.16	79.21	1122	49.98	7.92	32.16	79.21
Wave 3	920	48.63	7.56	32.16	79.21	1058	48.77	7.30	32.16	76.85
Positive Social Interaction										
Wave 1	1155	48.88	8.65	3.50	56.70	1230	50.77	7.37	3.50	56.70
Wave 2	1012	49.18	7.90	8.23	56.70	1122	51.10	7.03	13.78	56.70
Wave 3	920	48.48	8.01	15.92	56.70	1058	51.09	7.19	3.50	57.76

Correlations Between Independent Variables

Pearson correlations between the continuous IVs, Positive Social Interaction and Negative Social Interaction, were examined and no evidence of multicollinearity was found. Positive Social Interaction and Negative Social Interaction are moderately negatively correlated, and the degree of correlation between these variables is similar at each wave of data collection (see Table 3.5).

Table 3.5. Correlation between Positive Social Interaction and Negative Social Interaction. Split by Wave and Gender.

Wave	Statistic	Men	Women
Wave 1	r	-0.259	-0.356
	N	1155	1230
Wave 2	r	-0.301	-0.384
	N	1012	1122
Wave 3	r	-0.236	-0.295
	N	920	1058

Note: Correlations are Pearson correlations.

All correlations significant at the 0.001 level (2-tailed)

Binomial Logistic Regression for Drop-Out

A binomial logistic regression was performed on Drop-Out as the outcome and five variables of interest as the predictors: Physical Health, Mental Health, Positive Social Interaction, Negative Social Interaction and *AVPR1A RS3* genotype. This analysis was performed separately for male and female participants using data from Wave 1 for the predictor variables. A test of the full model with all five predictors against an intercept-only model was not found to be significant for the male (X^2 (6, N = 1102) = 9.408, p = 0.152) or female (X^2 (6, N = 1167) = 9.814, p = 0.133) group, indicating that as a set, the predictors could not reliably distinguish between participants who had completed all waves of data

collection and participants who had completed less than three waves of data collection. Examining the effects of individual parameters, a significant effect of Physical Health on Drop-Out was found for men and Positive Social Interaction had a significant effect on Drop-Out for women (Table 3.6). The parameter estimates suggested that men with poorer physical health were more likely to drop-out of the study and women with lower levels of Positive Social Interaction were more likely to drop-out of the study. However, the effect sizes for these differences are small, with odds ratios (OR) very close to 1, and are unlikely to significantly affect or bias the results of the subsequent analyses. The different pattern of results for men and women supports the separation of the male and female modelling analyses.

Table 3.6. Parameter estimates for logistic regression on participant drop out.

Drop out[a]	Gender	Variable	Parameter Estimate	Odds Ratio	95% Confidence Interval for OR	
					Lower Bound	Upper Bound
Data for < 3 Waves	Male	Intercept	-0.336			
		Physical Health	**-0.029**	**0.972**	**0.950**	**0.994**
		Mental Health	0.008	1.008	0.988	1.028
		Positive Social Interaction	-0.001	0.999	0.982	1.017
		Negative Social Interaction	0.005	1.005	0.987	1.024
		AVPR1A RS3 S/S[b]	-0.203	0.817	0.525	1.270
		AVPR1A RS3 S/L[b]	0.145	1.156	0.849	1.573
	Female	Intercept	0.045			
		Physical Health	-0.014	0.986	0.963	1.009
		Mental Health	0.010	1.010	0.990	1.030
		Positive Social Interaction	**-0.028**	**0.973**	**0.951**	**0.995**
		Negative Social Interaction	0.001	1.001	0.980	1.024
		AVPR1A RS3 S/S[b]	-0.352	0.703	0.427	1.159
		AVPR1A RS3 S/L[b]	-0.111	0.895	0.640	1.252

Note: a. Reference category is: Data for all 3 Waves

b. Reference category is: *AVPR1A RS3* L/L

Bolded results are significant at $\alpha = 0.05$, using a Wald test with 1 degree of freedom (df).

Repeated Cross-Sectional Analysis - Mental Health

Model 1. The fixed and random intercepts for this model are reported in Table 3.7 for women and Table 3.9 for men. Approximately 48.13% of the variance in Mental Health for men and 41.28% of the variance for women could be attributed to between person differences.

Model 2. *AVPR1A RS3* genotype was entered into the model as a fixed main effect. For women, genotype had a significant effect on Mental Health; individuals with the Short/Short or Short/Long genotype reported better Mental Health than those with the Long/Long genotype (Table 3.7). For men, individuals with the Short/Long genotype reported significantly worse Mental Health than those with the Long/Long genotype (Table 3.9). Mental Health did not differ between the male Short/Short and Long/Long genotype groups. *AVPR1A RS3* genotype was retained in the model.

Model 3. Time was entered into the model as a fixed main effect. Women (Table 3.7) and men (Table 3.9), reported that their Mental Health was worse at Wave 3 compared to Wave 1. This result was significant for women and just significant for men. The effects of *AVPR1A RS3* genotype were the same as for Model 2. Time was retained in the model.

Model 4. Positive Social Interaction and Negative Social Interaction were entered into the model as fixed main effects. For women (Table 3.7) and men (Table 3.9), Positive Social Interaction was positively associated with Mental Health and Negative Social Interaction was negatively associated with Mental Health. When controlling for Social Interaction, Time was not found to have a significant effect on women's Mental Health. This may be because women's Negative Social Interaction decreases over time. A paired-sample t-test found significant differences in women's Negative Social Interaction between Wave 1 and Wave 2 ($t(1110) = 7.235$, $p < 0.001$), and between Wave 2 and Wave 3 ($t(1023) = 4.607$, $p < 0.001$).

Conversely, the effect of the Short/Short genotype on Mental Health appeared stronger than in previous models (Table 3.7).

The opposite pattern was observed for men; when controlling for Social Interaction, *AVPR1A RS3* genotype did not have a significant effect on male Mental Health and the effect of Time increased. In Model 4, males reported worse Mental Health at both Wave 2 and Wave 3, compared to Wave 1 (Table 3.9). The change in the random intercept in Model 4, compared to previous models, indicates that Social Interaction explained the largest proportion of between-person differences in Mental Health for men and women (Table 3.7, Table 3.9). The information criteria for men and women suggested that, compared to the other variables in the model, *AVPR1A RS3* genotype had the greatest impact on improving model fit. All variables were retained in the model.

Interaction effects. Two interaction effects were tested in separate models:

1. *AVPR1A RS3* * Positive Social Interaction
2. *AVPR1A RS3* * Negative Social Interaction

A summary of the interaction effects and their impact on model fit can be found in Table 3.8 for women and Table 3.10 for men. No significant interaction effects were found for men. For women, a significant interaction effect was found between Negative Social Interaction and Short/Long genotype, and between Positive Social Interaction and Short/Long genotype when the interaction variables were entered into the model in "Positive only" and "Negative only" groups. The interaction between Social Interaction and *AVPR1A RS3* genotype suggested that women with the Short/Long genotype experience less negative effects from Negative Social Interaction and less positive effects from Positive Social Interaction, as the interaction effects were in the opposite direction of the Social Interaction main effects. The size of this effect was small and including interaction variables in the

model had little-to-no effect on model fit or, in the case of the BIC, worsened model fit. Thus, interaction effects were not included in the final model.

Best-fitting repeated cross-sectional mental health model. Model 4 had the best model fit of all the repeated cross-sectional mental health models according to information criteria. The model also included the main effects of all variables of interest. For men, Time, Negative Social Interaction and Positive Social Interaction were found to be significantly associated with Mental Health. Men reported worse Mental Health at Wave 2 and Wave 3, compared to Wave 1. High levels of Positive Social Interaction were associated with better Mental Health, and high levels of Negative Social Interaction were associated with poorer Mental Health. *AVPR1A RS3* genotype was not found to be significantly associated with male Mental Health. For women, similar effects for Positive Social Interaction and Negative Social Interaction were found; however, Time was not significantly associated with Mental Health, suggesting that on average, women's mental health did not change significantly between the three time-periods. *AVPR1A RS3* genotype was found to have a significant association with female Mental Health; women with the Short/Short and Short/Long genotypes both reported better Mental Health compared to women with the Long/Long genotype.

This page left intentionally blank.

Table 3.7. Female mental health models

	Parameter				95% Confidence Interval		Information Criteria		
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		45.967	<.001	45.758	46.147	24723.500	24729.500	24747.900
		Intercept	39.697	<.001	35.004	45.019			
		Residual	56.476	<.001	53.206	59.946			
Model 2	Intercept		45.449	<.001	45.107	45.756	23534.841	23544.841	23575.267
	AVPR1A RS3 S/S[a]		**0.860**	**.003**	**0.249**	**1.504**			
	AVPR1A RS3 S/L[a]		**0.927**	**<.001**	**0.494**	**1.352**			
		Intercept	39.320	<.001	34.564	44.730			
		Residual	56.052	<.001	52.734	59.579			
Model 3	Intercept		45.023	<.001	44.593	45.445	23527.577	23541.577	23584.174
	AVPR1A RS3 S/S[a]		**0.849**	**.003**	**0.242**	**1.480**			
	AVPR1A RS3 S/L[a]		**0.918**	**<.001**	**0.470**	**1.348**			
	Wave 2[b]		0.524	.052	0.026	1.036			
	Wave 3[b]		**0.867**	**.002**	**0.344**	**1.391**			
		Intercept	39.332	<.001	34.582	44.735			
		Residual	55.883	<.001	52.575	59.400			
Model 4	Intercept		45.631	<.001	46.548	52.960	23006.227	23024.227	23078.994
	AVPR1A RS3 S/S[a]		**1.058**	**<.001**	**0.496**	**1.675**			
	AVPR1A RS3 S/L[a]		**0.937**	**<.001**	**0.532**	**1.350**			
	Wave 2[b]		-0.155	.543	-0.602	0.352			
	Wave 3[b]		-0.132	.624	-0.582	0.390			

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Positive Social		**0.320**	**<.001**	**0.271**	**0.344**			
	Negative Social		**-0.326**	**<.001**	**-0.349**	**-0.281**			
		Intercept	24.402	<.001	20.933	28.445			
		Residual	52.015	<.001	48.939	55.285			
Model 5	Intercept		47.340	<.001	44.868	50.333	22862.495	22884.495	22951.371
	AVPR1A RS3 S/S[a]		**1.047**	**<.001**	**0.487**	**1.692**			
	AVPR1A RS3 S/L[a]		**0.867**	**<.001**	**0.444**	**1.289**			
	Wave 2[b]		-0.048	.854	-0.517	0.455			
	Wave 3[b]		-0.012	.964	-0.493	0.539			
	Positive Baseline		**0.342**	**<.001**	**0.305**	**0.371**			
	Positive Change		**0.296**	**<.001**	**0.239**	**0.331**			
	Negative Baseline		**-0.380**	**<.001**	**-0.411**	**-0.353**			
	Negative Change		**-0.274**	**<.001**	**-0.311**	**-0.225**			
		Intercept	24.552	<.001	21.091	28.580			
		Residual	51.624	<.001	48.571	54.869			

Note: a. Reference category is *AVPR1A RS3* L/L genotype.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except fixed intercepts). Information criteria are displayed in smaller-is-better form.

Table 3.8. Interaction effects for female repeated cross-sectional mental health model

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Random Intercept	24.402	<.001	20.933	28.445	23006.227	23024.227	23078.994
	Residual	52.015	<.001	48.939	55.285			
Positive	Positive * S/S	-0.044	.417	-0.146	0.072	23002.171	23024.171	23091.107
Social *	Positive * S/L[a]	**-0.097**	**.011**	**-0.164**	**-0.015**			
RS3	Random Intercept	24.291	<.001	20.830	28.328			
	Residual	51.995	<.001	48.919	55.264			
Negative	Negative * S/S	0.076	.134	-0.025	0.175	23002.249	23024.249	23091.186
Social *	Negative * S/L[b]	**0.088**	**.006**	**0.027**	**0.156**			
RS3	Random Intercept	24.379	<.001	20.915	28.416			
	Residual	51.946	<.001	48.875	55.211			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Social Interaction and Negative Social Interaction. Random effects and information criteria from Model 4 (Table 3.7) are presented for comparison.

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Fixed effects significant at the $p < 0.05$ level are bolded. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

a. The parameter estimate for the *AVPR1A RS3* S/L genotype is 5.860, 95% CI [1.671, 9.319], $p = 0.003$

b. The parameter estimate for the *AVPR1A RS3* S/L genotype is -3.493, 95% CI [-6.872, -0.365], $p = 0.032$.

This page left intentionally blank.

Table 3.9. Male mental health models

	Parameter				95% Confidence Interval		Information Criteria		
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		48.858	<.001	48.681	49.035	21782.697	21788.697	21806.793
		Intercept	40.593	<.001	36.018	45.749			
		Residual	43.751	<.001	41.070	46.607			
Model 2	Intercept		49.055	<.001	48.738	49.355	20794.318	20804.318	20834.244
	AVPR1A RS3 S/S[a]		-0.053	.835	-0.576	0.481			
	AVPR1A RS3 S/L[a]		**-0.466**	**.014**	**-0.870**	**-0.060**			
		Intercept	39.765	<.001	35.138	45.001			
		Residual	44.267	<.001	41.493	47.227			
Model 3	Intercept		49.356	<.001	48.951	49.736	20790.522	20804.522	20846.418
	AVPR1A RS3 S/S[a]		-0.051	.838	-0.556	0.472			
	AVPR1A RS3 S/L[a]		**-0.468**	**.013**	**-0.872**	**-0.067**			
	Wave 2[b]		-0.464	.064	-0.926	0.020			
	Wave 3[b]		**-0.541**	**.042**	**-1.056**	**-0.062**			
		Intercept	39.831	<.001	35.201	45.069			
		Residual	44.164	<.001	41.396	47.117			
Model 4	Intercept		49.755	<.001	47.145	52.237	20342.132	20360.132	20413.998
	AVPR1A RS3 S/S[a]		-0.206	.376	-0.706	0.336			
	AVPR1A RS3 S/L[a]		-0.204	.264	-0.600	0.190			
	Wave 2[b]		**-1.068**	**<.001**	**-1.482**	**-0.580**			
	Wave 3[b]		**-1.233**	**<.001**	**-1.690**	**-0.730**			
	Positive Social		**0.287**	**<.001**	**0.243**	**0.307**			

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Negative Social		**-0.280**	**<.001**	**-0.300**	**-0.237**			
		Intercept	25.244	<.001	21.826	29.197			
		Residual	41.818	<.001	39.188	44.625			
Model 5	Intercept		50.117	<.001	47.996	52.354	20255.015	20277.015	20342.818
	AVPR1A RS3 S/S[a]		-0.302	.210	-0.802	0.222			
	AVPR1A RS3 S/L[a]		-0.125	.485	-0.513	0.272			
	Wave 2[b]		**-0.961**	**<.001**	**-1.382**	**-0.482**			
	Wave 3[b]		**-1.087**	**<.001**	**-1.573**	**-0.600**			
	Positive Baseline		**0.326**	**<.001**	**0.296**	**0.351**			
	Positive Change		**0.245**	**<.001**	**0.193**	**0.279**			
	Negative Baseline		**-0.325**	**<.001**	**-0.350**	**-0.298**			
	Negative Change		**-0.221**	**<.001**	**-0.260**	**-0.171**			
		Intercept	25.087	<.001	21.716	28.982			
		Residual	41.446	<.001	38.846	44.221			

Note: a. Reference category is *AVPR1A RS3* L/L genotype.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except fixed intercepts). Information criteria are displayed in smaller-is-better form.

Table 3.10. Interaction effects for male repeated cross-sectional mental health model

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Random Intercept	25.244	<.001	21.826	29.197	20342.132	20360.132	20413.998
	Residual	41.818	<.001	39.188	44.625			
Positive	Positive * S/S	0.040	.410	-0.061	0.125	20338.568	20360.568	20426.404
Social *	Positive * S/L	0.040	.252	-0.019	0.117			
RS3	Random Intercept	25.247	<.001	21.825	29.206			
	Residual	41.798	<.001	39.167	44.606			
Negative	Negative * S/S	0.082	.094	-0.026	0.162	20341.140	20363.140	20428.977
Social *	Negative * S/L	-0.025	.481	-0.094	0.046			
RS3	Random Intercept	25.049	<.001	21.641	28.993			
	Residual	41.848	<.001	39.214	44.658			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Social Interaction and Negative Social Interaction. Random effects and information criteria from Model 4 (Table 3.9) are presented for comparison.

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

Repeated Cross-Sectional Analysis - Physical Health

Model 1. Approximately 36.54% of the variance in Physical Health could be attributed to between person differences for men (Table 3.13). For women, approximately 36.64% of the variance could be attributed to between person differences (Table 3.11).

Model 2. *AVPR1A RS3* genotype was added to the model as a fixed main effect. For women, there was a significant effect of *RS3* genotype on Physical Health; women with the Short/Short and Short/Long genotype reported significantly better Physical Health than women with the Long/Long genotype (Table 3.11). No genotype effects were found for men but including *RS3* genotype in the model substantially improved model fit (Table 3.13). *AVPR1A RS3* genotype was retained in the model for men and women.

Model 3. Time was entered into the model as a fixed main effect. For women (Table 3.11) and men (Table 3.13), Time had a significant effect on Physical Health. Participants reported significantly worse Physical Health at Wave 3, compared to Wave 1. Time was retained in the model.

Model 4. Positive Social Interaction and Negative Social Interaction were entered into the model as fixed main effects. For women (Table 3.11) and men (Table 3.13), Positive Social Interaction was significantly positively associated with Physical Health and Negative Social Interaction was significantly negatively associated with Physical Health. The effect of Negative Interaction appeared to be stronger than the effect of Positive Interaction. For women, including Social Interaction in the model appeared to strengthen then effect of Short/Short genotype and Time at Wave 3 (Table 3.11). For men, including Social Interaction in the model also appeared to increase the strength of the effect of Time at Wave 3. Compared to previous models, the change in the random intercept indicated that Social Interaction explained the largest proportion of between-person differences in Physical Health (Table 3.11,

3.13). The information criteria for men and women suggested that, compared to the other variables in the model, *AVPR1A RS3* genotype had the greatest impact on improving model fit. All variables were retained in the model.

Interaction effects. Two interaction effects were tested in separate models:

1. *AVPR1A RS3* * Positive Social Interaction
2. *AVPR1A RS3* * Negative Social Interaction

A summary of the interaction effects and their impact on model fit is provided in Table 3.12 for women and Table 3.14 for men; no significant interaction effects were found. Including the interaction effects in the model had little-to-no effect on model fit, and for most of the models tested, the AIC and BIC indicated that including an interaction effect slightly worsened model fit. Thus, interaction effects were not included in the final model.

Best-fitting repeated cross-sectional physical health model. Model 4 had the best model fit of all the repeated cross-sectional physical health models and included the main effects for all variables of interest. For men, Physical Health was significantly associated with Positive Social Interaction, Negative Social Interaction and Time. Higher levels of Positive Social Interaction were associated with better Physical Health and higher levels of Negative Social Interaction were associated with poorer Physical Health. Physical Health was also worse at Wave 3 compared to Wave 1, but there was no difference between Wave 2 and Wave 1. No *AVPR1A RS3* genotype effects were found for men. For women, the same pattern of results was found for Time, Positive Social Interaction and Negative Social Interaction. Significant *AVPR1A RS3* genotype effects were observed; women with a Short/Short or Short/Long genotype reported better Physical Health than women with a Long/Long genotype.

This page left intentionally blank.

Table 3.11. Female physical health models

	Parameter				95% Confidence Interval		Information Criteria		
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		50.588	<.001	50.255	50.921	23086.174	23092.174	23110.564
		Intercept	21.578	<.001	18.853	24.698			
		Residual	37.313	<.001	35.157	39.601			
Model 2	Intercept		50.163	<.001	49.895	50.454	21952.990	21962.990	21993.398
	AVPR1A RS3 S/S[a]		**0.790**	**<.001**	**0.292**	**1.252**			
	AVPR1A RS3 S/L[a]		**0.710**	**<.001**	**0.332**	**1.042**			
		Intercept	21.215	<.001	18.472	24.366			
		Residual	36.841	<.001	34.664	39.153			
Model 3	Intercept		50.347	<.001	50.020	50.710	21948.754	21962.754	22005.326
	AVPR1A RS3 S/S[a]		**0.795**	**<.001**	**0.283**	**1.276**			
	AVPR1A RS3 S/L[a]		**0.714**	**<.001**	**0.324**	**1.044**			
	Wave 2[b]		-0.092	.668	-0.507	0.301			
	Wave 3[b]		**-0.516**	**.023**	**-0.967**	**-0.126**			
		Intercept	21.267	<.001	18.522	24.419			
		Residual	36.753	<.001	34.582	39.061			
Model 4	Intercept		53.385	<.001	50.718	55.879	21813.236	21831.236	21885.972
	AVPR1A RS3 S/S[a]		**0.892**	**<.001**	**0.382**	**1.354**			
	AVPR1A RS3 S/L[a]		**0.721**	**<.001**	**0.335**	**1.058**			
	Wave 2[b]		-0.408	.059	-0.787	0.016			
	Wave 3[b]		**-0.997**	**<.001**	**-1.397**	**-0.548**			
	Positive Social		**0.103**	**<.001**	**0.059**	**0.125**			

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Negative Social		**-0.160**	**<.001**	**-0.176**	**-0.118**			
		Intercept	17.504	<.001	15.036	20.376			
		Residual	36.724	<.001	34.549	39.035			
Model 5	Intercept		52.902	<.001	50.472	55.047	21669.537	21691.537	21758.376
	AVPR1A RS3 S/S[a]		**0.875**	**<.001**	**0.366**	**1.346**			
	AVPR1A RS3 S/L[a]		**0.619**	**<.001**	**0.225**	**0.946**			
	Wave 2[b]		-0.323	.135	-0.722	0.083			
	Wave 3[b]		**-0.901**	**<.001**	**-1.324**	**-0.456**			
	Positive Baseline		**0.151**	**<.001**	**0.122**	**0.179**			
	Positive Change		**0.063**	**.005**	**0.015**	**0.098**			
	Negative Baseline		**-0.196**	**<.001**	**-0.217**	**-0.168**			
	Negative Change		**-0.120**	**<.001**	**-0.148**	**-0.074**			
		Intercept	17.319	<.001	14.885	20.152			
		Residual	36.514	<.001	34.354	38.809			

Note: a. Reference category is *AVPR1A RS3* L/L genotype.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except intercepts). Information criteria are displayed in smaller-is-better form.

Table 3.12. Interaction effects for female repeated cross-sectional physical health model

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Random Intercept	17.504	<.001	15.036	20.376	21813.236	21831.236	21885.972
	Residual	36.724	<.001	34.549	39.035			
Positive	Positive * S/S	-0.082	.085	-0.186	-0.004	21810.604	21832.604	21899.503
Social *	Positive * S/L[a]	-0.053	.131	-0.126	0.012			
RS3	Random Intercept	17.570	<.001	15.097	20.449			
	Residual	36.651	<.001	34.480	38.958			
Negative	Negative * S/S	-0.006	.900	-0.092	0.085	21812.941	21834.941	21901.841
Social *	Negative * S/L[b]	0.017	.565	-0.034	0.081			
RS3	Random Intercept	17.524	<.001	15.053	20.400			
	Residual	36.709	<.001	34.535	39.020			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Social Interaction and Negative Social Interaction. Random effects and information criteria from Model 4 (Table 3.11) are presented for comparison.

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

This page left intentionally blank.

Table 3.13. Male physical health models

	Parameter				95% Confidence Interval		Information Criteria		
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		52.169	<.001	52.018	52.288	19832.168	19838.168	19856.259
		Intercept	15.171	<.001	13.138	17.519			
		Residual	26.344	<.001	24.728	28.065			
Model 2	Intercept		52.248	<.001	52.014	52.481	18872.573	18882.573	18912.490
	AVPR1A RS3 S/S[a]		0.002	.992	-0.404	0.408			
	AVPR1A RS3 S/L[a]		-0.066	.647	-0.411	0.199			
		Intercept	15.101	<.001	13.040	17.489			
		Residual	25.810	<.001	24.189	27.541			
Model 3	Intercept		52.370	<.001	52.062	52.683	18865.842	18879.842	18921.726
	AVPR1A RS3 S/S[a]		0.005	.981	-0.395	0.416			
	AVPR1A RS3 S/L[a]		-0.069	.633	-0.412	0.207			
	Wave 2[b]		0.070	.714	-0.292	0.415			
	Wave 3[b]		**-0.503**	**.014**	**-0.899**	**-0.145**			
		Intercept	15.183	<.001	13.118	17.574			
		Residual	25.693	<.001	24.079	27.417			
Model 4	Intercept		54.540	<.001	52.693	56.327	18764.856	18782.856	18836.707
	AVPR1A RS3 S/S[a]		-0.070	.709	-0.467	0.342			
	AVPR1A RS3 S/L[a]		0.015	.922	-0.337	0.281			
	Wave 2[b]		-0.164	.383	-0.512	0.189			
	Wave 3[b]		**-0.816**	**<.001**	**-1.193**	**-0.444**			
	Positive Social		**0.079**	**<.001**	**0.049**	**0.093**			

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Negative Social		**-0.117**	**<.001**	**-0.132**	**-0.086**			
		Intercept	12.829	<.001	10.937	15.048			
		Residual	25.769	<.001	24.145	27.503			
Model 5	Intercept		54.012	<.001	52.408	55.416	18699.923	18721.923	18787.707
	AVPR1A RS3 S/S[a]		-0.142	.444	-0.563	0.264			
	AVPR1A RS3 S/L[a]		0.060	.683	-0.274	0.331			
	Wave 2[b]		-0.118	.526	-0.461	0.240			
	Wave 3[b]		**-0.747**	**.001**	**-1.138**	**-0.373**			
	Positive Baseline		**0.106**	**<.001**	**0.087**	**0.124**			
	Positive Change		**0.049**	**.002**	**0.014**	**0.071**			
	Negative Baseline		**-0.133**	**<.001**	**-0.149**	**-0.112**			
	Negative Change		**-0.091**	**<.001**	**-0.116**	**-0.053**			
		Intercept	12.897	<.001	11.009	15.109			
		Residual	25.655	<.001	24.039	27.379			

Note: a. Reference category is *AVPR1A RS3* L/L genotype.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except intercepts). Information criteria are displayed in smaller-is-better form.

Table 3.14. Interaction effects for male repeated cross-sectional physical health model

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Random Intercept	12.829	<.001	10.937	15.048	18764.856	18782.856	18836.707
	Residual	25.769	<.001	24.145	27.503			
Positive	Positive * S/S	0.005	.867	-0.059	0.066	18764.480	18786.480	18852.298
Social *	Positive * S/L	0.019	.415	-0.028	0.065			
RS3	Random Intercept	12.824	<.001	10.933	15.042			
	Residual	25.768	<.001	24.143	27.501			
Negative	Negative * S/S	0.058	.120	-0.026	0.110	18762.302	18784.302	18850.120
Social *	Negative * S/L	0.048	.047	-0.005	0.088			
RS3	Random Intercept	12.732	<.001	10.844	14.948			
	Residual	25.793	<.001	24.165	27.529			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Social Interaction and Negative Social Interaction. Random effects and information criteria from Model 4 (Table 3.13) are presented for comparison.

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

Analysis of Change in Social Interaction - Mental Health Models

Model 5. For this series of analyses, Models 1 to 3 were the same as those described for the repeated cross-sectional analysis, thus these steps are not described. This analysis tested one new model, Model 5, which included the fixed effects of *AVPR1A RS3* genotype, Time, between- and within-person differences in Positive Social Interaction and between- and within-person differences in Negative Social Interaction. Both Negative Social Interaction variables had a significant effect in women and men (Table 3.7, 3.9); higher reported Negative Social Interaction was associated with poorer Mental Health. This effect appeared to be stronger for between-person differences in Negative Social Interaction at baseline than for within-person change over time. Both Positive Social Interaction variables were significantly, positively associated with Mental Health in women and men (Table 3.7, 3.9), again this effect appeared to be slightly stronger for between-person differences than for within-person change.

As for Model 4, when social variables were controlled for *AVPR1A RS3* genotype had a significant effect on Mental Health for women (Table 3.7) and Time had a significant effect on Mental Health for men (Table 3.9).

Interaction effects. Four interaction variables were tested in the male and female models:

1. *AVPR1A RS3* * baseline Negative Social Interaction
2. *AVPR1A RS3* * baseline Positive Social Interaction
3. *AVPR1A RS3* * change in Negative Social Interaction
4. *AVPR1A RS3* * change in Positive Social Interaction

Due to the small sample size of the Short/Short genotype group the interaction terms were tested in small groups. Interactions 1 and 2 were tested in one model, and interactions 3

and 4 were tested in a second model. No significant interactions between within-person change in Social Interaction and *AVPR1A RS3* genotype were found for women (Table 3.15) or men (Table 3.16). This suggests that the effects of Social Interaction over time on Mental Health were consistent across genotype groups. For women, a significant interaction between *AVPR1A RS3* genotype and baseline Positive Social Interaction was found (Table 3.15). Women with the Short/Long genotype appeared to receive slightly less benefit from Positive Social Interaction on Mental Health than women with the Long/Long genotype. However, the effect of the interaction was small and including interaction terms in the model inflated the *AVPR1A RS3* Short/Long genotype coefficient (Table 3.15).

For men, a significant interaction was found between the Short/Short genotype and baseline Positive Social Interaction and the Short/Short genotype and baseline Negative Social Interaction (Table 3.16). The parameter estimates suggest that men with the Short/Short genotype experience more positive effects from Positive Social Interaction, and less negative effects of Negative Social Interaction on Mental Health compared to men with the Long/Long genotype. However, including interaction effects in the model substantially inflated the coefficient of the *AVPR1A RS3* Short/Short genotype, suggesting that these results may be due to the small size of the Short/Short genotype group (Table 3.16). Including the interaction terms in the model had little effect or a small negative effect on model fit, thus interaction terms were not included in the final male or female model.

Final mental health model. Model 5 was selected as the final mental health model as it had the best model fit and contained the key variables of interest. High levels of baseline Negative Social Interaction and an increase in Negative Interaction over time were associated with lower levels of Mental Health in men and women. Conversely, higher levels of baseline Positive Social Interaction and an increase in Positive Social Interaction over time were associated with higher levels of Mental Health. Results differed between men and women for

Time and *AVPR1A RS3* genotype. For men, Mental Health was found to significantly decrease over time but women's Mental Health remained stable across the three waves of data collection. In addition, *RS3* genotype had a significant main effect on Mental Health in women. Women with a Short/Short or Short/Long genotype reported significantly better Mental Health than women with a Long/Long genotype.

Table 3.15. Interaction effects for analysis of change in Social Interaction – Female mental health models

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 5	Random Intercept	24.552	<.001	39.300	45.567	22862.495	22884.495	22951.371
	Residual	51.624	<.001	31.640	36.057			
Baseline Social Interaction * *RS3*	Positive Baseline * S/S	-0.062	.188	-0.171	0.035	22858.530	22888.530	22979.724
	Positive Baseline * S/L[a]	**-0.095**	**.001**	**-0.157**	**-0.028**			
	Negative Baseline * S/S	0.027	.464	-0.053	0.113			
	Negative Baseline * S/L	0.038	.138	-0.016	0.096			
	Random Intercept	24.401	<.001	20.952	28.418			
	Residual	51.625	<.001	48.572	54.871			
Change in Social Interaction * *RS3*	Positive Change * S/S	0.037	.617	-0.108	0.179	22861.149	22891.149	22982.344
	Positive Change * S/L	-0.019	.692	-0.116	0.068			
	Negative Change * S/S	0.045	.482	-0.075	0.177			
	Negative Change * S/L	0.044	.326	-0.032	0.139			
	Random Intercept	24.641	<.001	21.172	28.677			
	Residual	51.550	<.001	48.500	54.791			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Baseline, Positive Change, Negative Baseline and Negative Change. Random effects and information criteria from Model 5 (Table 3.7) are presented for comparison.

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Results significant at the $p < 0.05$ level are bolded. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

a. The parameter estimate for the *AVPR1A RS3* S/L genotype is 3.731, 95% CI [-1.635, 8.670], $p = 0.109$.

Table 3.16. Interaction effects for analysis of change in Social Interaction – Male mental health models

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 5	Random Intercept	25.087	<.001	36.719	42.792	20255.015	20277.015	20342.818
	Residual	41.446	<.001	24.902	29.035			
Baseline Social Interaction * *RS3*	Positive Baseline * S/S[a]	**0.138**	**<.001**	**0.057**	**0.212**	20247.999	20277.999	20367.730
	Positive Baseline * S/L	-0.009	.722	-0.055	0.046			
	Negative Baseline * S/S[a]	**0.145**	**<.001**	**0.063**	**0.211**			
	Negative Baseline * S/L	-0.006	.816	-0.053	0.049			
	Random Intercept	24.752	<.001	21.403	28.626			
	Residual	41.481	<.001	38.877	44.260			
Change in Social Interaction * *RS3*	Positive Change * S/S	-0.025	.679	-0.137	0.096	20253.467	20283.467	20373.198
	Positive Change * S/L	0.037	.408	-0.044	0.130			
	Negative Change * S/S	0.011	.850	-0.104	0.133			
	Negative Change * S/L	-0.020	.685	-0.111	0.075			
	Random Intercept	25.123	<.001	21.750	29.020			
	Residual	41.401	<.001	38.803	44.173			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Baseline, Positive Change, Negative Baseline and Negative Change. Random effects and information criteria from Model 5 (Table 3.9) are presented for comparison. Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Results significant at the $p < 0.05$ level are bolded. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

a. The parameter estimate for the *AVPR1A RS3* S/S genotype is -14.562, 95% CI [-19.794, -8.051], $p = <.001$

Analysis of Change in Social Interaction - Physical Health Model

Model 5. For this series of analyses, Models 1 to 3 were the same as those described for the repeated cross-sectional analysis, thus these steps are not described. This analysis tested one new model, Model 5, which included the fixed effects of *AVPR1A RS3* genotype, Time, between- and within-person differences in Positive Social Interaction and between- and within-person differences in Negative Social Interaction. Both Negative Social Interaction variables had a significant effect in women and men (Table 3.11, 3.13); higher reported Negative Social Interaction was associated with poorer Physical Health. This effect appeared to be stronger for between-person differences in Negative Social Interaction at baseline than for within-person change over time. Both Positive Social Interaction variables were significantly, positively associated with Physical Health in women and men (Table 3.11, 3.13), again this effect appeared to be stronger for between-person differences than for within-person change.

As for Model 4, men and women reported significantly worse Physical Health at Wave 3, compared to Wave 1 (Table 3.11, 3.13). For women, *AVPR1A RS3* genotype had a significant effect on Physical Health; women with a Short/Short or Short/Long genotype reported significantly better health than women with the Long/Long genotype (Table 3.11). No significant *AVPR1A RS3* genotype effects were found for men.

Interaction effects. Four interaction variables were tested in the male and female models:

1. *AVPR1A RS3* * baseline Negative Social Interaction
2. *AVPR1A RS3* * baseline Positive Social Interaction
3. *AVPR1A RS3* * change in Negative Social Interaction
4. *AVPR1A RS3* * change in Positive Social Interaction

Due to the small sample size of the Short/Short genotype group the interaction terms were tested in small groups. Interactions 1 and 2 were tested in one model, and interactions 3 and 4 were tested in a second model. No significant interaction effects were found for women (Table 3.17) or men (Table 3.18) between *AVPR1A RS3* genotype and within-person changes in Social Interaction. This suggests that the effect of Social Interaction on Physical Health over time was consistent across genotype groups. For women, there were significant interactions between *AVPR1A RS3* genotype and baseline Positive Social Interaction, and genotype and baseline Negative Social Interaction (Table 3.17). Women with the Short/Long genotype experienced less positive effects from baseline Positive Social Support on Physical Health, compared to women with the Long/Long genotype. Women with a Short/Short or Short/Long genotype experienced more negative effects from baseline Negative Social Interaction on Physical Health, compared to women with a Long/Long genotype. However, these effects were small and including interaction effects in the model inflated the *AVPR1A RS3* genotype parameter estimates (Table 3.17).

For men, a significant interaction between *AVPR1A RS3* genotype and baseline Negative Social Interaction was found. Men with the Short/Short or Short/Long genotype experienced less negative effects from baseline Negative Social Interaction on Physical Health, compared to men with the Long/Long genotype. However, including interaction effects in the model substantially inflated the parameter estimates for *AVPR1A RS3* genotype main effects (Table 3.18). Overall, including interaction effects in the model had no effect, or a small negative effect on model fit for men (Table 3.18) and women (Table 3.17), thus interaction effects were not included in the final model.

Final physical health model. Model 5 was selected as the final physical health model as it had the best model fit according to the information criteria and contained all main effects of interest. Men and women reported significantly worse Physical Health at Wave 3,

compared to Wave 1 but there was no difference in physical health between Wave 1 and Wave 2. Higher Positive Social Interaction at baseline and an increase in Positive Social Interaction over time were associated with better Physical Health, and higher baseline Negative Social Interaction and an increase in Negative Social Interaction were associated with poorer Physical Health. In addition, there was a significant *AVPR1A RS3* genotype effect for women; women with a Short/Short or Short/Long genotype reported significantly better Physical Health than women with a Long/Long genotype.

Table 3.17. Fixed interaction effects for analysis of change in Social Interaction – Female physical health models

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 5	Random Intercept	17.319	<.001	27.044	32.771	21669.537	21691.537	21758.376
	Residual	36.514	<.001	22.009	25.876			
Baseline Social Interaction * *RS3*	Positive Baseline * S/S	-0.011	.790	-0.094	0.073	21666.211	21696.211	21787.354
	Positive Baseline * S/L[a]	**-0.066**	**.012**	**-0.111**	**-0.003**			
	Negative Baseline * S/S[b]	**-0.087**	**.005**	**-0.153**	**-0.017**			
	Negative Baseline * S/L[a]	**-0.053**	**.012**	**-0.095**	**-0.004**			
	Random Intercept	17.203	<.001	14.776	20.030			
	Residual	36.529	<.001	34.368	38.826			
Change in Social Interaction * *RS3*	Positive Change * S/S	-0.087	.143	-0.214	0.008	21666.491	21696.491	21787.634
	Positive Change * S/L	-0.017	.717	-0.117	0.053			
	Negative Change * S/S	0.018	.757	-0.099	0.114			
	Negative Change * S/L	0.050	.183	-0.024	0.122			
	Random Intercept	17.376	<.001	14.933	20.218			
	Residual	36.440	<.001	34.283	38.733			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Baseline, Positive Change, Negative Baseline and Negative Change. Random effects and information criteria from Model 5 (Table 3.11) are presented for comparison. Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Results significant at the $p < 0.05$ level are bolded. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

a. The parameter estimate for the *AVPR1A RS3* S/L genotype is 6.732, 95% CI [1.826, 10.281], $p = .001$.

b. The parameter estimate for the *AVPR1A RS3* S/S genotype is 5.956, 95% CI [-0.395, 12.027], $p = .037$.

Table 3.18. Fixed interaction effects for analysis of change in Social Interaction – Male physical health models

				95% Confidence Interval		Information Criteria		
Model	Parameter	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Main effects	Random Intercept	12.897	<.001	19.694	24.769	18699.923	18721.923	18787.707
	Residual	25.655	<.001	15.029	18.303			
Baseline Social Interaction * *RS3*	Positive Baseline * S/S	-0.007	.749	-0.060	0.037	18693.771	18723.771	18813.476
	Positive Baseline * S/L	0.012	.457	-0.022	0.052			
	Negative Baseline * S/S[a]	**0.116**	**<.001**	0.063	0.162			
	Negative Baseline * S/L[b]	**0.079**	**<.001**	0.039	0.111			
	Random Intercept	12.747	<.001	10.869	14.949			
	Residual	25.666	<.001	24.049	27.393			
Change in Social Interaction * *RS3*	Positive Change * S/S	0.032	.429	-0.051	0.104	18698.914	18728.914	18818.619
	Positive Change * S/L	0.021	.503	-0.042	0.078			
	Negative Change * S/S[a]	-0.035	.414	-0.112	0.054			
	Negative Change * S/L[b]	-0.002	.960	-0.055	0.069			
	Random Intercept	12.876	<.001	10.988	15.087			
	Residual	25.655	<.001	24.039	27.380			

Note: All models include the following fixed effects: *AVPR1A RS3* S/S, *AVPR1A RS3* S/L, Wave 2, Wave 3, Positive Baseline, Positive Change, Negative Baseline and Negative Change. Random effects and information criteria from Model 5 (Table 3.13) are presented for comparison.

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Fixed parameter estimates significant at the $p < 0.05$ level are bolded. The reference category for genotype is *AVPR1A RS3* L/L. Information criteria are displayed in a smaller-is-better format.

a. The parameter estimate for the *AVPR1A RS3* S/S genotype is -5.732, 95% CI [-9.171, -1.468], p = 0.001.

b. The parameter estimate for the *AVPR1A RS3* S/L genotype is -4.620, 95% CI [-7.496, -1.735], p <0.001.

Table 3.19. Summary of significant genetic main effects – Female Models

Model	Parameter	Estimate	Sig.	95% Confidence Interval	
				Lower Bound	Upper Bound
Outcome: Mental Health					
4: Repeated Cross-Sectional	*AVPR1A RS3* S/S	1.058	<.001	0.496	1.675
	AVPR1A RS3 S/L	0.937	<.001	0.532	1.350
5: Analysis of Change in Social Interaction	*AVPR1A RS3* S/S	1.047	<.001	0.487	1.692
	AVPR1A RS3 S/L	0.867	<.001	0.444	1.289
Outcome: Physical Health					
4: Repeated Cross-Sectional	*AVPR1A RS3* S/S	0.892	<.001	0.382	1.354
	AVPR1A RS3 S/L	0.721	<.001	0.335	1.058
5: Analysis of Change in Social Interaction	*AVPR1A RS3* S/S	0.875	<.001	0.366	1.346
	AVPR1A RS3 S/L	0.619	<.001	0.225	0.946

Notes: Bootstrapping used to produce 95% CIs around fixed effects. The reference category is *AVPR1A RS3* L/L genotype. Repeated cross-sectional models account for the effects of Wave, Positive Social Interaction and Negative Social Interaction. Analysis of change in Social Interaction models account for the effects of Wave, Positive Baseline, Positive Change, Negative Baseline and Negative Change.

Discussion

The aim of this project was to determine whether *AVPR1A RS3* genotype is associated with individual differences in physical and mental health, and whether genotype moderates the effect of social support and strain on health. *AVPR1A RS3* genotype was found to have a significant effect on women's self-reported mental and physical health. Women with one or two copies of the Short *RS3* allele reported significantly better health than women homozygous for the Long *RS3* allele. Based on previous research, it was hypothesised that the effect of *AVPR1A RS3* genotype would be strongest for male participants, however, no direct effect of genotype on mental or physical health was found for men. A summary of the significant genetic main effects can be found in Table 3.19. The results of this study also indicate that *AVPR1A RS3* genotype may moderate the impact of social interaction on mental and physical health for men and women, however, further research is required to confirm this. Together, these findings suggest that variation in the *AVPR1A* gene influences women's health and wellbeing but the mechanism for this effect remains unclear.

Evidence for Gene-Environment Interactions

The findings of this study provide some support for a gene-environment interaction between *AVPR1A RS3* genotype and social support and strain. However, the interpretations that can be drawn from the modelled interaction effects are restricted by the limitations of the study. There is some evidence that *AVPR1A RS3* genotype moderates the impact of positive and negative social interaction on mental and physical health. Moderation effects were found for women and men but were more consistently found across models for women. Overall, the pattern of results suggests that certain *AVPR1A RS3* genotypes, defined by relative allele length, could influence an individual's sensitivity to social interaction and thus moderate the impact of social interaction on mental and physical health. Allele length appears to have

different effects dependent on sex. For women, Short alleles appear to have a protective effect on health, whereas for men Long alleles may increase sensitivity to negative social stimuli. These results are in line with evidence from previous research which suggests that both *AVPR1A* genotype (Levin et al., 2009; Moons et al., 2014) and exogenous vasopressin administration (Shalev et al., 2011; Thompson et al., 2006) can affect an individual's response to social stimuli. No significant interaction effects were found for women or men between *AVPR1A RS3* genotype and within-person changes in social interaction, suggesting that the effect of within-person change in social interaction on physical and mental health over time was consistent across genotype groups.

Due to the small size of the Short/Short *RS3* genotype group, small related groups of interactions were tested in separate models, making it difficult to interpret the overall pattern of effects. Visualisations of the significant interaction effects indicated that the direction of the effects on health differed between models and the size of the moderation effect was small. Additionally, interaction effects were excluded from the final physical and mental health models because they failed to improve model fit. In the context of previous literature linking *AVPR1A* with individual differences in responses to social stimuli (Levin et al., 2009; Meyer-Lindenberg et al., 2009; Moons et al., 2014; Reuter et al., 2015; Vogel et al., 2012), the interaction effects in the present study suggest that *AVPR1A* variation may affect health by moderating the impact of social interaction. However, the true underlying mechanism linking *AVPR1A RS3* to women's mental and physical health remains unclear.

A Potential Mediation Model?

The relationship between *AVPR1A RS3* and health may instead be mediated by social interaction. *AVPR1A* variation is associated with differences in social behaviour (Avinun et al., 2011; Knafo et al., 2009; Wang et al., 2016) and the quality of social relationships (Avinun

et al., 2012; Bisceglia et al., 2012; Leerkes et al., 2017; Walum et al., 2008). However, in the present study there was no evidence of a direct association between *AVPR1A RS3* genotype and individual differences in positive or negative social interaction (Appendix A, Table A1 – A4). A functional SNP in the *AVPR1A* gene is associated with Behavioural Inhibition System (BIS) sensitivity in adults (Reuter et al., 2015); however, *AVPR1A RS3* genotype and BIS sensitivity were not associated in the present study (Appendix B, Table A5 – A6). BIS sensitivity did not differ between *RS3* genotype groups for men ($F(2, 1091) = 0.132$, $p = 0.876$) or women ($F(2, 1161) = 0.189$, $p = 0.828$). Lastly, research evidence indicates that *AVPR1A* variation may be particularly influential in spousal relationships (Maher et al., 2011; Walum et al., 2008). In the PATH study, the Dyadic Adjustment Scale-7 (DAS-7; Hunsley, Best, Lefebvre, & Vito, 2001) was administered in Wave 3 to measure the Dyadic Adjustment of participants currently in a relationship. Despite previous research suggesting that Dyadic Adjustment may be related to *AVPR1A RS3* genotype, no association was found between genotype and DAS-7 score in the present study (Appendix C, Table A7 – A8) for men ($F(2, 563) = 1.177$, $p = 0.309$) or women ($F(2, 702) = 0.303$, $p = 0.738$). The lack of an association may be due to the small sample of individuals with a partner in the 20+ age cohort. Overall, these results indicate that social interaction, dyadic adjustment and Behavioural Inhibition are unlikely to mediate the relationship between *AVPR1A RS3* genotype and mental and physical health. Instead, the relationship may be mediated by social factors that were not measured in the present study or through other biological pathways.

Peripheral Functions of Vasopressin

AVPR1A variation could impact on physical and mental health via the peripheral physiological functions of vasopressin. Vasopressin has multiple functional roles, including regulating water homeostasis by increasing renal water retention and elevating blood pressure via vasoconstriction (Kormos & Gaszner, 2013; Thibonnier et al., 2000). Vasopressin also

influences hypothalamic-pituitary-adrenal (HPA) axis function and has been implicated in the physiological response to chronic stress (Frank & Landgraf, 2008; Kormos & Gaszner, 2013). Vasopressin's role in stress physiology is likely primarily mediated by the V1-type receptors, including V1a which is expressed in the smooth muscle of blood vessels (Kormos & Gaszner, 2013; Thibonnier et al., 2000). It is possible that *AVPR1A RS3* genotype impacts on physical and mental health through these pathways. A previous study suggests that *AVPR1A* genotype is not associated with blood pressure (Thibonnier et al., 2000) but the association between the *RS3* polymorphism and blood pressure was not specifically tested. While the peripheral and higher-central actions of vasopressin are generally considered to be independent (Frank & Landgraf, 2008), evidence for the social and physiological functions of vasopressin supports the hypothesis that *AVPR1A* variation plays a role in the relationship between social interaction and health.

Social Interaction, Health and Wellbeing

The effects of positive and negative social interaction on mental and physical health found in the present study support findings reported in previous literature. Higher levels of positive social interaction were associated with higher reported levels of mental and physical health, in line with previous evidence for the relationship between social support and health (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016). Conversely, higher levels of negative social interaction were associated with lower reported levels of mental and physical health, supporting previous findings of a negative association between social strain and health (Schuster et al., 1990; Seeman et al., 2014; Yang et al., 2014). The significant association between social interaction and health was found for both men and women.

The effect of time on health differed between men and women for mental health outcomes but was similar for physical health outcomes. Women's reported mental health

levels did not significantly change across the three periods of data collection, whereas, on average, men's mental health declined over time. The effect of time on women's mental health may have been suppressed by a significant change in negative social interaction over time. By the third data collection period, reported physical health had declined for men and women, compared to their reported health at baseline. This suggests that by approximately age 30, participants were experiencing worse physical health compared to their health at approximately age 20.

Limitations

This study has two main limitations that qualify the conclusions that can be drawn from the results: sample size and allele categorisation. As discussed above, it was not possible to examine all gene-environment interaction effects in a single model due to the small sample size of the Short/Short genotype group. This makes it difficult to interpret the patterns of interaction effects that were found. Due to this limitation, and the failure of interaction effects to improve model fit, the main effects models are interpreted as the primary results of this study.

The full range of genetic variation available in the data was truncated by binning *AVPR1A RS3* alleles into two length categories. This categorisation was necessary to preserve power; it decreased the number of genotype categories in the analysis. *AVPR1A RS3* allele length was treated as a categorical variable because, at this time, the interaction of alleles at the *RS3* locus is too poorly understood to treat *RS3* allele length as a continuous variable. It is currently unclear whether there is a dominant or codominant relationship between *RS3* alleles. *RS3* is also a complex repeat, and it is unclear whether the lengths of the two repeat regions have differing effects on gene function (Prichard et al., 2007). While binning alleles into Short and Long categories sacrifices variation, it is currently the most common method for

dealing with these challenges. The other common approach is to examine the effect of the presence and absence of a target allele that has previously been found to be nominally over-transmitted in Autism Spectrum Disorders (Kim et al., 2002). This 334bp (270bp in the current study) allele has been subsequently associated with autistic traits in the general population and other social factors (Meyer-Lindenberg et al., 2009; Procyshyn et al., 2017; Walum et al., 2008).

In the present study, the 334bp allele corresponds to allele 20 and falls into the Long category. Models using the target allele categorisation strategy were examined for comparison to previous literature. No association between allele 20 and mental health was found for men or women (Appendix D, Table A9 – A10). An association was found for physical health; women carrying one or two copies of allele 20 reported significantly worse physical health than women carrying no copies of the allele (Appendix D, Table A11). Allele 20 was the most common allele in the Long length category in the present study and had the same effect on physical health as carrying two copies of a Long allele (i.e. poorer reported physical health). Thus, the two genotyping methods may be detecting the same effect. As the results of the relative length-based allele categories (Short/Short, Short/Long and Long/Long) were more consistent across analyses in the present study, this genotype categorisation method was used in all subsequent analyses reported in this thesis. Current evidence from animal (Hammock & Young, 2005), in vitro (Tansey et al., 2011) and human research (Knafo, Israel, et al., 2008; Vogel et al., 2012) suggests that relative allele length is an appropriate categorisation strategy, supporting this decision. An increased understanding of the specific relationship between allele length, allele interaction, and biological function at the *AVPR1A RS3* locus will facilitate a clearer understanding of the role of this variant in social behaviour and health in the future.

Future Directions

The direct association between *AVPR1A RS3* variation and women's mental and physical health is a novel result at the time of writing. It is important to replicate this result in an independent sample. The mechanism behind this association remains unclear and requires further investigation. Evidence from the literature suggests that the role of *AVPR1A RS3* in social behaviour is worthy of further exploration and may contribute to the link between *AVPR1A RS3* and health. The next chapter provides a report of a study that explored this possibility by investigating whether there is an association between *AVPR1A RS3* and empathy. Empathy is a trait that may impact a person's sensitivity to social stimuli, and the quality of social interactions and relationships (Chow et al., 2013; Tone & Tully, 2014).

Conclusions

AVPR1A RS3 is associated with women's mental and physical health. This finding supports the importance of understanding the role of the *AVPR1A* gene in human behaviour and cognition; variation in the promoter region of this gene appears to have consequences for health and wellbeing. A better understanding of the role of *AVPR1A* in behaviour and cognition may contribute to our understanding of how social support and strain impact on mental and physical health.

Chapter 4: Mind-reading receptors: Exploring the relationship between *AVPR1A RS3* genotype and individual differences in empathy

Chapter Summary

Chapter 3 investigated the relationship between *AVPR1A RS3* genotype, positive and negative social interaction, and health. Social interaction permeates almost every domain of daily life. Current evidence suggests that *AVPR1A RS3* may influence individual differences in empathy, a trait that is important for social functioning and may provide a link between *AVPR1A* and health outcomes. The aim of this study was to determine whether *AVPR1A RS3* genotype was associated with individual differences in self-reported tendency to empathise and with cognitive empathy skill. These associations were explored through population-level regression analyses and family-level twin-modelling analyses. Due to the limitations of the study, strong conclusions could not be made about the relationship between *AVPR1A* and empathy. The results suggested that *AVPR1A RS3* variation is associated with cognitive empathy skill in men. Men homozygous for longer *RS3* alleles displayed higher cognitive empathy skills than men homozygous for shorter *RS3* alleles. A similar relationship was observed for middle-aged and older women, but not for young women. Limited evidence also suggests that young women homozygous for the longer *RS3* allele have a greater tendency to empathise than women carrying at least one shorter allele. While the results of this study must be interpreted with caution, the significant findings align with previous research on the role of *AVPR1A RS3* in empathy-related behaviours. Combined with evidence from the literature, the results of this study suggest that the relationship between *AVPR1A RS3* and empathy is worthy of further investigation.

Mind-reading receptors: Exploring the relationship between *AVPR1A RS3* genotype and individual differences in empathy

Social interaction permeates almost every domain of everyday life, from the home to the classroom or workplace. As demonstrated in the previous chapter, social support and social strain can affect our mental and physical health (Ditzen & Heinrichs, 2014; Seeman et al., 2014; Yang et al., 2014). Difficulties with social interaction and communication play a role in many mental health disorders and can impact on day-to-day functioning (Tone & Tully, 2014). To address these challenges, it is important to understand human social behaviour and cognition, including their underlying biological basis. The *AVPR1A* gene has been associated with a variety of social behaviours (Albers, 2012; Hammock & Young, 2005; Insel, 2010) and, as described in the previous chapter, it is associated with physical and mental health in women. Current evidence suggests that *AVPR1A* variation may be associated with individual differences in empathy, a trait that is important for social functioning. This chapter explores whether there is an association between the *AVPR1A RS3* locus and individual differences in dispositional empathy and cognitive empathy ability.

Empathy

The definition of empathy used in the project is described in detail in Chapter 2. Briefly, empathy is defined as the ability to recognise and understand the emotions, thoughts and intentions of other people (Decety & Jackson, 2004). It is split into two domains; emotional empathy and cognitive empathy (Decety & Jackson, 2004; Singer, 2006). Emotional empathy refers to the ability of an individual to share or experience the emotions of another person upon observing their emotional state, to recognise that emotion state, and to be aware that it originated within the other, not within the self (de Vignemont & Singer, 2006; Decety & Jackson, 2004; Fan et al., 2011; Singer, 2006). Cognitive empathy, also referred to

as mentalising or Theory of Mind, encompasses the ability to represent and understand the mental states of other people, i.e. the ability to take another person's perspective (Decety & Jackson, 2004; Singer, 2006). Empathy is distinct from sympathy and from personal distress (Davis, 1980; de Vignemont & Singer, 2006; Decety & Jackson, 2004; Maibom, 2012).

Deficits in empathy abilities play a role in disorders featuring communication and social interaction difficulties. For example, autism spectrum disorders, schizophrenia and conduct disorder are associated with lower or disordered cognitive empathy (Baron-Cohen & Wheelwright, 2004; Baron-Cohen, Wheelwright, Hill, et al., 2001; Bons et al., 2013; Golimbet, Alfimova, Abramova, Kaleda, & Gritsenko, 2015). As discussed in Chapter 2, while empathy is usually protective for mental health, high empathic sensitivity may be a 'risky skill' (Tone & Tully, 2014). Being highly sensitive to the negative emotions of other people can be detrimental, leading an individual to experience a relationship partner's distress as if it were their own (Smith & Rose, 2011; Tone & Tully, 2014). This high sensitivity could lead to maladaptive self-focussed responses, including increased personal distress and excessive interpersonal guilt (Tone & Tully, 2014); cognitive and affective responses that have been linked to internalising disorders, social isolation, neuroticism and burnout (Smith & Rose, 2011). Thus, if *AVPR1A* is associated with individual differences in empathy, this trait may provide a link between *AVPR1A RS3* variation and health.

AVPR1A and Empathy

A recent study has found an association between *AVPR1A* variation and self-reported empathy, particularly cognitive empathy (Uzefovsky et al., 2015). An allele length previously associated with autism, 327bp (or 334bp depending on the genotyping method used), was found to predict lower total self-reported empathy and lower self-reported cognitive empathy, compared to individuals who did not carry the allele. The study included both male and

female participants and sex was found to predict self-reported empathy, but the interaction between genotype and sex was not examined (Uzefovsky et al., 2015). Total empathy was measured using a composite score from three measures: (1) the Interpersonal Reactivity Index (IRI; Davis, 1983), (2) the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004), and (3) the Questionnaire Measure of Emotional Empathy (QMEE; Mehrabian & Epstein, 1972). The cognitive and emotional domains of empathy were measured using the subscales of the IRI (Uzefovsky et al., 2015).

The association between *AVPR1A RS3* and self-reported empathy has not been replicated. To best understand the relationship between a candidate gene and a complex trait, it is important to amass evidence about that relationship (Tabor et al., 2002). It is also important to examine whether the relationship between *AVPR1A* and empathy varies by sex, given the frequently observed sex differences in self-reported empathy (Maibom, 2012) and in the effects of vasopressin and *AVPR1A* variation (Albers, 2012; McCall & Singer, 2012; Thompson et al., 2006; Walum et al., 2008). The sex differences observed for self-reported empathy are reduced or absent in skill-based assessments of empathy (Maibom, 2012). Thus, it is of interest to examine the association between *AVPR1A* and empathy skills, particularly cognitive empathy skills, given the findings of Uzefovsky and colleagues (2015).

AVPR1A and Empathy-related Behaviour

Although the association between *AVPR1A RS3* variation and self-reported empathy has not been replicated, current evidence suggests that *AVPR1A* is associated with empathy-related behaviours. These behaviours include altruism (Avinun et al., 2011; Knafo, Israel, et al., 2008; Wang et al., 2016) and maternal sensitivity (Avinun et al., 2012; Bisceglia et al., 2012; Leerkes et al., 2017). The association between *AVPR1A RS3* variation and these

behaviours supports the potential association between this locus and the underlying trait of empathy.

Altruism. Empathy can motivate altruistic behaviour. In a biological sense altruism is defined as a behaviour that increases the fitness of the recipient at the cost of the actor (de Waal, 2008). Empathy can make the distress of another person our own by reflecting that person's emotions within the self (de Waal, 2008; Decety & Jackson, 2004; Singer, 2006). We are thus motivated to engage in altruistic behaviour to relieve both their distress and our own (de Waal, 2008). Multiple studies support an association between *AVPR1A RS3* genotype and individual differences in altruistic behaviour. These studies provide indirect evidence for an association between *AVPR1A* variation and empathy.

AVPR1A RS3 genotype has been associated with altruistic behaviour in adults (Knafo, Israel, et al., 2008; Wang et al., 2016) and in children (Avinun et al., 2011). Altruistic behaviour can be measured using a protocol called the Dictator Game (a detailed description of the game can be found in Chapter 2). In this game, player one must allocated a sum of real money between themselves and an unknown second player (Forsythe et al., 1994). Differences in the amount of money player one allocates to player two is used as a measure of individual differences in altruism; the higher the amount allocated, the more altruistic the behaviour (Forsythe et al., 1994; Knafo, Israel, et al., 2008).

Adults with longer *AVPR1A RS3* alleles have been found to display more altruism in the Dictator Game compared to adults carrying shorter *RS3* alleles (Knafo, Israel, et al., 2008; Wang et al., 2016). That is, those with longer alleles allocate significantly more money to player two than those with short alleles. This association was initially found in a population-based analysis conducted with men and women (Knafo, Israel, et al., 2008). Participants homozygous for the long allele displayed significantly more altruistic behaviour in the

Dictator Game than participants homozygous for the short allele. Money allocation was also significantly associated with the *RS3* locus in a family-based analysis (Knafo, Israel, et al., 2008). A more recent population-based analysis replicated this association for men but not for women (Wang et al., 2016). Men with one or two copies of the long allele allocated more money to player two than men homozygous for the short allele.

AVPR1A RS3 is also associated with altruism in children. This association was investigated using a version of the Dictator Game adapted for children (Avinun et al., 2011). The child-version of the game is delivered in-person (rather than on a computer) and uses sticker sheets as a stimulus in place of money. Children are asked if they would give some of their sticker charts to another unknown child who had no sticker charts (Avinun et al., 2011). Male and female three-and-a-half-year-old twin-pairs participated in this version of the Dictator Game. In a population-based and a family-based analysis, children carrying at least one copy of the 334bp target allele were found to be less likely to donate at least one sticker sheet to another child, compared to non-carrier children (Avinun et al., 2011).

The variations in the *AVPR1A RS3* locus associated with altruism are also associated with neural differences between individuals. In post-mortem hippocampal tissue, samples from male and female individuals with longer *RS3* alleles were found to have higher levels of *AVPR1A* mRNA, compared to samples from individuals with shorter *RS3* alleles (Knafo, Israel, et al., 2008). Men carrying longer *AVPR1A RS3* alleles have also been found to have greater grey matter volume in the right fusiform face area (FFA), compared to men homozygous for shorter alleles (Wang et al., 2016). No genotype-group differences in grey matter volume were observed for women. Further, in men, right FFA grey matter volume mediated the relationship between *AVPR1A RS3* genotype and altruistic behaviour in the Dictator Game (Wang et al., 2016).

Together, these studies provide evidence for an association between *AVPR1A* variation and altruism. However, the precise relationship between *RS3* allele length and behaviour varies between the studies. In adults, longer *RS3* allele lengths were associated with more altruistic behaviour (i.e. the allocation of more money to player two) in the Dictator Game (Knafo, Israel, et al., 2008; Wang et al., 2016). In children, the 334bp target allele was associated with less altruistic behaviour in the Dictator Game (Avinun et al., 2011). The 334bp allele was included in the Long category of both adult studies (Knafo, Israel, et al., 2008; Wang et al., 2016). Thus, allele length appears to have different effects in adults compared to children. Avinun and colleagues (2011) suggest that this discrepancy could be due to the age and developmental stage differences between adult and pre-schooler participants. Cognitive and emotional empathy develop and change with age (Knafo et al., 2009; Knafo, Zahn-Waxler, et al., 2008), and as an empathy-motivated behaviour (de Waal, 2008) altruism is likely to differ between age groups as well.

The relationship between *AVPR1A RS3* and altruistic behaviour is further complicated by Wang and colleagues' (2016) findings that while longer *RS3* alleles were associated with greater grey matter volume in the right FFA compared to shorter *RS3* alleles, a post-hoc analysis suggested that the 334bp allele was associated with lower grey matter volume in this region. However, the presence or absence of this allele was not associated with differences in altruistic behaviour (Wang et al., 2016). This discrepancy may, in part, be due to issues with the accuracy of allele binning. *AVPR1A RS3* is a complex tandem repeat locus to genotype and the process is prone to error (Bakhtiari et al., 2018). Binning alleles into narrow categories is more likely to result in categorisation errors than binning them into broader categories, such as the Long/Short division typically utilised in the literature.

The studies also differ in the presence or absence of sex-specific effects. Wang and colleagues (2016) found significant differences in altruistic behaviour and right FFA grey

matter volume between *AVPR1A RS3* genotype groups for male participants but not for female participants. In contrast, Knafo, Israel and colleagues (2008) investigated the effects of sex in a bivariate ANOVA and found no main effect of sex and no interaction between sex and genotype. Sex differences in genotype effects were not tested in the study of altruism in children (Avinun et al., 2011). Given the reported sex differences in the effects of vasopressin and *AVPR1A* genotype in human and animal studies (Albers, 2012; McCall & Singer, 2012; Thompson et al., 2006; Walum et al., 2008), the potential for sex differences in the effect of genotype on altruism, and by extension empathy, warrants further investigation.

Despite some discrepancies in the associations found for specific participant groups, the evidence reviewed above supports an association between *AVPR1A* and individual differences in altruistic behaviour. Together, these studies indicate that longer *RS3* alleles are associated with increased altruistic behaviour in adult men (Knafo, Israel, et al., 2008; Wang et al., 2016), however, the direction or pattern of genotype effects may differ for women and children (Avinun et al., 2011; Wang et al., 2016). Further research is needed to explore this possibility. The association between *AVPR1A RS3* and altruism suggests that *AVPR1A* variation is also likely to be associated with individual differences in empathy.

Maternal sensitivity. Evidence of an association between *AVPR1A RS3* and empathy in women can be found in studies of the genetic correlates of maternal sensitivity. Maternal sensitivity is a mother's ability to accurately perceive and respond to her infant's cues promptly and appropriately (Ainsworth et al., 1978). This concept is similar to empathy and, as such, associations between maternal sensitivity and *AVPR1A* provide evidence for an association between this gene and empathy.

AVPR1A RS3 allele length is associated with observed maternal sensitivity and self-reported maternal cry-processing. Mothers homozygous for longer *RS3* alleles have been

found to display less maternal sensitivity (as rated by observers) than mothers carrying at least one copy of a shorter *RS3* allele (Bisceglia et al., 2012). Mother's carrying one or two copies of the 334bp allele have also been observed to display less sensitivity, compared to mothers not carrying this allele (Avinun et al., 2012). Specifically, mothers carrying the target allele were less likely to support their child's functioning during play through structuring (setting goals, preventing distractions) and gentle guidance (Avinun et al., 2012). A recent study failed to find a direct association between *AVPR1A RS3* allele length and observed maternal sensitivity, however, *RS3* genotype was associated with self-reported mother-oriented cry-processing (Leerkes et al., 2017). Compared to non-carriers, mothers carrying long *RS3* alleles were more likely to focus on their own needs and endorse negative cognitions about their infant's distress. *AVPR1A RS3* was also indirectly associated with maternal sensitivity via self-reported cry-processing (Leerkes et al., 2017).

The association between maternal sensitivity and *AVPR1A RS3* provides support for a link between *AVPR1A* variation and empathy in women, suggesting that women with longer *RS3* alleles may have lower levels of empathy. These results also support the hypothesis that *AVPR1A* variation may have sex-specific effects on empathy and related behaviours. In men, carrying longer *RS3* alleles was associated with more altruistic behaviour, indicative of higher empathy (Knafo, Israel, et al., 2008; Wang et al., 2016). In contrast, maternal sensitivity studies suggest that longer *RS3* alleles are associated with lower empathy in women (Bisceglia et al., 2012; Leerkes et al., 2017). A sex difference in the impact of *AVPR1A* genotype is in-line with the findings of the previous chapter (Chapter 3) and with the literature (Albers, 2012; McCall & Singer, 2012; Shalev et al., 2011; Walum et al., 2008). This highlights the importance of considering sex differences when investigating the role of vasopressin and *AVPR1A*.

Empathy is an important contributor to healthy social, emotional and moral development (Allemand et al., 2015; Tone & Tully, 2014). Individual differences in empathy may, in part, explain the link between *AVPR1A RS3* variation and health described in Chapter 3. One study has found an association between *AVPR1A RS3* and self-reported empathy, particularly cognitive empathy (Uzefovsky et al., 2015). Accumulating evidence about the behavioural associations of *AVPR1A* variation is key to understanding its function, and it is important to replicate the association between *AVPR1A RS3* and empathy (Tabor et al., 2002). It is also important to determine whether the relationship between self-reported empathy and genotype differs between men and women. Given the differences between self-reported tendency to empathise and empathy skills (Maibom, 2012), it is of interest to determine whether *AVPR1A* is associated with cognitive empathy accuracy in a similar manner to the previously reported association with self-reported cognitive empathy (Uzefovsky et al., 2015). This chapter describes a study that aimed to fill these gaps in knowledge, exploring the sex-specific associations between *AVPR1A RS3*, self-reported dispositional empathy and cognitive empathy skills.

Aim

The aim of this study was to determine whether *AVPR1A RS3* genotype was associated with individual differences in self-reported dispositional empathy and cognitive empathy skill. Potential associations were examined separately for male and female participants to determine whether there were sex-based differences in genotype effects. Based on existing literature and the results of the previous chapter, individuals with at least one Short *RS3* allele were hypothesised to have lower dispositional empathy and cognitive empathy abilities compared to individuals homozygous for the Long *RS3* allele. This hypothesis was tested using two techniques; a population-based regression analysis

comparing genotype groups in subsamples of unrelated individuals and a twin analysis of the measured empathy traits and their association with *AVPR1A RS3*.

Methods

Participants

Participants were a subset of the twins from the Brisbane Longitudinal Twin Study conducted at the Queensland Institute of Medical Research Berghofer Medical Research Institute (Wright & Martin, 2004). This subset of twins had completed the "Genetic and Environmental Foundations of Political and Economic Behaviours: A Panel Study of Twins and Families" sub-study (Hatemi, Smith, Alford, Martin, & Hibbing, 2015). The focus of that study (subsequently referred to as the Twin Study) was to measure economic, risk-taking, moral, and political behaviour, and explore genetic and environmental influences on their variation. Data was collected across two waves, which were completed approximately 12 months apart. Genotype data for a panel of 16 short tandem repeat (STR) loci was available for the participating twins and their parents and siblings (where present). Phenotypic data was collected from the twins at Wave 1 and Wave 2 but was only collected from parents and siblings at Wave 2.

Genetic Data

Genotyping. Participants were genotyped as described in Chapter 3[2]. Briefly, the *AVPR1A RS3* locus was genotyped as part of a custom fragment analysis multiplex assay of 16 short tandem repeat loci based on the methods described in Prichard and Easteal (2006). The *RS3* locus was amplified with primers: 5'-gctcaaaggcacactgttctc-3' (forward) and 5'-

[2] Genotyping was conducted by Susan Tan, John Curtin School of Medical Research, The Australian National University

gtcttgggaatctggtcagg-3' (reverse). The 16 loci were amplified by PCR and separated by capillary electrophoresis using an ABI 3730 gene analyzer (Biomolecular Resource Facility, John Curtin School of Medical Research, The Australian National University) using a multiplex design. The PCR program used was adapted for multiplex PCR of microsatellite loci: "denaturation: 94°C, 30s, annealing: 57°C, 90s, extension: 72°C, 60s for 28 cycles with an additional final step of 60°C, 30 minutes to minimize presence of artefact fragments" (Prichard & Easteal, 2006, p.3).

Allele and genotype frequencies. Hardy-Weinberg (HW) equilibrium and allele frequencies for the *AVPR1A RS3* locus were examined in a subsample of unrelated individuals. Using the Monte Carlo method from the ExactoHW software application (Engels, 2009) the locus was found to be in HW equilibrium. The allele labels are those suggested by Prichard and Easteal (2006) and approximately correspond to the number of repeats present. Allele frequencies are reported in Table 4.1. As described in Chapter 3, *AVPR1A RS3* contains two dinucleotide repeats and it was not possible to determine the relative contributions of each repeat motif to the fragment length (Prichard & Easteal, 2006). Allele frequencies were compared to those found in the 20+ cohort of the Personality and Total Health (PATH) Through Life study (Anstey et al., 2012) described in Chapter 3 (Table 4.1). Differences in allele frequencies between the two samples were tested with a Chi-square analysis in IBM SPSS version 22; no significant differences were found (X^2 (17, N = 6570) = 12.965, p = 0.739).

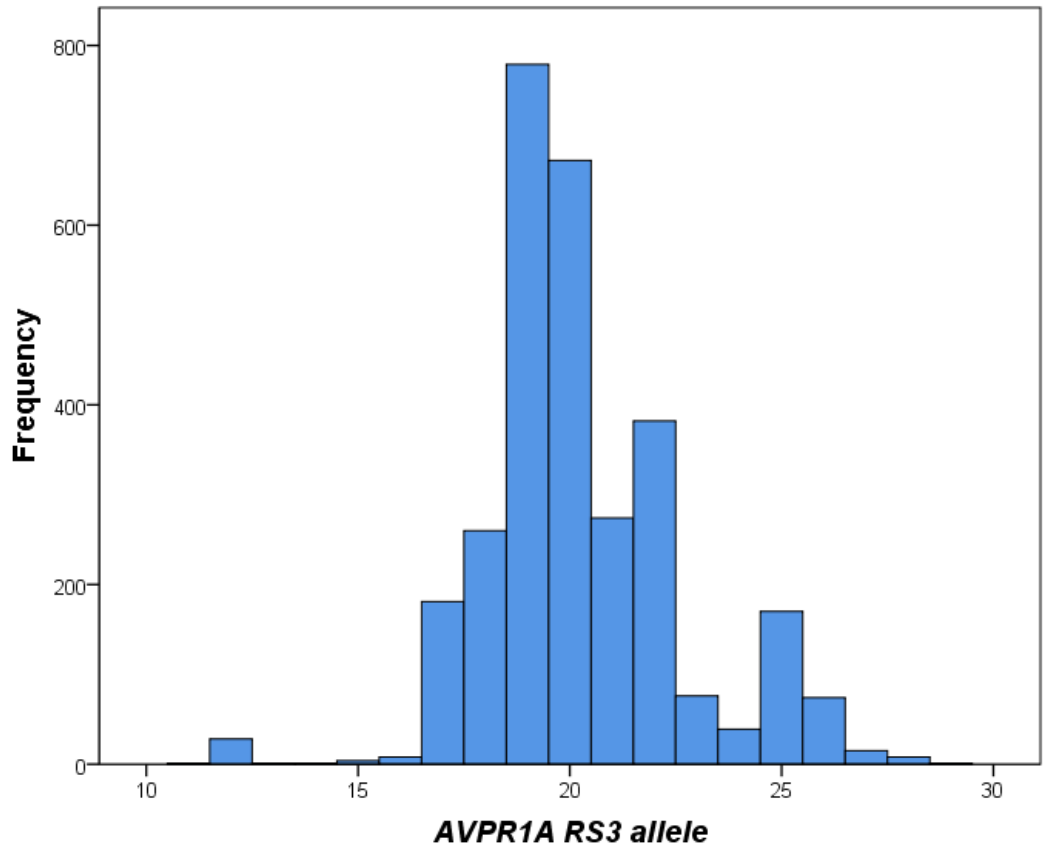
Table 4.1. *AVPR1A RS3* allele frequencies for the PATH 20+ cohort and the Twin Study.

Allele	PATH 20+	Twin Study
12	0.013	0.011
13	0.001	0.001
14	0	0.001
15	0	0.002
16	0.005	0.003
17	0.057	0.058
18	0.084	0.087
19	0.242	0.261
20	0.218	0.224
21	0.096	0.088
22	0.132	0.132
23	0.03	0.026
24	0.012	0.015
25	0.070	0.060
26	0.031	0.025
27	0.006	0.006
28	0.003	0.003
29	0	0.001

Note: Allele names represent the number of repeats present. Allele frequencies do not differ significantly between the studies (X^2 (17, N = 6570) = 12.965, p = 0.739).

Genotype categories. As discussed in Chapter 3, the *AVPR1A RS3* locus has a range of alleles lengths and alleles at the extreme ends of the distribution are present at low frequencies in the population (Figure 4.1). In the Twin Study sample, allele lengths ranged from 11 repeats to 28 repeats, with a median length of 20 repeats and a mode of 19 repeats. In previous literature, STR loci with similar allele distributions are commonly analysed by binning allele lengths into short and long groups (e.g. Knafo, Israel, et al., 2008). This improves the power of statistical analyses when there are multiple alleles of low frequency present in the population. Binning *AVPR1A RS3* alleles into short and long categories is also supported by evidence from animal and human studies suggesting that allele length is associated with differences in gene transcription (Hammock & Young, 2002; Hammock & Young, 2005; Knafo, Israel, et al., 2008; Tansey et al., 2011). As was done in Chapter 3, *AVPR1A RS3* alleles were binned into categories based on median repeat number. *RS3* allele lengths of 19 repeats or less were classified as Short and lengths of 20 repeats or more were classified as Long. Participants were then classified into three genotypes: Short/Short, Short/Long and Long/Long. For more complex twin analyses where it was necessary to further preserve power in the small sample size, participants were categorised into two genotype categories: Short Allele Present (Short/Short and Short/Long) and Short Allele Absent (Long/Long); i.e., dominance of the Short Allele. For analyses using three genotype categories, the Long/Long genotype was used as the reference category so that the effects of having one or two copies of the Short allele could be examined in a dose dependent manner. For analyses using two genotype categories, Short Allele Present was used as the reference category as this was the larger group.

Figure 4.1. *AVPR1A RS3* allele frequencies for the Twin Study data set.



Alleles are named by number of repeats present. Note the low-frequency allele lengths at the extreme ends of the distribution. The median length is 20 repeats and the modal length is 19 repeats.

Questionnaires

The Empathy Quotient. Self-reported dispositional empathy was measured using the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004). A short 18-item version of the EQ was included in the Wave 1 survey. This version contained 18 of the original 40 items and will subsequently be referred to as the EQ-18. Participants were asked to rate how well each item described them on a 5-point Likert scale ranging from 1 = very inaccurate to 5 = very accurate, with 3 representing a neutral option. This approach was a departure from the usual forced-choice scoring format of the scale (Baron-Cohen & Wheelwright, 2004). The

minimum possible score was 18 and the maximum was 90. Higher scores indicate a higher level of self-reported dispositional empathy. This version of the EQ had not been reported in the literature prior to the Twin Study, however, a psychometric analysis reported in this thesis demonstrated that the scale performs similarly to the original version and other short versions. See Chapter 5 for a report of the psychometric properties of the EQ-18.

The Reading the Mind in the Eyes Test. The Reading the Mind in the Eyes Test (RMET) is a task that aims to measure individual differences in cognitive empathy or Theory of Mind ability (Baron-Cohen, Wheelwright, Hill, et al., 2001). The RMET assesses an individual's ability to identify complex mental and emotional states from images of the eye region of the face. Response options are one-word descriptions of the mental state being displayed and participants must select the correct description from four multiple-choice options; one correct answer and three-foils (Baron-Cohen, Wheelwright, Hill, et al., 2001).

Wave 1. In Wave 1, a shortened 17-item version of the RMET was administered (subsequently referred to as RMET-17). Participants received 1 point for a correct answer and 0 points for an incorrect answer. The lowest possible score was 0 and the highest was 17, with a higher score indicating better performance on the test. This version of the test had not been reported in the literature prior to the Twin Study, however, a psychometric analysis reported in Chapter 5 suggests that it is performing similarly to the original version of the RMET.

Wave 2. Wave 2 included a 14-item version of the RMET (RMET-14), created by dropping three of the items used in Wave 1. Additionally, a "Don't know" response was added to the multiple-choice response options available. In this version of the test, participants received 1 point for a correct answer and 0 points for an incorrect answer or a "Don't know" response. The maximum possible score on the test was 14 and the minimum

possible score was 0, with a higher score indicating better performance on the task. This version of the test had not been reported in the literature prior to the Twin Study, however, a psychometric analysis reported in Chapter 5 suggests that it is performing similarly to the original version of the RMET and the 17-item version of the RMET, despite the inclusion of a new, previously unused response option.

Statistical Analyses

Software. Descriptive statistics, regression models and twin correlations were examined using IBM SPSS version 22. Assumption testing and twin modelling was conducted in the OpenMx package for R and R Studio (Boker et al., 2011).

Exploratory genetic associations. An exploratory population-level analysis of the relationship between *AVPR1A RS3* genotype and empathy measures was conducted in unrelated subsamples of participants. Analyses were conducted separately for men and women as performance on empathy measures and the effects of *AVPR1A* genotype and vasopressin have been found to vary by sex (Albers, 2012; Feng, Hackett, et al., 2015; Maibom, 2012; McCall & Singer, 2012; Walum et al., 2008). Three linear regression models were tested to determine whether empathy scores differed significantly between *AVPR1A RS3* genotype groups. The independent variable for all three models was *AVPR1A RS3* genotype. Genotype categories (Short/Short, Short/Long and Long/Long) were dummy-coded with Long/Long as the reference category. The dependent variable differed between the regression models; each model tested the effect of genotype on one of the three empathy measures: the EQ-18, RMET-17, or RMET-14. Participants with missing genetic data were excluded from these analyses. Participants were also excluded from an analysis if they were missing the relevant phenotypic data for that analysis.

AVPR1A RS3 genotype and EQ-18 score. Twin pairs were split to produce two unrelated subsamples: Twin 1 and Twin 2. The relationship between *AVPR1A RS3* genotype and EQ-18 score was initially carried out in the Twin 1 participant group and was subsequently replicated in the Twin 2 participant group. Analyses were conducted separately for male and female twins.

AVPR1A RS3 genotype and RMET-17 score. As for the EQ-18 model, the association between *AVPR1A RS3* genotype and RMET-17 score was initially tested in the Twin 1 participant group and was replicated in the Twin 2 participant group. Analyses were conducted separately for male and female twins.

AVPR1A RS3 genotype and RMET-14 score. The relationship between *AVPR1A RS3* genotype and RMET-14 was tested in four participant groups: (1) Twin 1 (male only), (2) Twin 2 (male only), (3) Mothers, and (4) Fathers. Female twins were excluded from this analysis because no significant association between *RS3* genotype and RMET-17 or EQ-18 score was found for this group. Despite the lack of association for female twins, the association between genotype and RMET-14 score was tested for Mothers because empathy and the heritability of empathy have been found to differ significantly with age (Tone & Tully, 2014).

Twin analysis. To determine whether scores on the RMET-17, RMET-14 and the EQ-18 were appropriate measures for twin modelling, twin correlations for each phenotype variable were examined and assumptions tests were performed.

Assumption testing. Assumptions tests were performed for each of the phenotype measures using the method described in Evans, Frazer, and Martin (1999). It was assumed that the twins participating in this study had been randomly sampled from the population; therefore, there should be no significant differences in the means or variances of phenotype

measures between zygosity groups (Evans et al., 1999). To establish the validity of this assumption, a series of hypotheses were tested to determine whether means and variances of the EQ-18, RMET-17 and RMET-14 could be equated across twin pairs, sex and zygosity. Hypotheses were tested in a series of nested models using Full Information Maximum Likelihood (FIML) estimation. Significance was determined by comparing the difference in –2 log likelihood (–2LL) against the preceding model using a X^2 statistic. As multiple hypotheses were tested, a conservative alpha level of 0.01 was used for all tests.

Twins were sorted into 5 zygosity groups:

(1) female monozygotic (MZ),

(2) male MZ,

(3) female dizygotic (DZ),

(4) male DZ,

(5) mixed DZ.

Each zygosity group had 2 sample means (one each for Twin 1 and Twin 2), 2 sample variances (one each for Twin 1 and Twin 2) and one sample covariance (between Twin 1 and Twin 2). A saturated base model was estimated from these 25 parameters.

The first sub-model, M1, equated means across Twin 1 and Twin 2. The second sub-model, M2, equated means across same-sex MZ and DZ twins, testing for homogeneity of means across zygosity groups of the same sex. M3 tested for homogeneity of means among all male twins and all female twins. M4 tested for homogeneity of means across sex; a significant loss of fit when male and female means were equated would suggest sex-based differences in performance (Evans et al., 1999). Sub-models V1 to V4 tested the same set of constraints for homogeneity of variance (Evans et al., 1999).

Models C1 through C3 tested for expected patterns of covariance among zygosities. If a trait is genetically influenced, it is expected that MZ correlations will be greater than DZ correlations (Evans et al., 1999; Evans, Gillespie, & Martin, 2002). Sex differences, or sex limitations, were also tested for. The C1 sub-model equated covariance across MZ twins and across same-sex DZ twins, testing for scalar sex limitation. C2 equated covariance across MZ twins and across all DZ twins, testing for non-scalar sex limitation. C3 equated covariance across MZ and DZ twins, testing whether MZ covariance and DZ covariance differ significantly. A significant loss of fit for sub-model C3 would mean that MZ and DZ correlations differ significantly. This would suggest that the trait being measured may be heritable (Evans et al., 1999).

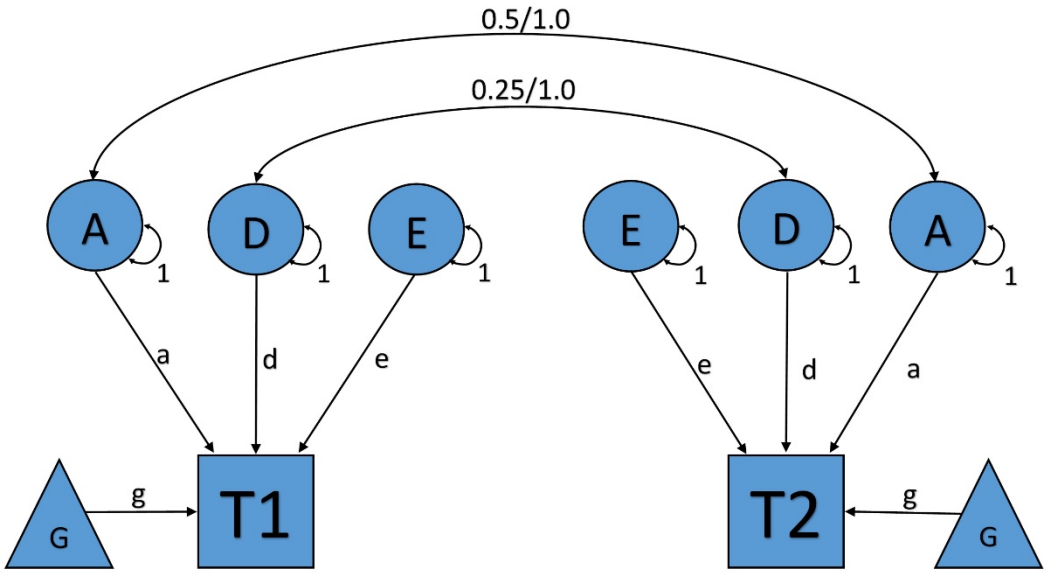
Twin Modelling. Based on the results of the twin correlations and assumption tests, the RMET-14 and RMET-17 were excluded from further analysis. All subsequent twin modelling analyses were carried out with EQ-18 as the sole empathy variable.

Twin models were constructed and tested using the structural equations modelling (SEM) approach described in Evans and colleagues (2002). Based on the results of the twin correlations and assumption tests, ADE models were constructed and tested with EQ-18 as the phenotype variable and *AVPR1A RS3* genotype included as a covariate. An ADE model was selected over an ACE model because the EQ-18 twin correlation for DZ twin pairs was not significant. If EQ-18 scores were influenced by common environment factors (C), DZ twin scores would be expected to significantly correlate.

Using a classical twin modelling approach, structural equation models were used to estimate the proportion of variance in the phenotype (EQ-18 score) that could be accounted for by variation in genetic factors (additive and dominant) and environmental factors (unique). The path diagram for the saturated ADE model is shown in Figure 4.2. The model

represents the linear relationship between two observed variables, EQ-18 score variation in Twin 1 and Twin 2, and three latent variables, additive genetic factors (A), dominance genetic factors (D), and unique environmental factors (E, including measurement error; Evans et al., 2002). The influence of *AVPR1A RS3* genotype on EQ-18 score was tested by modelling EQ-18 score means as a function of genotype (Neale, de Knijff, Havekes, & Boomsma, 2000). The latent variables were standardized and have a mean of zero and a standard deviation of 1. The path coefficients from the latent variables to the observed variables (a, d, e) and from the covariate to the observed variables (g) were estimated using FIML estimation.

Figure 4.2. The saturated ADE model.



The model contains three latent variables: additive genetic factors (A), dominance genetic factors (D), and unique environmental factors and measurement error (E). Correlations between latent variables are show for DZ/MZ twin pairs. The two observed variables represent EQ-18 score for Twin 1 (T1) and Twin 2 (T2). EQ-18 score means are modelled as a function of the *AVPR1A RS3* genotype covariate (G).

The correlations between the latent variables were determined by biometrical genetic theory (Evans et al., 2002). The additive genetic variance (A) represents the average effects of individual alleles across the genome (Griffiths et al., 2012). The additive genetic correlation between twin pairs is 1 for MZ twins (they share 100% of their segregating genetic material) and 0.5 for DZ twins (they share approximately 50% of their segregating genetic material; Evans et al., 2002; Griffiths et al., 2012). The dominance genetic variance (D) represents the interaction between alleles at the same locus, i.e. whether the effects of one allele display dominance over the effects of the other (Evans et al., 2002; Griffiths et al., 2012). The dominance genetic correlation between twin pairs is 1 for MZ twins (they share all their dominance variation) and 0.25 for DZ twins (they receive the same alleles from both parents 25% of the time; Evans et al., 2002; Griffiths et al., 2012). The unique environment variance (E) represents the environmental influences that are unique to each twin and incorporates measurement error. This variance is not correlated between twins (Evans et al., 2002; Griffiths et al., 2012).

The significance of the effect of the genotype covariate and the a, d, and e variance components was determined by comparing the goodness-of-fit of the saturated model to reduced, nested models where these components were individually removed. Significance was determined using log-likelihood ratio tests; the difference in –2LL between a reduced model and the saturated model was compared using a X^2 statistic. (Evans et al., 2002; Neale et al., 2000). The following nested models were compared to the saturated ADE model:

- AE model: Path d estimate set to 0
- DE model: Path a estimate set to 0
- E model: Path a and d estimates set to 0
- No Gene model: Path g estimate set to 0

Assumption testing identified a sex difference in EQ-18 means, so analyses were conducted separately for men and women. Due to the limitations of the OpenMx application, both members of a twin pair were required to have complete covariate data, i.e. genetic data, for the *AVPR1A RS3* locus. Participants were excluded from the analysis if they were missing *AVPR1A RS3* genotype data, or if their twin pair was missing genotype data. Individuals with missing phenotype data were included in the analysis as the FIML estimation was able to account for this missing data.

Results

Participants

A summary of the demographics of the total Twin Study sample is provided in Table 4.2. Sample sizes and other relevant information for analysis sub-groups are described with the results of the relevant analyses.

Table 4.2. Summary of participants in Wave 1 and Wave 2.

Wave 1		**Wave 2**	
Total Participants	*N* = 584 (338 Female; 246 Male)	**Total Participants**	*N* = 1641 (1004 Female; 637 Male)[a]
Average Age (years)	*M* = 25 *SD* = 3 Range = 19 − 31	**Average Age (years)**	*M* = 41 *SD* = 14 Range = 20 − 75
Complete MZ Twin Pairs	MZF: *n* = 61 MZM: *n* = 36	**Complete MZ Twin Pairs**	MZF: *n* = 61 MZM: *n* = 26
Complete DZ Twin Pairs	DZF: *n* = 55 DZM: *n* = 29 DZMF: *n* = 73	**Complete DZ Twin Pairs**	DZF: *n* = 33 DZM: *n* = 20 DZMF: *n* = 41
Participants with genetic data	*n* = 443	**Participants with genetic data**	*n* = 1394
Participants with complete phenotypic data (EQ-18 and RMET-17)	*n* = 578	**Participants with complete phenotypic data (RMET-14)**	*n* = 1495

[a]Includes: 365 fathers, 583 mothers, 124 siblings.

MZF = female monozygotic twins, MZM = male monozygotic twins, DZF = female dizygotic twins, DZM = male dizygotic twins, DZMF = mixed-sex dizygotic twins.

Questionnaires

The response rate and distribution of the empathy measures for the whole sample are reported below, for a detailed psychometric analysis of these measures see Chapter 5.

EQ-18. The EQ-18 was completed by 581 participants. Two participants were excluded from analysis as their attempts were judged to be non-serious. Additionally, one participant scored more than 3 standard deviations lower than the group mean, was classified as an outlier and excluded from analysis. This left a total sample of 578 participants (333 female; 245 male). The mean score on the EQ-18 was 66.654 (*SD* = 7.983), the minimum score was 41 and the maximum score was 86. The distribution of scores was approximately normal and scores were not transformed. EQ-18 scores differed significantly between men and women, with women scoring higher on average than men (Women: *M* = 68.090, *SD* = 7.556; Men: *M* = 64.702, *SD* = 8.147; $t(576) = 5.153$, $p = <0.001$; Cohen's $d = 0.43$).

RMET-17. In Wave 1, 580 participants completed the RMET. Two participants were excluded, as their attempts were non-serious, leaving a total of 578 participants (333 female; 245 male) in subsequent analyses. The mean RMET-17 score was 12.284 (*SD* = 2.259), the minimum score was 4 and the maximum score was 17. The distribution of scores was approximately normal with some negative skew present. Scores were not transformed. No sex differences were found for performance on the RMET-17 (Women: *M* = 12.34, *SD* = 2.32; Men: *M* = 12.21, *SD* = 2.18; $t(576) = 0.69$, $p = 0.491$).

RMET-14. Of the 1641 individuals who participated in Wave 2, 1495 (897 female; 598 male) completed all items of the RMET. The mean score on the RMET-14 was 9.029 (*SD* = 2.253), the minimum score was 0 and the maximum score was 14. The distribution of scores was approximately normal with some negative skew present. Scores were not transformed. A significant difference in performance was found between men and women,

with women scoring higher than men (Women: $M = 9.21$, $SD = 2.19$; Men: $M = 8.76$, $SD = 2.32$; $t(1493) = 3.85$, $p < 0.001$; Cohen's $d = 0.20$).

RMET-14: Family member differences. Means and variances for each family position group are presented in Table 4.3. As scores on empathy measures have been found to decrease with age in previous literature (Tone & Tully, 2014), twin scores on the RMET-14 were compared to parent scores. Performance on the RMET-14 had been found to differ significantly by sex, thus female twins were compared to Mothers and male twins were compared to Fathers. A one-way ANOVA indicated that there was a significant difference between the male groups ($F(2, 552) = 17.175$, $p < 0.001$). Post hoc comparisons using the Tukey honestly significant difference (HSD) test indicated that Fathers ($M = 8.249$, $SD = 2.394$) scored significantly lower that male Twin 1 ($M = 9.365$, $SD = 2.032$) and male Twin 2 ($M = 9.475$, $SD = 2.135$). There was no significant difference between male Twin 1 and Twin 2. Similar results were found for the female groups; a one-way ANOVA found a significant difference between groups ($F(2, 822) = 6.778$, $p = 0.001$). Mothers' mean RMET-14 score ($M = 8.957$, $SD = 2.241$) was found to be significantly lower than female Twin 1 ($M = 9.588$, $SD = 2.129$) but not lower than female Twin 2 ($M = 9.433$, $SD = 2.101$). As found for the male twins, there was no significant difference between female Twin 1 and Twin 2.

Table 4.3. Descriptive statistics for phenotypic variables, sorted by participant group.

	EQ-18			RMET-17			RMET-14		
Participant Group	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Twin 1 Male	65.513	7.921	160	12.213	2.294	160	9.365	2.032	137
Twin 1 Female	68.137	7.667	205	12.424	2.295	205	9.588	2.129	194
Twin 2 Male	63.177	8.392	85	12.200	1.944	85	9.475	2.135	80
Twin 2 Female	68.016	7.405	128	12.203	2.366	128	9.433	2.101	120
Fathers							8.249	2.394	338
Mothers							8.957	2.241	511

Exploratory Genetic Associations: Wave 1

Analyses were carried out using the Twin 1 subsample and then replicated in the Twin 2 subsample. Sample sizes for each *AVPR1A RS3* genotype category are presented in Table 4.4. Note that all participants included in the exploratory genetic association analyses were required to have complete genetic and phenotypic data.

***AVPR1A RS3* genotype and EQ-18 score.** No associations were found between *AVPR1A RS3* genotype and individual differences in EQ-18 score for men or women (Table 4.5).

Table 4.4. Sample size (*n*) for exploratory genetic associations, split by *AVPR1A RS3* genotype group.

	Wave 1			Wave 2		
Genotype Group	Short/Short	Short/Long	Long/Long	Short/Short	Short/Long	Long/Long
Twin 1 Male	25	58	33	24	59	29
Twin 1 Female	27	88	49	28	91	55
Twin 2 Male	11	33	19	10	32	19
Twin 2 Female	13	50	36	15	53	37
Fathers				59	159	111
Mothers				86	252	148

Note: All participants have complete genetic and phenotype data.

Table 4.5. Regression parameter estimates for association between *AVPR1A RS3* genotype and EQ-18 score.

					95% Confidence Interval	
Twin Group	Sex	Parameter	*B*	*p*-value	Upper Bound	Lower Bound
Twin 1	Female	Intercept	69.204	<0.001	67.034	71.374
		RS3 Short/Short	-0.982	0.595	-4.623	2.659
		RS3 Short/Long	-1.738	0.207	-4.446	0.970
	Male	Intercept	66.424	<0.001	63.770	69.079
		RS3 Short/Short	-3.104	0.131	-7.147	0.939
		RS3 Short/Long	-1.045	0.535	-4.370	2.280
Twin 2	Female	Intercept	69.639	<0.001	67.049	72.229
		RS3 Short/Short	-3.716	0.146	-8.744	1.312
		RS3 Short/Long	-1.639	0.341	-5.035	1.758
	Male	Intercept	63.000	<0.001	59.246	66.754
		RS3 Short/Short	-1.545	0.620	-7.745	4.654
		RS3 Short/Long	0.091	0.969	-4.622	4.803

Note: The reference category for all models is *AVPR1A RS3* Long/Long genotype, the intercept represents the mean of the Long/Long genotype group.

***AVPR1A RS3* genotype and RMET-17 score.** An association between *AVPR1A RS3* genotype and individual differences in cognitive empathy was found for the male Twin 1 sample (Table 4.6). Individuals with a Short/Short genotype performed significantly worse on the RMET-17 than individuals with a Long/Long genotype. However, this association was not replicated in the male Twin 2 sample (Table 4.6). Further, though not significant, the relationship between *AVPR1A RS3* genotype and RMET score appeared non-linear for Twin 2 (Figure 4.3). This difference in results may have been due to the smaller sample size in the Twin 2 group (Table 4.3). A retrospective power analysis indicated the power of the Twin 1 regression analysis was 0.90, whereas the Twin 2 sample had an estimated power of 0.64 to detect an effect of the same size ($R^2 = 0.103$). No association between RMET-17 and genotype was found for women (Table 4.6).

Table 4.6. Regression parameter estimates for association between *AVPR1A RS3* genotype and RMET-17 score.

Twin Group	Sex	Parameter	*B*	*p*-value	95% Confidence Interval	
					Upper Bound	Lower Bound
Twin 1	Female	Intercept	12.143	<0.001	11.476	12.809
		RS3 Short/Short	0.265	0.641	-0.854	1.383
		RS3 Short/Long	0.369	0.383	-0.463	1.200
	Male[a]	Intercept	12.606	<0.001	11.874	13.338
		***RS3* Short/Short**	**-1.886**	**0.001**	**-3.001**	**-0.771**
		RS3 Short/Long	-0.296	0.524	-1.213	0.621
Twin 2	Female	Intercept	12.611	<0.001	11.815	13.408
		RS3 Short/Short	-1.226	0.119	-2.773	0.320
		RS3 Short/Long	-0.551	0.298	-1.596	0.493
	Male	Intercept	12.842	<0.001	12.009	13.676
		RS3 Short/Short	-0.388	0.575	-1.764	0.989
		RS3 Short/Long	-0.933	0.080	-1.979	0.113

Note: The reference category is *AVPR1A RS3* Long/Long genotype, the intercept represents the mean of the Long/Long genotype group. Significant results are bolded.

[a] $R^2 = 0.103$ ($F(2, 113) = 6.474$, $p = 0.002$)

Figure 4.3. Relationship between *AVPR1A RS3* genotype and estimated marginal means for RMET-17 scores.

A comparison of the differences in RMET-17 score (estimated marginal means) between *AVPR1A RS3* genotype groups. While the relationship between RMET-17 and *RS3* genotype appears linear for Twin 1, in the Twin 2 sample the relationship may be non-linear.

Exploratory Genetic Associations: Wave 2

Analyses were carried out in the male Twin 1 subsample and then replicated in the male Twin 2 subsample. The analysis was then replicated using a sample of Fathers and Mothers. Sample sizes for each participant and *AVPR1A RS3* genotype group are presented in Table 4.3.

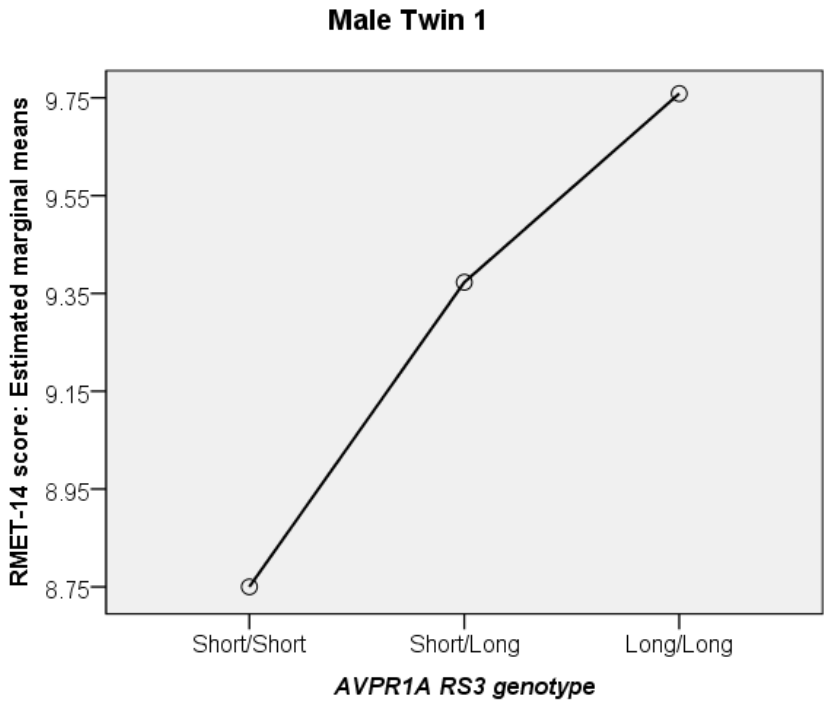
***AVPR1A RS3* genotype and RMET-14.** The association between *AVPR1A RS3* genotype and RMET-14 performance neared significance for Twin 1 and was not significant for Twin 2 (Table 4.7). For Twin 1, a negative trend was observed for the effect of *AVPR1A RS3* genotype on RMET-14 score, suggesting that individuals with the Short/Short genotype may score lower on average than individuals with the Long/Long genotype. Additionally, on visual inspection the relationship between RMET-14 score and *AVPR1A RS3* genotype for Twin 1 was similar to the relationship observed between RMET-17 score and *RS3* genotype for this group in Wave 1 (Figure 4.4).

The association between *AVPR1A RS3* genotype and RMET-14 score was significant for Fathers (Table 4.7); Fathers with a Short/Short genotype scored significantly lower on the test than Fathers with a Long/Long genotype. It should be noted that the five lowest-scoring individuals in the Father sample were in the Short/Short and Short/Long genotype groups. The two lowest-scoring individuals (receiving scores of 0 and 1) were in the Short/Short group and three individuals who received a score of 2 were in the Short/Long group. While these were the most extreme low scores, the overall distribution of scores is near-normal (Figure 4.5). Additionally, in the Father distribution, every possible score on the RMET-14 is represented, excluding a score of 14 (100% correct). As such, these individuals were kept in the analysis and scores were not transformed. It should also be noted that while the difference between the Short/Short and Long/Long genotype groups was found to be significant, the

regression equation explained a small amount of variation in RMET-14 score (approximately 2.1%).

Initially, no association between *AVPR1A RS3* genotype and RMET-14 score was found for Mothers (Table 4.7), supporting the results found for the female twin groups in Wave 1. However, the overall regression equation was found to be just significant ($F(2, 483) = 3.190$, $p = 0.042$), with an R^2 of 0.013. Visual inspection of the pattern of group means suggested that Mothers with a Short/Short genotype may significantly differ from Mothers with a Short/Long genotype (Figure 4.6). This difference was not tested by the parameter estimates in the regression model. The difference between these groups was found to be significant ($t(336) = -2.534$, $p = 0.012$), indicating that Mothers with a Short/Long genotype scored significantly higher on the RMET-14 than Mothers with a Short/Short genotype (Cohen's $d = 0.312$).

Figure 4.4. Relationship between *AVPR1A RS3* genotype and estimated marginal means for RMET-14 scores.



A comparison of the difference in RMET-14 score (estimated marginal means) between *AVPR1A RS3* genotype groups. While the relationship between RMET-14 and genotype appears linear for Twin 1, in the Twin 2 sample the relationship may be non-linear. The pattern of mean differences is similar to the pattern observed for these participants in Wave 1 (see Figure 4.3).

Table 4.7. Regression parameter estimates for the association between *AVPR1A RS3* genotype and RMET-14 score.

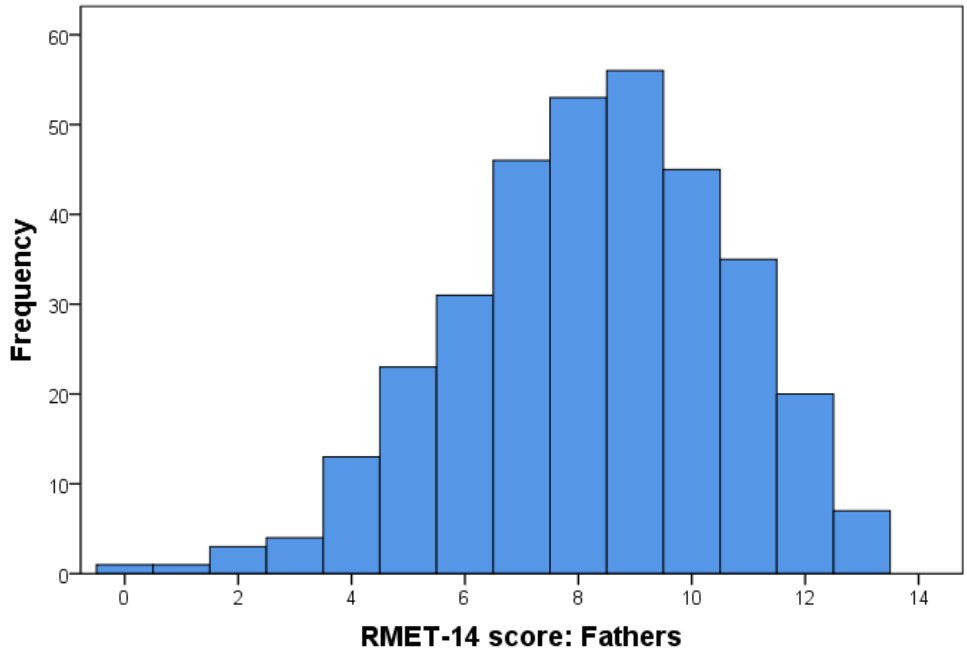
				95% Confidence Interval	
Twin Group	Parameter	*B*	*p*-value	Upper Bound	Lower Bound
Twin 1 (Male)	Intercept	9.759	0.001	9.033	10.484
	RS3 Short/Short	-1.009	0.066	-2.087	0.070
	RS3 Short/Long	-0.386	0.390	-1.272	0.500
Twin 2 (Male)	Intercept	9.632	<0.001	8.605	10.658
	RS3 Short/Short	0.068	0.938	-1.680	1.817
	RS3 Short/Long	-0.507	0.437	-1.803	.790
Fathers[a]	Intercept	8.613	<0.001	8.168	9.057
	***RS3* Short/Short**	**-1.019**	**0.008**	**-1.774**	**-.265**
	RS3 Short/Long	-.348	0.237	-.927	.230
Mothers[b]	Intercept	8.932	<0.001	8.572	9.293
	RS3 Short/Short	-0.444	0.143	-1.039	0.151
	RS3 Short/Long	0.254	0.273	-0.201	0.709

Note: The reference category is the Long/Long *AVPR1A RS3* genotype group, the intercept represents the mean of the Long/Long genotype group. Significant results are bolded.

[a] $R^2 = 0.021$ ($F(2, 326) = 3.536$, $p = 0.030$)

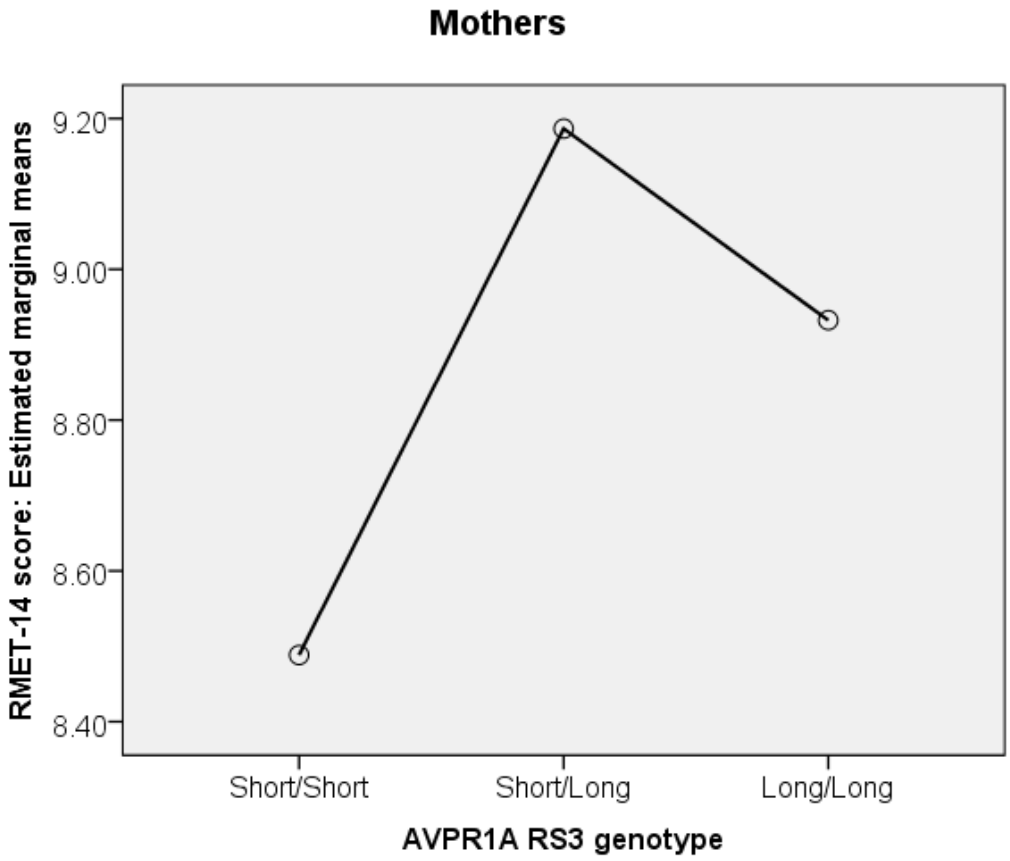
[b] $R^2 = 0.013$ ($F(2, 483) = 3.190$, $p = 0.042$)

Figure 4.5. Distribution of RMET-14 scores for Fathers



There is some negative skew present in the distribution, with a small number of individuals receiving a score between 0 and 3 points.

Figure 4.6. Relationship between *AVPR1A RS3* genotype and estimated marginal means for RMET-14 scores.



Estimated marginal means of RMET-14 score for each *AVPR1A RS3* genotype group in the Mothers sample. It appears that the greatest difference in means is between the Short/Short and Short/Long genotype groups. This is difference was likely to be driving the significance of the overall regression model (see Table 4.7).

Allele frequencies: Male twin group comparisons. To explore the difference in results between the male Twin 1 and Twin 2 groups, *AVPR1A RS3* allele frequencies were compared between the groups. No significant differences were found (X^2 (12, N = 360) = 7.05, p = 0.85). Allele frequencies were also compared between the Short/Long and Short/Short *AVPR1A RS3* genotype groups. No differences were found in allele frequencies between the Twin 1 and Twin 2 Short/Long genotype groups (X^2 (10, N = 182) = 4.75, p = 0.91), or between the Short/Short groups (X^2 (3, N = 72) = 2.97, p = 0.40). However, these results must be interpreted with caution as the Chi Square analysis contained multiple categories with $n < 5$, which reduces the reliability of the results.

Twin Correlations: EQ-18, RMET-17 and RMET-14

EQ-18. The EQ-18 was included in the twin analyses despite the lack of a genetic association in the population-level analysis, as it was possible that an association between EQ-18 score and *AVPR1A RS3* genotype may be apparent when familial factors were accounted for. A significant correlation was found between Twin 1 and Twin 2 EQ-18 scores for male and female MZ twins (Table 4.8). The correlation between male twins appeared to be stronger than the correlation between female twins. No significant correlations were found between DZ twins (Table 4.8). The correlation between female DZ twins was negative but non-significant. The greater similarity of MZ twins, compared to DZ twins suggests that genetic factors may contribute to individual differences in empathy (as measured by the EQ) and that the EQ may be a suitable candidate for twin modelling. An assumption test was conducted to confirm this (see below).

RMET-17. Significant correlations between Twin 1 and Twin 2 RMET-17 scores were found for female MZ twins and male DZ twins (Table 4.9). No other significant correlations were found. The correlation between female DZ twins was negative but non-

significant. This pattern of results could indicate that individual differences in RMET-17 score are influenced by genetic factors in women, but not in men. However, men and women are not typically found to score differently on the RMET (Harkness et al., 2010; Olderbak et al., 2015; Söderstrand & Almkvist, 2012; Spek, Scholte, & Berckelaer-Onnes, 2010). Overall, these results suggest that the RMET-17 is not an appropriate candidate for twin modelling. An assumption test was conducted to confirm this (see below).

RMET-14. The pattern of twin correlations differed between the RMET-17 and the RMET-14. For the RMET-14, significant correlations between Twin 1 and Twin 2 were found for female DZ twins and female/male DZ twins (Table 4.10). The twin correlation for female MZ twins also neared significance ($r = 0.27$, $p = 0.065$). The correlation between male DZ twins was negative but this result was non-significant and near zero ($r = -0.06$, $p = 0.800$). Combined with the RMET-17 correlations, these results suggest that the RMET is not an appropriate measure for twin modelling. An assumption test was conducted to confirm these results.

Table 4.8. Descriptive statistics and twin correlations EQ-18

					Twin Correlation	
Twin Group		Mean	*SD*	*N*	*r* (*n*)	*p*
Female MZ	Twin 1	65.43	7.12	65		
	Twin 2	63.93	8.42	68	**0.34 (60)**	**0.007**
Male MZ	Twin 1	60.61	8.61	46		
	Twin 2	63.27	8.35	41	**0.57 (35)**	**<0.001**
Female DZ	Twin 1	64.49	7.72	59		
	Twin 2	65.11	7.51	61	-0.12 (54)	0.407
Male DZ	Twin 1	59.46	8.28	39		
	Twin 2	62.53	7.09	34	0.14 (29)	0.478
Female/Male DZ	Twin 1	65.06	7.86	81		
	Twin 2	61.94	8.23	85	0.04 (71)	0.746

Note: Significant correlations are bolded.

Coefficients are Pearson correlation coefficients.

Table 4.9. Descriptive statistics and twin correlations RMET-17

					Twin Correlation	
Twin Group		Mean	*SD*	*N*	*r* (*n*)	*p*
Female MZ	Twin 1	12.09	2.62	65		
	Twin 2	11.81	2.26	67	**0.34 (60)**	**0.008**
Male MZ	Twin 1	12.43	2.03	46		
	Twin 2	11.80	2.29	41	0.26 (35)	0.137
Female DZ	Twin 1	12.68	2.08	59		
	Twin 2	12.64	2.42	61	-0.08 (54)	0.574
Male DZ	Twin 1	11.92	1.83	39		
	Twin 2	12.82	2.01	34	**0.50 (29)**	**0.005**
Female/Male DZ	Twin 1	12.51	2.16	81		
	Twin 2	12.61	2.37	85	0.03 (71)	0.823

Note: significant correlations are bolded.

Coefficients are Pearson correlation coefficients.

Table 4.10. Descriptive statistics and twin correlations RMET-14

					Twin Correlation	
Twin Group		Mean	*SD*	*N*	*r* (*n*)	*p*
Female MZ	Twin 1	9.67	2.25	76		
	Twin 2	9.32	2.07	72	0.27 (49)	0.065
Male MZ	Twin 1	9.61	2.39	44		
	Twin 2	9.24	1.95	49	0.30 (25)	0.143
Female DZ	Twin 1	10.10	2.00	41		
	Twin 2	9.60	2.15	48	**0.40 (27)**	**0.041**
Male DZ	Twin 1	9.31	1.79	36		
	Twin 2	9.07	1.98	27	-0.06 (18)	0.800
Female/Male DZ	Twin 1	9.23	2.03	77		
	Twin 2	9.59	2.12	61	**0.37 (39)**	**0.021**

Note: Significant correlations are bolded.
Coefficients are Pearson correlation coefficients.

Assumption Testing: EQ-18, RMET-17 and RMET-14

For the EQ-18 and RMET-17, 332 twin pairs (complete and incomplete) were included in the assumption tests. For the RMET-14, 388 twin pairs (complete and incomplete) were included in the assumption tests. Zygosity frequencies and sample sizes for each empathy measure are presented in Table 4.11.

Table 4.11. Sample sizes for the assumption tests.

			EQ-18		RMET-17		RMET-14	
Zygosity	No. Twin Pairs EG1	No. Twin Pairs EG2	Individuals with Data	Individuals Missing	Individuals with Data	Individuals Missing	Individuals with Data	Individuals Missing
MZF	74	105	133	15	132	16	148	62
MZM	52	68	87	17	87	17	93	43
DZF	66	67	120	12	120	12	89	45
DZM	44	47	73	15	73	15	63	31
DZOF	51	53	84	18	84	18	70	36
DZOM	45	48	82	8	82	8	68	28
Total	332	388	579	85	578	86	531	245

Note: Both complete and incomplete twin-pairs are included in the zygosity frequency counts. The substantial increase in missing phenotype data in Wave 2 is partially due to an increase in incomplete twin-pairs. MZF = female monozygotic twins, MZM = male monozygotic twins, DZF = female dizygotic twins, DZM = male dizygotic twins, DZOF = mixed-sex dizygotic twins with first-born female, DZOM = mixed-sex dizygotic twins with first-born male.

EQ-18. Scores on the EQ-18 fulfilled the assumptions of homogenous means and variances across zygosities and across sex (Table 4.12). A sex limitation was found for the EQ-18; equating female and male means resulted in a significant loss of model fit. This limitation was accounted for in subsequent analyses. Female and male variances were homogenous. MZ and DZ correlations differed significantly, indicated by a significant loss of model fit when these correlations were equated (Table 4.12). This indicates that the trait measured by the EQ-18 may be heritable.

Table 4.12. Assumptions test for EQ-18

Model	-2LL	df	Δ-2LL	Δdf	p
Base	4012.63	554			
M1	4022.24	558	9.61	4	0.048
M2	4022.78	560	0.55	2	0.761
M3	4029.70	562	6.91	2	0.032
M4	**4037.35**	**563**	**7.65**	**1**	**0.006**
V1	4033.83	566	4.13	4	0.388
V2	4034.60	568	0.78	2	0.679
V3	4035.18	570	0.58	2	0.750
V4	4035.64	571	0.46	1	0.497
C1	4037.59	573	1.95	2	0.377
C2	4037.59	574	0.00	1	0.959
C3	**4048.73**	**575**	**11.14**	**1**	**0.001**

Note: Models are nested and each model is compared to the preceding model, except for V1, which is compared to M3 due to the significant reduction in model fit for M4. Due to the number of comparisons, a conservative alpha level of 0.01 is used for significance testing, significant changes in fit are bolded.

RMET-17. Scores on the RMET-17 fulfilled the assumptions of homogenous means and variances across zygosities and sex, and no sex limitations were found. However, no significant differences between MZ and DZ correlations were found (Table 4.13). This indicates that twin similarity in RMET scores does not differ between MZ and DZ twin-pairs and suggests that the trait measured by the RMET is not heritable. This result could also be due to the low internal consistency and reliability of the RMET-17 (see Chapter 5).

Table 4.13. Assumptions test for RMET-17

Model	-2LL	df	Δ-2LL	Δdf	p
Base	2541.01	553			
M1	2550.99	557	9.98	4	0.041
M2	2556.68	559	5.68	2	0.058
M3	2556.88	561	0.20	2	0.904
M4	2557.21	562	0.34	1	0.563
V1	2560.90	566	3.69	4	0.450
V2	2562.80	568	1.90	2	0.387
V3	2564.40	570	1.61	2	0.448
V4	2566.21	571	1.81	1	0.178
C1	2571.40	573	5.18	2	0.075
C2	2571.54	574	0.15	1	0.700
C3	2575.61	575	4.06	1	0.044

Note: Models are nested and each model is compared to the preceding model. Due to the number of comparisons, a conservative alpha level of 0.01 is used for significance testing.

RMET-14. As for the RMET-17, scores on the RMET-14 fulfilled the assumptions of homogenous means and variances across zygosities and sex. No sex limitations were found and no significant difference between MZ and DZ correlations were found (Table 4.14). This supports the conclusion that either the trait measured by the RMET is not heritable, or the reliability of the RMET-14 is too low to be an appropriate measure for twin modelling.

Table 4.14. Assumptions test for RMET-14

Model	-2LL	df	Δ-2LL	Δdf	p
Base	2262.88	506			
M1	2268.31	510	5.43	4	0.245
M2	2270.07	512	1.76	2	0.416
M3	2270.53	514	0.46	2	0.795
M4	2272.2	515	1.67	1	0.196
V1	2275.12	519	2.92	4	0.571
V2	2276.70	521	1.90	2	0.455
V3	2278.07	523	1.37	2	0.503
V4	2278.08	524	0.01	1	0.953
C1	2279.10	526	1.03	2	0.598
C2	2279.57	527	0.47	1	0.494
C3	2279.62	528	0.05	1	0.821

Note: Models are nested and each model is compared to the preceding model. Due to the number of comparisons, a conservative alpha level of 0.01 is used for significance testing.

Twin Models: EQ-18 ADE

A total of 51 complete male twin pairs (33 MZ, 18 DZ) and 91 complete female twin pairs (50 MZ, 41 DZ) with no missing genetic data were included in the twin-modelling analysis.

Male ADE model. An ADE model with an *AVPR1A RS3* genotype covariate was constructed. The saturated ADE model suggested that variation in empathy is primarily explained by dominance genetic factors and unique environment factors (including measurement error). The variation in empathy accounted for by additive genetic variance was near 0 (Table 4.15).

Removing the 'a' or 'd' path from the saturated model did not result in a significant loss of model fit (Table 4.15). When the 'd' path was removed, the dominance variance was accounted for by the additive latent variable (a), without any significant loss in fit. Removing both the 'a' and 'd' paths from the model resulted in the largest increase in -2LL but this difference did not reach significance when compared to the saturated model ($p = 0.074$). Reduced models (AE and DE) were not compared to the E model, as both additive and dominant genetic effects were non-zero in the saturated model. Removing the *AVPR1A RS3* genotype path from the model also did not result in a significant loss of fit (Table 4.15). However, it did result in an increase in the amount of empathy variation accounted for by dominance genetics factors, suggesting that *AVPR1A RS3* may have been contributing to dominance genetic influences on the trait. Overall these results suggest that variation in empathy (or EQ-18 score) is primarily due to unique environmental factors and measurement error. However, it should be noted that 51 twin-pairs is a small sample size for twin analysis and may not have provided the power necessary to detect genetic effects.

Table 4.15. Male ADE model EQ-18

	Model				
Parameter	ADE	AE	DE	E	No Gene
μ	62.520	62.544	62.517	62.509	63.650
g	3.317	3.369	3.311	3.562	[0.000]
a	1.600 (0.199)	5.054 (0.627)	[0.000]	[0.000]	0.000
d	4.833 (0.601)	[0.000]	5.093 (0.633)	[0.000]	5.364 (0.654)
e	6.232 (0.774)	6.281 (0.779)	6.229 (0.774)	8.088 (1.000)	6.205 (0.757)
a^2	0.040	0.393	[0.000]	[0.000]	0.000
d^2	0.361	[0.000]	0.401	[0.000]	0.428
e^2	0.600	0.607	0.599	1	0.572
σ^2	64.759				
ΔX^2		0.061	0.001	5.212	2.518
Δdf		1	1	2	1
p		0.804	0.979	0.074	0.113

Note: Estimates are rounded to 3 decimal places. Unstandardised estimates are reported (with standardised estimates in brackets when available). All reduced models are compared to the saturated (ADE) model.

μ is the grand mean for a model. The symbols, g, a, d and e refer to the path coefficients for *RS3* genotype, additive genetic effects, dominance genetic effects and unique environment/error, respectively. The symbols a^2, d^2, e^2 refer to the proportion of the total variance in EQ-18 score accounted for by additive genetic sources, dominance genetic sources and unique environment/error sources, respectively. σ^2 is the total variance in the model. ΔX^2 is the difference in -2 log likelihood fit compared to the saturated ADE model.

Female ADE model. An ADE model with an *AVPR1A RS3* genotype covariate was constructed. Variance in empathy was explained by dominance genetic factors and unique environmental factors (Table 4.16). Additive genetic factors did not contribute any variance to the saturated model and dropping the 'a' path from the model did not result in a loss of model fit (Table 4.16). When the 'd' path was dropped from the model, dominance variance was accounted for by the additive latent variable and the 'a' path coefficient became negative. This did not result in a significant loss of fit (Table 4.16). Dropping both the 'a' and 'd' paths from the model resulted in the largest increase in -2LL, however, this difference was not significant (p = 0.104, df = 2). As the 'a' variable explained zero variance in the saturated model, the E model was compared to the reduced DE model. Removing path 'd' from the DE model resulted in a significant loss of fit (ΔX^2 = 4.520, df = 1, p = 0.034). Removing the *AVPR1A RS3* genotype path from the model resulted in a just significant loss of model fit at the 0.05 alpha-level (ΔX^2 = 3.856, df = 1, p = 0.050). The path estimate for genotype suggests that individuals with the Long/Long genotype report higher levels of empathy than individuals with at least one copy of the Short allele. Removing the *RS3* genotype path also resulted in an increase in the amount of variation in EQ-18 score accounted for by dominance genetic factors. This suggests that *AVPR1A RS3* variation may contribute to the dominance genetic influence on empathy.

Table 4.16. Female ADE model EQ-18

	Model				
Parameter	ADE	AE	DE	E	No Gene
μ	67.877	67.878	67.877	67.787	68.695
g	2.572	2.548	2.571	2.821	[0.000]
a	0.000	-3.634 (-0.520)	[0.000]	[0.000]	0.000
d	4.008 (0.573)	[0.000]	4.008 (0.573)	[0.000]	4.324 (0.608)
e	5.735 (0.820)	5.978 (0.855)	5.735 (0.820)	6.988 (1.000)	5.649 (0.794)
a^2	0.000	0.270	[0.000]	[0.000]	0.000
d^2	0.328	[0.000]	0.328	[0.000]	0.370
e^2	0.672	0.730	0.672	1	0.631
σ^2	48.955				
ΔX^2		1.083	0	4.520	3.856
Δdf		1	1	2	1
p		0.298	1	0.104	0.050

Note: Estimates are rounded to 3 decimal places. Unstandardised estimates are reported (with standardised estimates in brackets when available). All reduced models are compared to the saturated (ADE) model.

μ is the grand mean for a model. The symbols, g, a, d and e refer to the path coefficients for *RS3* genotype, additive genetic effects, dominance genetic effects and unique environment/error, respectively. The symbols a^2, d^2, e^2 refer to the proportion of the total variance in EQ-18 score accounted for by additive genetic sources, shared environmental sources and unique environment/error sources, respectively. σ^2 is the total variance in the model. ΔX^2 *is* the difference in -2 log likelihood fit compared to the saturated ADE model.

Discussion

The aim of this study was to determine whether *AVPR1A RS3* variation was associated with individual differences in self-reported empathy and cognitive empathy skill. Due to the limitations of the data set, it is not possible to comment conclusively on the effect of *AVPR1A RS3*. However, the results provide some support for an association between *RS3* genotype and empathy, though specific findings differ between sex and age groups. The results suggest that *AVPR1A RS3* variation is associated with differences in cognitive empathy skill in men. Men homozygous for the shorter *RS3* allele may have lower cognitive empathy skills than men homozygous for the longer *RS3* allele. However, this effect was not consistently replicated in all male participant groups. A similar relationship may also occur in middle-aged and older women, but not young women. There is limited evidence to support an association between self-reported empathy and *AVPR1A RS3* genotype in young women, suggesting that women homozygous for the longer *RS3* allele report a greater tendency to empathise than women with at least on copy of the shorter allele. While the evidence from this study must be interpreted cautiously, the significant findings align with previous research on empathy-related behaviours. Combined with evidence from the literature, the results of this study suggest that the relationship between *AVPR1A RS3* and empathy is worthy of further investigation in future research.

The results of this study partially support the findings of Uzefovsky and colleagues (2015); *AVPR1A RS3* variation was associated with empathy, and was most consistently associated with cognitive empathy. However, unlike Uzefovsky and colleagues (2015) the current findings were impacted by age and sex, and a different relationship between allele length and empathy was found. The previous study found that a target allele, 327bp, was associated with lower levels of self-reported cognitive empathy (Uzefovsky et al., 2015). In the present study, this allele was included in the Long category and was associated with

higher levels of cognitive empathy skills. This discrepancy may be due to the allele-binning methods used, the difference in empathy measures, or differences in the age and sex of participants included in the two studies.

The findings of the present study do align with evidence for the relationship between *AVPR1A RS3* and altruistic behaviour. Previous research has found that adults carrying longer *AVPR1A RS3* alleles display higher levels of altruistic behaviour during the Dictator Game compared to adults carrying shorter *AVPR1A RS3* alleles (Knafo, Israel, et al., 2008; Wang et al., 2016). As would be expected, the present study found that longer *RS3* alleles were associated with higher cognitive empathy skills in men and were potentially also associated with higher dispositional empathy in women.

Longer *AVPR1A RS3* alleles have also been associated with lower levels of maternal sensitivity (Avinun et al., 2012; Bisceglia et al., 2012; Leerkes et al., 2017) and poorer self-reported mental and physical health in women (Chapter 3). In combination with the results of the present study, this may support the theory that empathy can be a risky skill (Tone & Tully, 2014). Being highly sensitive to the negative emotions of other people could lead to higher levels of personal distress for an individual (Davis, 1980; Tone & Tully, 2014). In turn, personal distress may lead to more self-focussed responses to the distress of other people (e.g. mother-oriented cry processing in response to infant distress; Leerkes et al., 2017) and may result in higher stress levels, resulting in poorer mental and physical health outcomes (Tone & Tully, 2014). More research is required to confirm and understand the role of *AVPR1A* in women's empathy, social behaviour, and mental and physical health.

Limitations

The interpretations that can be made from the results of this study are limited by several factors. The study may not have had sufficient power to consistently detect small

genotype effects. Total sample size was limited by the data set used and genotype group sizes were limited by the relatively uncommon Short/Short genotype group. To account for this limitation, simple population-level analyses were performed, including only essential independent variables. The impact of sample size was most marked in the twin analysis. Due to the limitations of the OpenMx application, which required all participants to have complete covariate data, the sample was substantially reduced to only include complete twin pairs with genetic data.

The results may have also been influenced by the reliability of the short-form empathy measures used. The psychometric properties of these measures have not been previously reported in the literature, and the results of the twin correlations and assumption tests suggest that the short-form RMET measures may be unreliable. A psychometric analysis of these scales is reported in Chapter 5. Additionally, the short-form measures may have truncated the possible range of scores, reducing potential differences between genotype groups. Lastly, the results of this study should be considered exploratory. Multiple comparisons were made, and significance levels were not adjusted for multiple testing, thus, some significant findings may be due to Type I error. Several results were near-significant or nominal and would not have survived corrections for multiple testing. Replication is required to draw strong conclusions about the role of *AVPR1A RS3* in empathy.

Future Directions

Although the conclusions that can be drawn from this study are limited, they do not rule out a relationship between *AVPR1A RS3* variation and individual differences in empathy. The results are in alignment with previous research linking *AVPR1A* with empathy-related behaviours and make sense in the context of the findings reported in Chapter 3. In the context of evidence from the literature, the results of this study suggest that future research should

continue to investigate the potential relationship between *AVPR1A RS3* and empathy. To ensure success future studies will require adequate power; the effects of a single gene on a complex trait are likely to be small. It will also be important to use valid and reliable measures of both self-reported tendency to empathise and empathy ability. Self-report measures should allow for a clear distinction between the cognitive and emotional domains of empathy (e.g. the Basic Empathy Scale; Jolliffe & Farrington, 2006). A test of cognitive empathy ability, the RMET, was included in the present study. Emotional empathy could be tested by using physiological responses, such as heart rate or electrodermal (skin) conductance, to indicate emotional arousal in response to empathy-provoking stimuli (Blair, 1999; Blair, Jones, Clark, & Smith, 1997; Eisenberg & Fabes, 1990; Eisenberg et al., 1988). The next chapter (Chapter 5) investigates the psychometric properties of the short-form measures used in this study and assesses whether they are appropriate for use in future research. Lastly, future studies should account for the effects of participant age and sex, on both empathy and the potential effects of *AVPR1A RS3* genotype. Investigating the role of *AVPR1A RS3* variation in empathy may improve our understanding of empathy-based social and communication difficulties, and of the relationship between *AVPR1A RS3,* social interaction and health.

Conclusion

Strong conclusions about the relationship between *AVPR1A RS3* and individual differences in empathy cannot be drawn from the findings of this study. However, in the context of previous research they provide some evidence for an association between *AVPR1A RS3* and empathy, both self-reported dispositional empathy and cognitive empathy skills. To better understand this relationship, future research will require large samples, well-validated empathy measures and an acknowledgement of the potential impacts of age and sex.

Chapter 5: Psychometric properties of three short-form measures of empathy in an Australian sample

Chapter Summary

Chapter 4 explored the relationship between *AVPR1A RS3* genotype and individual differences in empathy using three short-form measures. The psychometric properties of these short-form measures have not been reported in the literature. Accurate and efficient measures of empathy may assist researchers to understand the trait's role in social behaviour and health, and to understand its biological basis. For example, valid short-forms of common empathy measures may facilitate the measurement of empathy in the large-scale studies required to explore complex genetic associations. In the present study, the psychometric properties of one short version of the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004) and two short forms of the Reading the Mind in the Eyes Test (RMET; Baron-Cohen, Wheelwright, Hill, et al., 2001) were examined. The properties of the short-form measures were compared to the properties of the full-length scales reported in previous literature. Age and sex differences in performance, internal consistency, and factor structure were examined for each short-form scale. The test-retest reliability of the short form RMET and its correlation with the short-form EQ were also calculated.

The short-form scales were found to perform similarly to the full-length versions. The psychometric properties of the scales were robust to changes in scale length and response options. The properties of the short-form EQ suggested that the scale is appropriate for future research use. Despite performing similarly to the full-length scale, the properties of the two short-form versions of the RMET suggested that they are not appropriate measures for research use. It is recommended that the results of investigations using the RMET as a measure of cognitive empathy, including the results reported in Chapter 4, be interpreted with caution.

Psychometric properties of three short-form measures of empathy in an Australian sample

Social support and social strain have important effects on our physical and mental health (Ditzen & Heinrichs, 2014). The results reported in Chapter 3 suggest that *AVPR1A RS3* genotype is also associated with individual differences in self-reported mental and physical health in women. Evidence from previous literature suggests that the *AVPR1A* gene is associated with a variety of social behaviours (Albers, 2012; Ebstein et al., 2012) and it was theorised that social behaviour or cognition may provide a link between genetic variation at the *RS3* locus and individual differences in health. Previous research suggests that individual differences in empathy may have health impacts; both high and low levels of empathy can have negative effects on an individual's mental health (Smith & Rose, 2011; Tone & Tully, 2014). Chapter 4 explored the hypothesis that *AVPR1A RS3* genotype was associated with individual differences in self-reported empathy and cognitive empathy ability. However, due to the limitations of the study, the results of Chapter 4 were inconclusive.

Accurate and efficient measures of empathy can facilitate research to understand the trait's role in social behaviour and health, and to understand its biological basis. For example, valid short-forms of common empathy measures may facilitate the measurement of empathy in the large-scale studies required to explore complex genetic associations. The data set used in Chapter 4 contained three new short-forms of common empathy measures: one short version of the Empathy Quotient (EQ-18), and two short versions of the Reading the Mind in the Eyes Test (RMET-17 and RMET-14). In this chapter, the psychometric properties of these short-form empathy measures are examined and reported to aid in the interpretation of the findings of Chapter 4, and to determine if the measures are appropriate for future research use.

Measuring Empathy

Empathy is a key part of social cognition, encompassing the ability to recognise and understand the emotions, thoughts and intentions of other people (Decety & Jackson, 2004). The concept of empathy is generally divided into two related domains; emotional empathy and cognitive empathy (Decety & Jackson, 2004; Singer, 2006). Many instruments have been developed to measure cognitive and emotional empathy, including self-report questionnaires (e.g. Baron-Cohen & Wheelwright, 2004; Davis, 1980; Jolliffe & Farrington, 2006) and tests of ability (e.g. Baron-Cohen, Wheelwright, Hill, et al., 2001; Dziobek et al., 2008; Happé, 1994). Self-report questionnaires measure an individual's tendency to experience empathy or their perceived empathy abilities. Objective measures of cognitive empathy test emotion recognition and perspective taking abilities. Emotional empathy can be objectively assessed by measuring changes in an individual's physiological arousal in response to empathy-provoking images.

The EQ (Baron-Cohen & Wheelwright, 2004) and the RMET (Baron-Cohen, Wheelwright, Hill, et al., 2001) are commonly used measures of empathy. While the short-forms of these scales used in Chapter 4 have not previously been psychometrically assessed, the psychometric properties of the full-length versions of the scale are available in the literature. The goal of the present study was to determine whether the short-form scales had similar psychometric properties to the full-length scales and whether they were appropriate for research use.

Empathy Quotient

The EQ was designed to be a single-factor self-report measurement of empathy, combining the cognitive and affective aspects of the construct (Baron-Cohen & Wheelwright, 2004). The original scale consists of 40 empathy-related statements (e.g. "Other people tell

me I am good at understanding how they are feeling and what they are thinking") and 20 filler or distractor items (Baron-Cohen & Wheelwright, 2004). Participants are asked to rate how well each statement describes them on a 4-point forced-choice Likert scale, from "strongly disagree" to "strongly agree". On the empathy related items, participants receive a score of 0 for non-empathic responses of any level of agreement and score 1 or 2 for empathic responses depending on the strength of their agreement (Baron-Cohen & Wheelwright, 2004).

Psychometric properties. The EQ has been found to have good internal consistency as measured by Cronbach's alpha (0.85 – 0.92; Baron-Cohen & Wheelwright, 2004; Muncer & Ling, 2006; Wakabayashi et al., 2006) and good test-retest reliability (r = 0.84 – 0.97; Baron-Cohen & Wheelwright, 2004; Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004). The scale was originally proposed to consist of a single factor (Baron-Cohen & Wheelwright, 2004), and one subsequent study supports this finding (Wakabayashi et al., 2006). Wakabayashi and colleagues (2006) aimed to create a short version of the EQ containing only key contributing items. A principle components analysis (PCA) was conducted and the scree plot was reported to suggest a single-component solution was adequate, however, three components had Eigen values greater than one and a possible three-factor solution was not explored (Wakabayashi et al., 2006). The results of two studies suggest that the scale is better explained by a three-factor model than a single-factor model (Lawrence et al., 2004; Muncer & Ling, 2006). An exploratory (Lawrence et al., 2004) and confirmatory (Muncer & Ling, 2006) factor analysis suggest that the EQ items fall into three thematic categories; cognitive empathy, emotional reactivity and social skills. Small-to-moderate positive correlations have been found between these factors suggesting that they are measuring distinct but related aspects of the empathy construct (Lawrence et al., 2004; Muncer & Ling, 2006).

Women are typically found to score higher than men on the full-length EQ (Baron-Cohen & Wheelwright, 2004; Lawrence et al., 2004; Muncer & Ling, 2006; Wakabayashi et

al., 2006), which is a common finding for self-report measures of empathy (Maibom, 2012). Women have also been found to score higher than men on the cognitive empathy and emotional reactivity factors of the EQ, but no sex differences have been found for the social skills factor (Lawrence et al., 2004; Muncer & Ling, 2006). Scores on the EQ have also been found to negatively correlate with scores on the Autism-Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, et al., 2001), a measure of autistic traits in the general adult population (Baron-Cohen & Wheelwright, 2004; Melchers, Montag, Markett, & Reuter, 2015), and Adults with Asperger's Syndrome (AS) or High Functioning Autism (HFA) score significantly lower on the EQ than age- and gender-matched controls (Baron-Cohen et al., 2014; Baron-Cohen & Wheelwright, 2004).

Reading the Mind in the Eyes Test

The RMET is a task which aims to measure individual differences in cognitive empathy or Theory of Mind (ToM) ability (Baron-Cohen, Wheelwright, Hill, et al., 2001). The RMET tests an individual's ability to identify complex mental and emotional states from images of the eye region of the face. The test includes 36 images taken from movie stills and other photos of male and female ancestrally European actors expressing emotions or other complex mental states (Baron-Cohen, Wheelwright, Hill, et al., 2001). Response options are one-word descriptions of the mental state being displayed and participants must select the correct description from four multiple-choice options; one correct answer and three-foils (Baron-Cohen, Wheelwright, Hill, et al., 2001). To increase the difficulty of the test all three foils have the same valance (positive, negative or neutral) as the correct answer. Items vary in difficulty to produce a fine-grade measure of individual differences in cognitive empathy ability (Baron-Cohen, Wheelwright, Hill, et al., 2001).

The RMET is designed to test recognition of complex mental states (i.e. cognitive mental states such as beliefs or intentions), rather than the more easily recognised basic emotions (i.e. happiness, sadness, anger, disgust, fear and surprise; Baron-Cohen, Wheelwright, Hill, et al., 2001). In order to describe these complex mental states, response options consist of high-level vocabulary words. In the recommended administration of the test, participants are provided with a glossary of the adjectives used and are encouraged to consult it if there are words they do not understand (Baron-Cohen, Wheelwright, Hill, et al., 2001).

Psychometric properties. Although the RMET is a commonly used measure in empathy and Theory of Mind research, few studies have specifically investigated its psychometric properties. This is particularly true for the English version of the test; the available psychometric data primarily comes from translations of the test (Vellante et al., 2013). The RMET has been translated into several languages including; Dutch, French, German, Italian, Japanese, Persian, Spanish, Swedish, and Turkish. Multiple translations have been made for some languages. The translated versions of the test vary in item number, the difficulty of the words used, and the translation issues encountered.

The reported psychometric properties also vary across the English and translated versions of the RMET. Previous studies have utilised a variety of assessment methods, however, few studies have performed a comprehensive psychometric analysis of the test. Internal consistency measured using Cronbach's alpha ranges from 0.37 to 0.77 (Khorashad et al., 2015; Prevost et al., 2014), with the majority of studies reporting values between 0.58 and 0.64 (Harkness, Jacobson, Duong, & Sabbagh, 2010; Söderstrand & Almkvist, 2012; Vellante et al., 2013; Voracek & Dressler, 2006). The most commonly found alpha values are borderline acceptable for a 36-item scale. Only two studies have found acceptable alpha values (Dehning et al., 2012; Prevost et al., 2014) and of those, Prevost et al. (2014) only

found acceptable values for the English version of the scale (Cronbach's alpha = 0.77), but not for the French translation (Cronbach's alpha = 0.53).

The reported evidence for sex differences also varies. There is some evidence that women score significantly higher than men on the RMET (Hallerbäck, Lugnegård, Hjärthag, & Gillberg, 2009; Khorashad et al., 2015; Vellante et al., 2013; Voracek & Dressler, 2006). However, the size of this effect is small (Voracek & Dressler, 2006) and several studies have found no sex differences in performance (Harkness et al., 2010; Olderbak et al., 2015; Söderstrand & Almkvist, 2012; Spek, Scholte, & Berckelaer-Onnes, 2010). While women are typically found to score higher on tests of dispositional empathy, this sex difference is not consistently found for tests of situational empathy or empathy ability (Eisenberg & Lennon, 1983; Ickes, Gesn, & Graham, 2000; Maibom, 2012). In contrast, evidence for the test-retest reliability of the RMET is consistent. Several studies have reported good reliability over time periods ranging from one week to one year (Fernández-Abascal, Cabello, Fernández-Berrocal, & Baron-Cohen, 2013; Hallerbäck et al., 2009; Khorashad et al., 2015; Prevost et al., 2014; Vellante et al., 2013; Yıldırım et al., 2011).

Explanations for variation. The variability in psychometric properties across studies may be due to differences between the translated versions of the RMET and differences between the populations the test has been studied in. While the cross-cultural expression and perception of basic emotions has been well-researched (Elfenbein & Ambady, 2003), the expression and perception of complex mental states is not as well understood. Thus, it may be difficult to produce an accurate, culturally appropriate translation of the RMET. The difficulty level of the vocabulary used is likely to vary between translations and the relationship between the response options and images, and between the target word and foils may change in the translation process (Hallerbäck et al., 2009).

Additionally, individual performance on the RMET is likely to be affected by the "other-race effect", a well-documented effect in which individuals are found to be poorer at recognising faces from other races than faces from their own race (Elfenbein & Ambady, 2002). A study conducted by Adams and colleagues (2010) found evidence of a performance advantage on the RMET when judging the mental states of own-race faces. Japanese participants and European American participants were compared on their performance on the original European RMET and on a Japanese version of the RMET (Adams et al., 2010). The Japanese version included the same (translated) target and foil words but the images used were of Asian faces. The American participants performed better on the original European version of the test and the Japanese participants performed better on the Japanese version (Adams et al., 2010). Thus, the properties of the RMET are likely to vary between different language versions and between participant groups.

Factor structure. The factor structure of the original, English language version of the RMET has been investigated in two studies. The test is conceptualised as tapping a single underlying construct, however, neither study found evidence for any particular factor structure (Olderbak et al., 2015; Ragsdale & Foley, 2011). Ragsdale and Foley (2011) found that the RMET had low inter-item correlations; most item correlations were below 0.2 and many correlations were negative. A Principal Components Analysis produced 15 components with Eigenvalues greater than 1 and there were no thematic similarities between items loading onto the same factor (Ragsdale & Foley, 2011). No relationships were found between items falling into thematic subgroups of emotion or mental state, thus participant performance on items depicting the same or similar expressions was apparently unrelated. Direction of eye gaze (towards or away from the viewer) was not associated with performance and there were no thematic relationships between items answered correctly most often or incorrectly most often (Ragsdale & Foley, 2011). The authors suggest that accuracy on test items may therefore rely

more on the content of the foil words than on the target emotion (Ragsdale & Foley, 2011). These results indicate that the RMET may not be measuring emotion or mental state recognition and further investigation is needed to determine what is actually being measured.

Olderbak and colleagues (2015) also found that the RMET had poor inter-item correlations and a confirmatory factor analysis did not find evidence for a single factor solution. The confirmatory factor analysis also found that a two factor and a three factor model suggested by previous literature were unsatisfactory (Harkness, Sabbagh, Jacobson, Chowdrey, & Chen, 2005; Konrath, Corneille, Bushman, & Luminet, 2013; Olderbak et al., 2015). The unsatisfactory thematic models split the scale into positive and negative affect categories (Konrath et al., 2013) or positive, negative and neutral affect categories (Harkness et al., 2005). Using Ant Colony Maximisation, a data-driven method, Olderbak and colleagues (2015) produced a 10-item version of the RMET with acceptable internal-consistency and average inter-item correlations greater than 0.2. This version of the test was strongly related to a measure of vocabulary but weakly-to-moderately related to measures of emotion perception and cognitive empathy (Olderbak et al., 2015). Together, these studies suggest that the original RMET has poor psychometric properties and performance on the test may measure multiple constructs or rely on multiple cognitive abilities.

Relationship Between the EQ and RMET

The EQ and RMET are designed to tap the underlying construct of cognitive empathy to different extents; the EQ measures both the emotional and cognitive aspects of dispositional empathy (Baron-Cohen & Wheelwright, 2004) and the RMET is designed to only measure cognitive empathy ability (Baron-Cohen, Wheelwright, Hill, et al., 2001). Additionally, the EQ is a self-report measure of perceived empathy ability, whereas the RMET is a test of cognitive empathy performance. While scores on the RMET and EQ are

likely to reflect each other, an individual may over- or underestimate their empathy abilities. Thus, the EQ and RMET are expected to moderately positively correlate. Some studies have found moderate positive correlations between RMET and EQ score (Lawrence et al., 2004; Ragsdale & Foley, 2011; Spek et al., 2010; Voracek & Dressler, 2006), however, others have found no relationship between the measures (Melchers et al., 2015), or a difference in RMET score only for individuals scoring below a certain cut-off point on the EQ (Vellante et al., 2013). The variability in these results might be due to the variable performance of the different versions of the RMET discussed above. Together, previous evidence suggests that the RMET and EQ may measure some of the same underlying constructs, or related constructs, and are expected to moderately positively correlate.

Aims

The present study aimed to determine whether the EQ-18, RMET-17, and RMET-14 displayed similar psychometric properties to those previously reported for the full-length scales. The EQ-18 was expected to have good internal consistency and a three-factor structure (cognitive empathy, emotional reactivity and social skills; Lawrence et al., 2004; Muncer & Ling, 2006). EQ-18 scores were predicted to differ significantly between men and women and to correlate moderately with the RMET-17. Due to the variability of previous reports on the properties of the RMET, few predictions about the performance of the two short-form measures could be made. The RMET-17 and RMET-14 were expected to have no clear factor structure and poor inter-item correlations. However, performance on the test was expected to be stable over time.

Methods

Participants

The participants for this study were a subset of the participants from the Brisbane Longitudinal Twin Study conducted at the Queensland Institute of Medical Research Berghofer Medical Research Institute (Wright & Martin, 2004). The sample included adolescent twins, their siblings and their parents. This subset of participants had completed the "Genetic and Environmental Foundations of Political and Economic Behaviours: A Panel Study of Twins and Families" sub-study (Hatemi et al., 2015). Data was collected in two waves completed approximately 12 months apart. The RMET-17 and EQ-18 were administered to the twins in Wave 1 and the RMET-14 was administered to twins, their parents and their siblings in Wave 2. Demographic data is presented in Table 5.1.

Table 5.1. Sample size, age and family position of participants

	Wave 1	**Wave 2**
Total Participants (*N*)	584 (338 female, 246 male)	1641 (1004 female; 637 male)
Twins (*n*)	584 (338 female, 246 male)	614 (390 female, 224 male)
Parents (*n*)		903 (538 female, 365 male)
Siblings (*n*)		124 (76 female, 48 male)
Age (years)		
Mean	25	41
Standard Deviation	3	14
Range	19 - 31	20 - 75

Scales

The Empathy Quotient. A shortened version of the EQ (EQ-18) containing 18 of the original 40 empathy-related items was included in the Wave 1 questionnaire. Participants were asked to rate how well each item described them on a 5-point Likert scale ranging from 1 = very inaccurate to 5 = very accurate, with 3 representing a neutral option. The minimum possible score was 18 and the maximum was 90. Higher scores indicate a higher level of self-reported dispositional empathy.

The Reading the Mind in the Eyes Test. In Wave 1, participants were asked to complete a 17-item version of the RMET (RMET-17). Participants were awarded 1 point for a correct answer and 0 points for an incorrect answer. The lowest possible score was 0 and the highest was 17, with a higher score indicating better performance on the test. Wave 2 included a 14-item version of the RMET (RMET-14), created by dropping three of the items used in Wave 1. Additionally, a "Don't know" response was added to the multiple-choice response options available. In this version of the test, participants were scored 1 point for a correct answer and 0 points for an incorrect answer or a "Don't know" response. The maximum possible score on the test was 14 and the minimum possible score was 0, with a higher score indicating better performance on the task.

Demographics. Demographic information about participants was collected in Wave 1 and Wave 2. Descriptive statistics for participant age and sex are presented in Table 5.1.

Statistical Methods

All analyses were carried out using the IBM SPSS 22 statistical package. A significance value of $\alpha = 0.05$ was used for all applicable analyses. Descriptive statistics were calculated for the total score on each scale and the distribution of scores was examined for normality. Based on the results of Chapter 4, the correlation between age and each empathy

variable (EQ-18, RMET-17 and RMET-14) was examined. Sex differences in performance were determined for each scale, and the factors of the EQ, using an independent samples *t*-test. The internal consistency of the scales was estimated using Cronbach's alpha. The factor structure of each scale was examined using an exploratory Principle Components Analysis (PCA). For the EQ-18, the correlations between suggested factors were calculated. Correlations between EQ total score and RMET-17 score, and between the RMET-17 and each of the EQ factors were also calculated. All correlations reported in text are Pearson correlation coefficients.

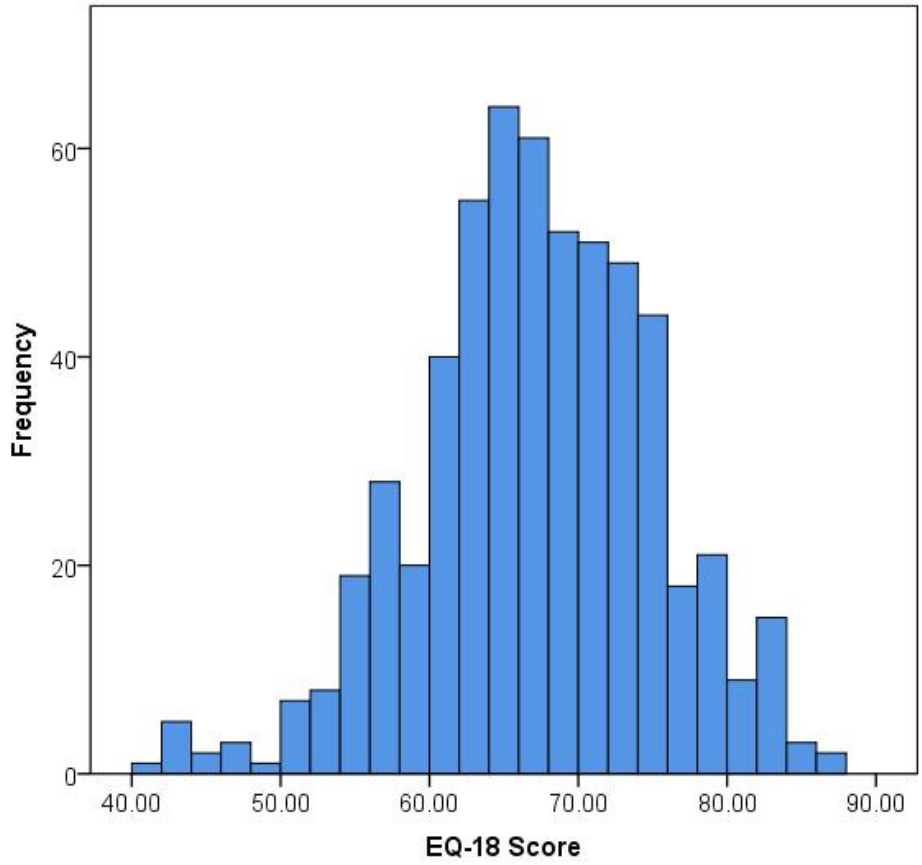
For the RMET, test-retest reliability over a 12-month time-period was estimated using the 361 participants who completed the test in both Wave 1 and Wave 2. An RMET-14 score was produced for Wave 1 by excluding the three items not included in Wave 2, and the correlation between RMET-14 scores from Wave 1 and Wave 2 was calculated. Test-retest reliability could not be calculated for the EQ-18 because the measure was only administered in Wave 1.

Results

EQ-18

The EQ was administered in Wave 1 and was completed by 581 participants. Two participants were excluded from the analyses as their attempts were judged to be non-serious. Additionally, one participant scored more than 3 standard deviations lower than the group mean and was classified as an outlier and excluded from the analyses. This left a total sample of 578 participants (333 female, 245 male). The mean score on the EQ-18 was 66.65 (SD = 7.98), the minimum score was 41 and the maximum score was 86. The distribution of scores was approximately normal and scores were not transformed (Figure 5.1). EQ-18 score was not significantly correlated with participant age ($r = 0.024$, $p = 0.573$).

Figure 5.1. Distribution of EQ-18 scores.



Exploratory PCA. A Pearson's correlation matrix showed that all items correlated with at least one other item at 0.2 or above, except for item 18 ("I don't consciously work out the rules of social situations"; Baron-Cohen & Wheelwright, 2004). A PCA with varimax rotation showed commonalities to be above 0.25, except for item 16 and item 18. Item 16 was retained as it loaded onto the second factor in the final model. Item 18 did not load onto any factor and was removed, leaving a total of 17 items.

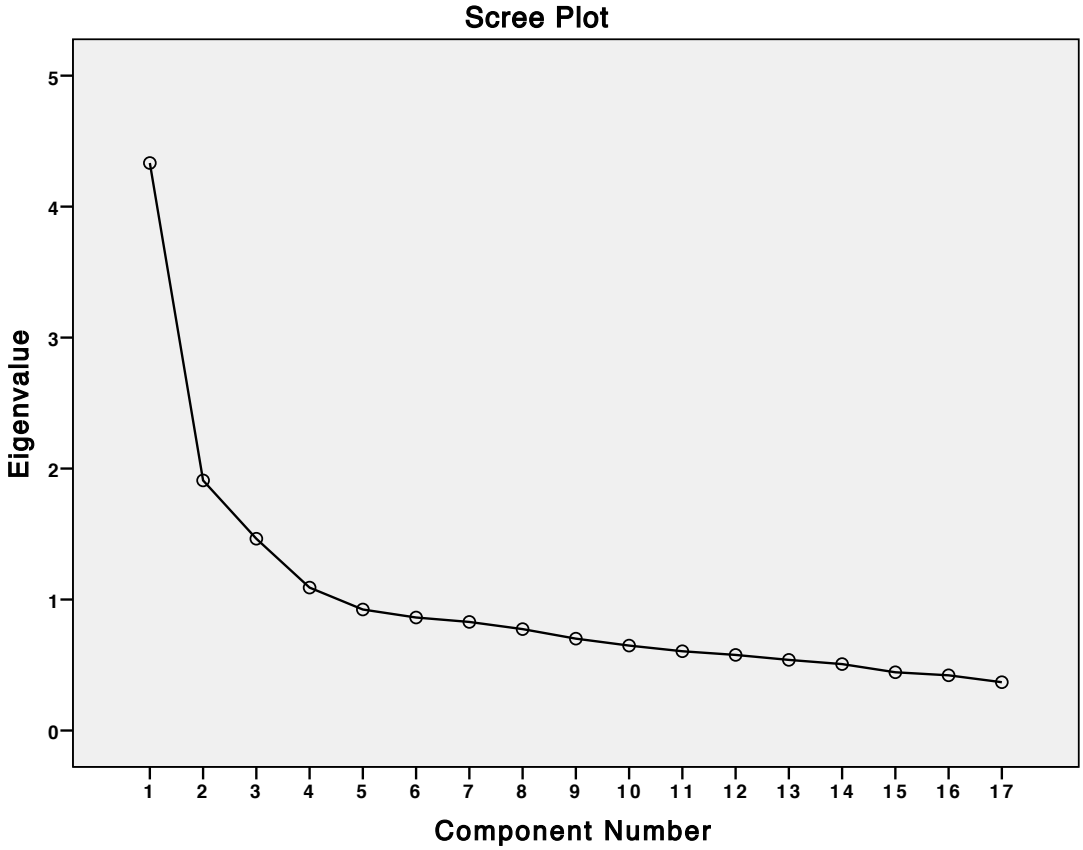
The scree plot (Figure 5.2) and eigenvalues (Table 5.2) suggested a model with three factors, explaining 43% of the variance. Item loadings for these factors in the rotated solution are shown in Table 5.3. Some items loaded on multiple factors, as found in previous analyses (Lawrence et al., 2004; Muncer & Ling, 2006). Following Lawrence and colleagues (2004), items loading onto more than one factor were allocated based on their content. Factor 1 has 7

items and corresponds to the cognitive empathy factor, factor 2 has 6 items and corresponds to the emotional reactivity factor and factor 3 has 4 items and corresponds to the social skills factor. A full description of the items can be found in Appendix E.

Table 5.2. EQ-17: Eigenvalues for the first 3 components from an exploratory PCA with varimax rotation.

Component	Eigenvalue	% Variance explained
1	2.958	16.44
2	2.651	14.73
3	2.124	11.80

Figure 5.2. Scree plot for the EQ-18.



The scree plot supports the presence of three components, as suggested by the eigenvalues (Table 5.2).

Table 5.3. Final loadings from the EQ-17 PCA.

Item	1	2	3
6	0.597		
9	0.757		
10	0.704		
15	0.502		
17	0.702		
1	0.518		0.360
8	0.509	0.425	
7	0.311	0.619	
11		0.707	
12		0.724	
13		0.629	
16		0.447	
5		0.413	0.314
2			0.809
3			0.713
4			0.599
14			0.552

Factor 1 corresponds to the previously reported cognitive empathy factor, factor 2 corresponds to the emotional reactivity factor and factor 3 corresponds to the social skills factor. Item numbers indicate the order in which items were presented in Wave 1, a full description of items can be found in Appendix E.

Cronbach's alpha. The Cronbach's alpha for the 18-item scale was 0.79 and was 0.80 for the 17-items included in the final PCA. For the factors, Cronbach's alpha was 0.76 for cognitive empathy, 0.69 for emotional reactivity and 0.65 for social skills. These values are lower than the full scale but acceptable given that the factors contained seven, six and four items respectively.

Inter-factor correlations. Cognitive empathy and emotional reactivity were significantly positively correlated ($r = 0.410$, $p < 0.001$), as were cognitive empathy and social skills ($r = 0.381$, $p < 0.001$) and emotional reactivity and social skills ($r = 0.247$, $p < 0.001$).

Sex differences. Women scored significantly higher than men on the 17-item EQ and the cognitive empathy and emotional reactivity factors (Table 5.4). No sex difference was found for the social skills factor.

Table 5.4. Sex differences in performance on the EQ-17 and its factors.

Scale		Mean (*SD*)	*t*	*p*	Cohen's *d*
EQ-17					
	Women	64.92 (7.48)	5.055	**< 0.001**	0.42
	Men	61.60 (8.21)			
Cognitive Empathy					
	Women	26.65 (3.59)	3.06	**0.002**	0.25
	Men	25.69 (3.94)			
Emotional Reactivity					
	Women	22.65 (3.77)	6.36	**<0.001**	0.54
	Men	20.63 (3.77)			
Social Skills					
	Women	15.62 (2.91)	1.40	1.62	
	Men	15.28 (2.82)			

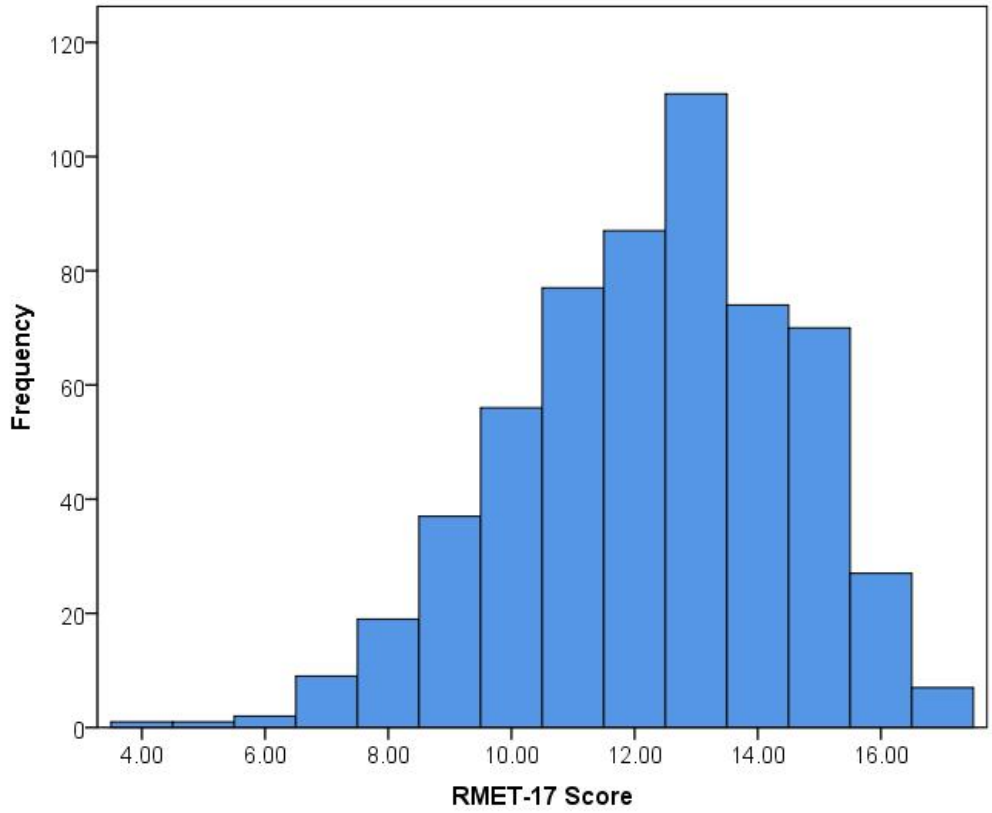
Note: $df = 576$

Significant results are bolded.

RMET-17

In Wave 1, 580 participants completed the RMET. Two participants were excluded, as their attempts were non-serious, leaving a total of 578 participants (333 female, 245 male) in subsequent analyses. The mean RMET-17 score was 12.28 (*SD* = 2.26), the minimum score was 4 and the maximum score was 17. The distribution of scores was approximately normal with some negative skew present (Figure 5.3). Scores were not transformed. RMET-17 score was not significantly correlated with participant age ($r = 0.003$, $p = 0.950$).

Figure 5.3. Distribution of RMET-17 scores from Wave 1.



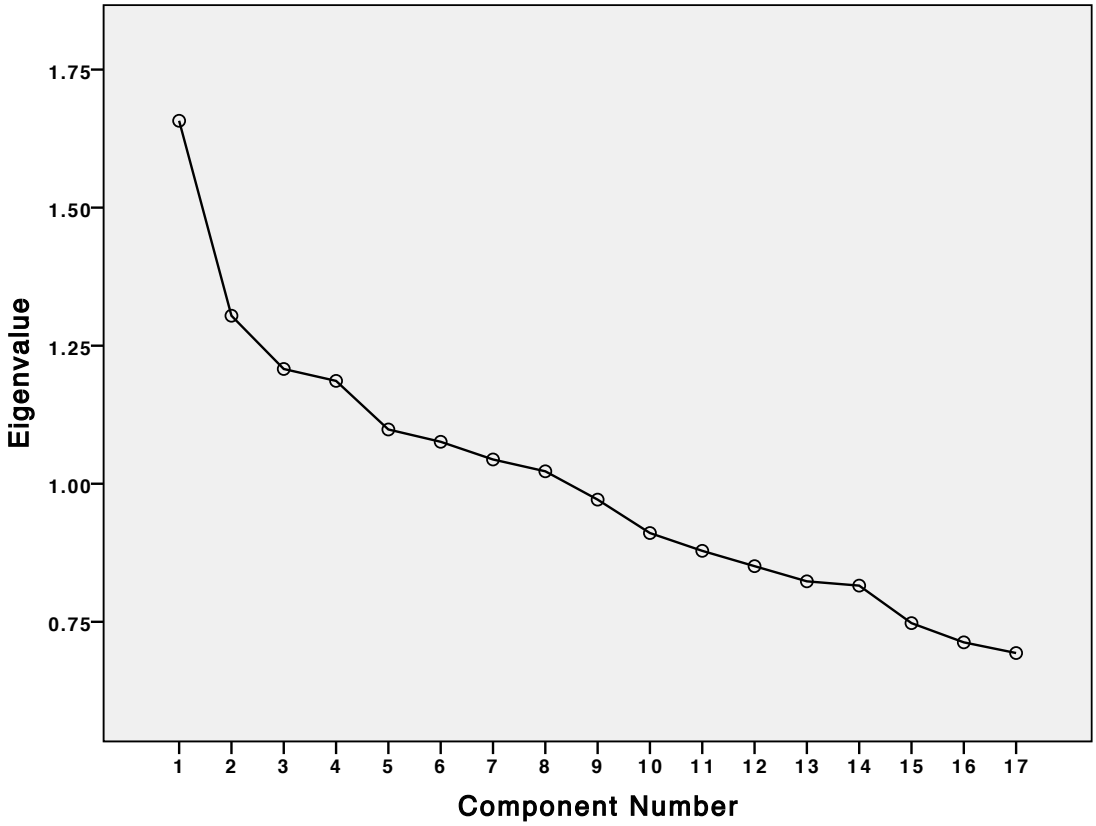
Cronbach's alpha. The Cronbach's alpha for the RMET-17 was 0.40, which is low and not acceptable for a 17-item scale.

Exploratory PCA. A Pearson's correlation matrix showed that inter-item correlations were low, with no correlations exceeding 0.2 and few correlations exceeding 0.1. The PCA failed to reveal any clear item groupings or factor structure in the data, producing 8 components with Eigenvalues greater than 1, accounting for 56% of the variance (Table 5.5, Figure 5.4).

Table 5.5. RMET-17: Eigenvalues for the first 8 components from exploratory PCA with varimax rotation.

Component	Eigenvalue	% Variance explained
1	1.349	7.94
2	1.292	7.60
3	1.225	7.21
4	1.192	7.01
5	1.149	6.76
6	1.148	6.75
7	1.125	6.62
8	1.117	6.57

Figure 5.4. Scree plot for the RMET-17.



The scree plot suggests that there are not clearly distinguishable components.

Correlation with EQ-17. Small, significant correlations were found between RMET-17 score and EQ-17 score ($r = 0.11$, $p = 0.009$), and scores on the cognitive empathy ($r = 0.10$, $p = 0.019$) and emotional reactivity ($r = 0.10$, $p = 0.015$) factors. No correlation was found between RMET-17 score and score on the social skills factor ($r = 0.04$, $p = 0.362$).

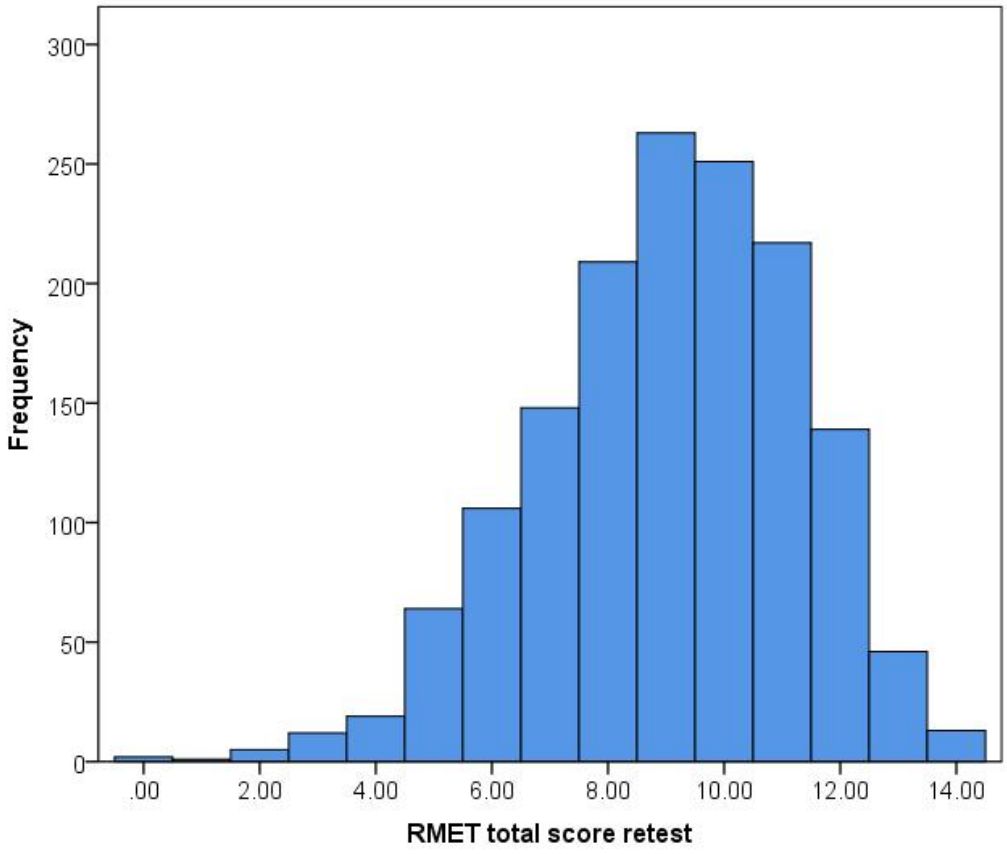
Sex differences. No sex differences were found for performance on the RMET-17 (Women: $M = 12.34$, $SD = 2.32$; Men: $M = 12.21$, $SD = 2.18$; $t(576) = 0.69$, $p = 0.491$).

RMET-14

The RMET-14 was administered in Wave 2. Of the 1641 individuals who participated in this wave, 1495 (897 female, 598 male) completed all items of the RMET-14. The mean score on the RMET-14 was 9.03 ($SD = 2.25$), the minimum score was 0 and the maximum

score was 14. The distribution of scores was approximately normal with some negative skew present (Figure 5.5). Scores were not transformed. RMET-14 score was significantly, negatively correlated with participant age (r = -0.196, p < 0.001).

Figure 5.5. Distribution of RMET-14 scores from Wave 2.



Cronbach's alpha. Cronbach's alpha was 0.44, which is low and not acceptable for a 14-item scale.

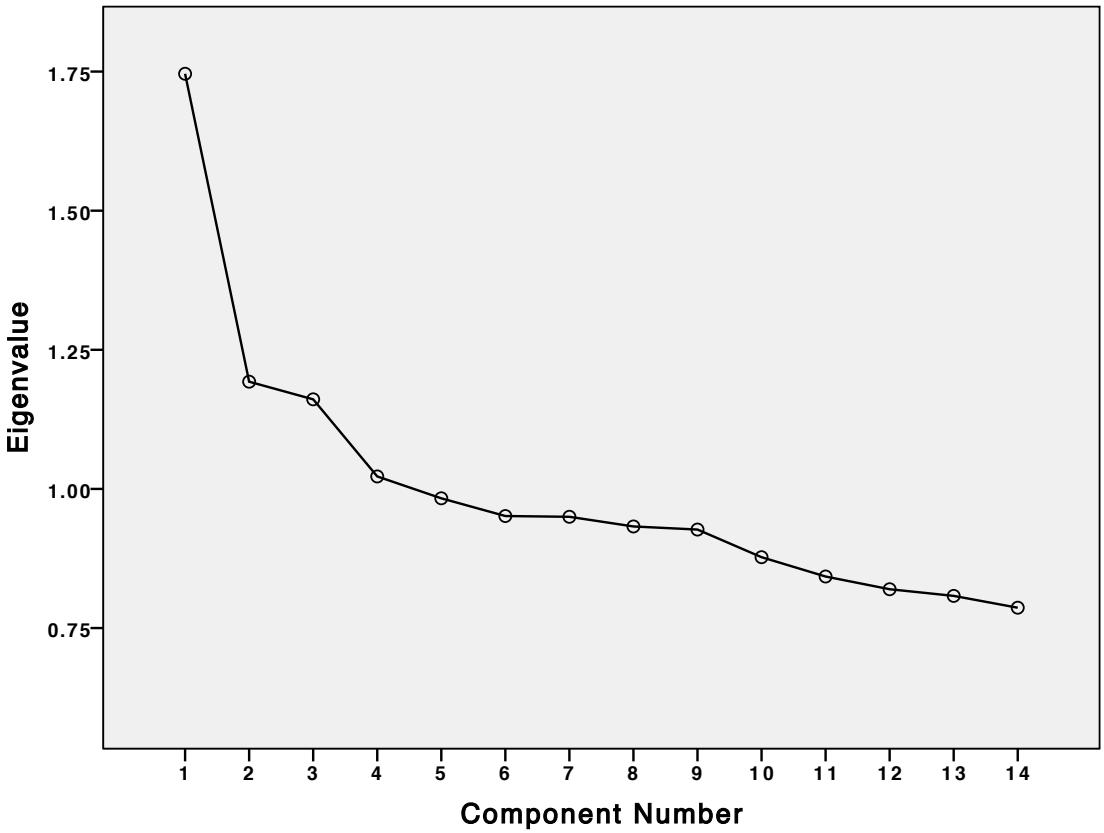
Exploratory PCA. A Pearson's correlation matrix showed that inter-item correlations were low, with no correlations exceeding 0.2, few exceeding 0.1 and several small negative correlations. A PCA with varimax rotation suggested four components might be present, explaining 37% of the variance (Table 5.6, Figure 5.6). However, no clear item groupings

were suggested and items loading onto the same factor had no apparent thematic similarities (e.g. type or valence of emotion, gender of image, gaze direction of image).

Table 5.6. RMET-14: Eigenvalues for the first 4 components from exploratory PCA with varimax rotation

Component	Eigenvalue	% Variance explained
1	1.337	9.55
2	1.317	9.40
3	1.260	9.00
4	1.209	8.63

Figure 5.6. Scree plot for the RMET-14.



The scree plot suggests three components may be present.

Sex differences. A significant difference in performance was found between men and women, with women scoring higher than men (Women: $M = 9.21$, $SD = 2.19$; Men: $M = 8.76$, $SD = 2.32$; $t(1493) = 3.85$, $p < 0.001$). The effect size of sex on RMET-14 performance was small (Cohen's $d = 0.20$).

Test-retest reliability. A moderate positive correlation between RMET-14 score in Wave 1 and Wave 2 was found ($r = 0.47$, $p < 0.001$).

Discussion

The EQ-18, RMET-17 and RMET-14 were examined to determine if they had similar psychometric properties to their full-length equivalents. Overall, the psychometric properties of the short-form scales were found to be similar to the properties of the full-length scales reported in the literature. The properties of the EQ-17 suggest that the scale is appropriate for research use. The RMET-17 and RMET-14 also performed similarly to the original test. However, the results of this study and previous research suggest that the RMET has poor internal consistency and may be measuring multiple underlying constructs or rely on multiple cognitive abilities (Olderbak et al., 2015; Ragsdale & Foley, 2011).

EQ-17

The results of these analyses suggest that the EQ-17 performs similarly to the original EQ scale. The scale had good internal consistency and a similar factor structure to that found in previous validation studies (Lawrence et al., 2004; Muncer & Ling, 2006). Items included in Lawrence and colleagues (2004) and Muncer and Ling's (2006) final models loaded onto similar factors in the present analysis. The content of items that were not included in previously published models matched the content of previously described factors (emotional reactivity, cognitive empathy, and social skills). However, the items included in the emotional

reactivity factor may be better described as "concern for others" in the present study (see Appendix E for a description of items). The pattern of correlations between factors was also similar to previous findings, with the highest correlation between emotional reactivity and cognitive empathy, and the lowest correlation between emotional reactivity and social skills (Lawrence et al., 2004; Muncer & Ling, 2006).

The EQ-17 also displayed the expected patterns of sex differences. As predicted from previous studies (Baron-Cohen & Wheelwright, 2004; Lawrence et al., 2004; Muncer & Ling, 2006; Wakabayashi et al., 2006), women scored significantly higher on the EQ-17 than men. Additionally, women were found to score higher than men on the cognitive empathy and emotional reactivity factors, but no sex differences were found for the social skills factor, replicating previous findings (Lawrence et al., 2004; Muncer & Ling, 2006). Overall, these results suggest that the EQ-17 is tapping the same constructs of dispositional empathy as the full-length EQ scale.

RMET-17

The RMET-17 had low reliability and no clear factor structure. Previous evidence suggests these features were not unique to the 17-item version of the scale. Ragsdale and Foley (2011) found similar results in an examination of the full 36-item RMET: Cronbach's alpha was low (0.48) and most inter-item correlations were less than 0.2 with many negative correlations present. Olderbak and colleagues (2015) found a similar pattern of inter-item correlations. The low internal consistency of the scale could result from the wide variety of emotions and mental states represented by the items; different processes could underlie the recognition of the different complex emotions and mental states. If this is the case, performance on certain items may be unrelated to performance on other items but the RMET would be expected to contain several factors based on groups of similar emotions and mental

states. No clear item groupings were found in the present study. The lack of a clear factor structure for the RMET-17 is surprising as many of the items overlap in theme (e.g. items displaying suspicious, cautious and nervous expressions) or represent the same expression (e.g. for two items the correct answer was "fantasising"). The RMET is generally assumed to have a single underlying factor (Baron-Cohen, Wheelwright, Hill, et al., 2001), however, this assumption is not supported by the results of the present study. Ragsdale and Foley (2011) also found little evidence of item grouping and found no evidence of a thematic relationship between the items that were most often answered correctly and incorrectly. Olderbak and colleagues (2015) tested factor models based on emotion and mental state valence but found they were not a good fit for their data. These results make it difficult to determine what construct (or constructs) the RMET is measuring.

Ragsdale and Foley (2011) suggest that the RMET may actually be measuring general executive functioning. They entered RMET scores into a PCA with EQ, AQ and Behavioural Inhibition System sensitivity (BIS; Carver & White, 1994) subscale scores. BIS sensitivity was used a measure of executive function. The RMET loaded onto the same component as BIS sensitivity and the Attention Switching and Attention to Detail subscales of the AQ (Ragsdale & Foley, 2011). This suggests the RMET may be measuring executive functioning, rather than Theory of Mind or emotion recognition. The items of the test display a range of emotions and mental states which may require different processes to recognise. Thus, an individual's performance on the RMET may be driven by the executive functioning skills required to complete the task, rather than by emotion or mental state recognition abilities (Ragsdale & Foley, 2011).

The RMET and EQ are designed to measure similar constructs, therefore, some correlation between the scales was expected. Small significant positive correlations were found between the RMET-17 and the EQ-17, and between the RMET-17 and two of the EQ-

17 factors: cognitive empathy and emotional reactivity. Though small, these correlations suggest that the RMET and EQ are to some extent measuring similar or related constructs. Lawrence and colleagues (2004) found a moderate positive relationship between total RMET score and total EQ score ($n = 48$, $r = 0.294$, $p = 0.033$). They also found a significant positive correlation between the RMET and their social skills factor ($n = 53$, $r = 0.273$, $p = 0.048$), but no correlations with the other factors. The RMET was not found to correlate with the social skills factor in the present study; however, this may be because social skills was the smallest factor, containing four items. It is possible that the strength of the correlations in the present study were reduced due to the truncation of the EQ and RMET, either through reduced inter-individual variation or through the loss of items potentially driving the relationship between the RMET and EQ.

Unlike the EQ-17, no sex difference was found in performance on the RMET-17. While women are commonly found to score higher than men on self-report measures of empathy, like the EQ, sex differences are less commonly found on task-based measures of empathy, like the RMET (Eisenberg & Lennon, 1983; Ickes et al., 2000; Maibom, 2012). Additionally, in studies where sex differences on the RMET have been found, the effect size has been small (Baron-Cohen, Wheelwright, Hill, et al., 2001; Lawrence et al., 2004).

Overall, these results suggest that the RMET-17 is not a reliable scale and they do not support the conceptualisation of the RMET as a single-factor measure of cognitive empathy. Previous research suggests that this is not unique to the RMET-17 and may be a feature of the original RMET scale (Olderbak et al., 2015; Ragsdale & Foley, 2011).

RMET-14

The properties of the RMET-14 are largely consistent with those found for the RMET-17. Cronbach's alpha and inter-item correlations were low. The exploratory PCA was more

suggestive of a potential factor structure, however, no clear thematic item groupings were identified. In contrast to the RMET-17, a significant difference in performance on the RMET-14 was found between men and women. As in previous studies, the effect of sex on RMET performance was small (Baron-Cohen, Wheelwright, Hill, et al., 2001; Lawrence et al., 2004). The difference between the RMET-17 and RMET-14 may be explained by the larger sample size in Wave 2; comparisons using Wave 2 data had more power to detect small effects.

The sample in Wave 2 also included a greater range of ages, resulting from the inclusion of the parents of the twin participants. A correlation between age and RMET-14 performance was found in Wave 2 but not in Wave 1. In combination with the ANOVA reported in Chapter 4 comparing twin and parent performance (Chapter 4, p148), this indicates that the correlation between age and RMET performance was driven by the difference between the young-adult twins and their middle-aged or older-adult parents. This difference may be the result of a cohort effect; the twin sample and parent sample may have responded differently to the RMET-14 measure. The age difference may also indicate a developmental difference in the trait measured by the RMET-14 between young adulthood and middle/older adulthood. Empathy has been found to change across adulthood (Tone & Tully, 2014), however, previous evidence suggests that the RMET may not be specifically measuring cognitive empathy (Olderbak et al., 2015; Ragsdale & Foley, 2011).

The test-retest reliability of the RMET-14 was low. The variation in participants' performance over a 12-month period could indicate changes in cognitive empathy ability (through practice or development), it could result from alterations made to the test in Wave 2 (i.e. the inclusion of a "Don't know" response option), or it could provide evidence that the RMET-14 is an unreliable measure. It is also possible that motivation to complete the task differed between Wave 1 and Wave 2 for some individuals. In Wave 1, the EQ-18 was

presented to participants immediately before the RMET-17, thus, the social aspect of the RMET may have been more salient in Wave 1 than in Wave 2, potentially altering performance (Ickes et al., 2000). The inconsistent patterns of twin correlations for the RMET-17 and RMET-14 reported in Chapter 4 (Tables 4.9 and 4.10) provide further evidence that the scale may be unreliable. However, previous research suggests that the full-length RMET has good test-retest reliability over periods of time ranging from one week to one year (Fernández-Abascal et al., 2013; Hallerbäck et al., 2009; Khorashad et al., 2015; Prevost et al., 2014; Vellante et al., 2013; Yıldırım et al., 2011). This suggests that the poor test-retest reliability observed in the present study is due to the changes made to the RMET-17 and RMET-14 or the conditions in which they were administered. The results of analyses utilising the RMET-17 and RMET-14 should be interpreted with caution.

The combined evidence from Wave 1 and Wave 2 suggests that the short-form RMET scales are not suitable measures of cognitive empathy ability for research applications. However, this does not appear to be unique to the present study or to the shortened versions of the scale (Olderbak et al., 2015; Ragsdale & Foley, 2011). Both the RMET-17 and RMET-14 appear to be performing in a similar manner to the full-length RMET, although the stability of participant performance over time may be reduced. Performance on the full-length and shortened versions of the RMET appears to rely on multiple traits and abilities, but it is unclear what constructs are being measured by the test (Ragsdale & Foley, 2011). While the results of these studies suggest that the scale is unreliable, the RMET is a popular and easy to administer scale that has been widely used in studies of cognitive empathy and Theory of Mind. It may remain useful as a tool for comparison with previous work.

Re-examining the Results of Chapter 4

The results of this study have consequences for the interpretation of the results of Chapter 4. The EQ-17 performs similarly to the full-length EQ scale. This supports the finding in Chapter 4 that *AVPR1A RS3* genotype is unlikely to be associated with individual differences in self-reported tendency to empathise (measured with the EQ-18). Additionally, using the methods described in Chapter 4 (pp138-139), a series of regression analyses failed to find an association between *AVPR1A RS3* genotype and individual differences in EQ-17 score, or between genotype and the cognitive empathy, emotional reactivity, and social skills factors (see Appendix F). The lack of a significant association between self-reported empathy and *RS3* genotype does not appear to be caused by changes made to the EQ scale. However, due to the small sample size available to test these relationships, a replication in a larger sample is required to conclusively confirm this non-association. As one previous study suggests that there is an association between *AVPR1A RS3* genotype and self-reported empathy (Uzefovsky et al., 2015), it is of particular interest to follow-up the near-significant association found between genotype and EQ-18 score in young-adult female twins (Chapter 4, Table 4.16) in a larger sample size. The results of the present study suggest that the EQ-17 is an appropriate measure to use in future investigations of this relationship.

In Chapter 4, the RMET-17 and RMET-14 were used as measures of cognitive empathy skill. However, the results of the present chapter, in combination with evidence from previous research, suggest that the RMET may not measure the single cognitive empathy factor it was designed to measure (Olderbak et al., 2015; Ragsdale & Foley, 2011). Instead, the scale may provide an indication of the executive functioning skills needed to successfully complete the task (Ragsdale & Foley, 2011). Thus, the small number of significant associations between RMET score and *AVPR1A RS3* genotype observed in Chapter 4 may not provide evidence for the biological basis of cognitive empathy.

Additionally, the unclear factor structure, low internal consistency, and poor test-retest reliability of the RMET-17 and RMET-14 may explain the variable pattern of gene associations and twin correlations observed in Chapter 4. One previous study has found an association between self-reported cognitive empathy and *AVPR1A RS3* variation (Uzefovsky et al., 2015). Future research should investigate whether genetic variation at the *RS3* locus is associated with cognitive empathy skill using a valid measure of this trait. The matching and labelling tasks described by Palermo, O'Connor, Davis, Irons, and McKone (2013) could be a potential alternative to the RMET. These tasks test the perception and identification of facial expressions for the six basic emotions. Like the RMET they are relatively easy to administer and they can detect individual differences in expression recognition in the general population (Palermo et al., 2013).

Conclusion

The results of the present study suggest that the RMET-17, RMET-14 and EQ-17 have similar psychometric properties to the full-length versions of the scales. These properties appear to be robust to changes in scale length and response options. The EQ-17 is an appropriate short-form of the EQ for use in research. The RMET-17 and RMET-14 do not have appropriate psychometric properties for research use, however, the RMET is a widely used scale in the literature and it may remain an important tool for comparing results with previous work. In light of the results of this study and evidence from the literature, analyses using the short-form or full-length versions of the RMET should be interpreted with caution.

Chapter 6: General Discussion

General Discussion

Social support, particularly perceived social support, is beneficial for our mental and physical health. Conversely, social strain is detrimental to mental and physical health (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016; Seeman et al., 2014). The impact of social support and strain varies between individuals, and this variation may be driven by biological factors (Ditzen & Heinrichs, 2014). Variants in the *AVPR1A* gene are associated with behaviours and traits that influence how people respond to their social environment, including empathy and empathy-related traits (Bisceglia et al., 2012; Knafo, Israel, et al., 2008; Uzefovsky et al., 2015). By influencing how people respond to social support and social strain, variation in the *AVPR1A* gene may modulate the health effects of social interaction. Exploring this relationship may help to elucidate the mechanisms underlying the influence of social support and strain on health.

Summary of Research Findings

The first study (Chapter 3) described in this thesis aimed to determine whether variation at the *AVPR1A RS3* locus was directly associated with individual differences in mental and physical health, and whether *AVPR1A RS3* genotype moderated the impact of social support and social strain on health. Based on previous research, it was hypothesised that the effect of *AVPR1A RS3* genotype would be strongest for male participants, however, no direct effect of genotype on mental or physical health was found for men. *AVPR1A RS3* genotype was found to have a significant effect on women's self-reported mental and physical health. Women with one or two copies of the Short *RS3* allele reported significantly better health than women homozygous for the Long *RS3* allele. The results of this study also indicate that *AVPR1A RS3* genotype may moderate the impact of social interaction on mental and physical health for men and women. However, the number of variables that could be

included in a single model was limited by the small size of the Short/Short genotype group. As a result, strong conclusions about moderation effects could not be drawn. A mediation model could not be tested as no direct relationship between *AVPR1A RS3* genotype and positive or negative social interaction was found. Together, these findings suggest that variation in the *AVPR1A* gene influences women's health and wellbeing but the mechanism for this effect remains unclear.

The second study (Chapter 4) aimed to explore the relationship between *AVPR1A* variation and empathy; a social trait that could link the *AVPR1A* gene to mental and physical health outcomes. This study sought to replicate the previously reported relationship between *AVPR1A RS3* variation and self-reported empathy (Uzefovsky et al., 2015), and determine whether genotype at this locus was associated with cognitive empathy ability. The results described in Chapter 4 provide some support for an association between *RS3* genotype and empathy; however, due to the limitations of the data set, it was not possible to conclusively demonstrate an effect of *AVPR1A RS3*. While overall, carrying two copies of the Long *RS3* allele was associated with higher empathy levels, specific findings differed between participant sex and age groups. The results of the population-based analysis suggested that *AVPR1A RS3* variation was associated with differences in cognitive empathy skill in men and in middle-aged and older women. However, this effect was not consistently replicated across all male participant groups. The family-based analysis produced limited evidence of an association between self-reported empathy and *AVPR1A RS3* genotype in young women. While the evidence presented in Chapter 4 must be interpreted cautiously, the significant findings aligned with previous research on the relationship between *AVPR1A RS3* and empathy-related behaviours (Avinun et al., 2012; Knafo, Israel, et al., 2008; Wang et al., 2016). The results of this study, combined with evidence from the literature, suggest a

relationship between *AVPR1A RS3* and empathy that is worthy of further investigation. Potential approaches to exploring this relationship are discussed below.

To clarify the interpretation of the results from Chapter 4 and to facilitate future research, the final study (Chapter 5) aimed to determine whether three previously unreported short-form empathy measures (the EQ-17, RMET-17, and RMET-14) displayed similar psychometric properties to those previously reported for the original full-length scales (the EQ and the RMET; Baron-Cohen & Wheelwright, 2004; Baron-Cohen, Wheelwright, Hill, et al., 2001). Overall, the psychometric properties of the EQ-17, RMET-17 and RMET-14 were similar to the properties of the full-length scales. The EQ-17 displayed the expected factor structure, internal consistency and sex differences in performance, suggesting that the scale is appropriate for research use. The RMET was also robust to changes in scale length; both the RMET-17 and RMET-14 performed similarly to the full-length test. However, the results of this study and previous research indicate that the RMET has poor internal consistency and an unclear factor structure (Olderbak et al., 2015; Ragsdale & Foley, 2011). The poor reliability and internal consistency of the RMET short-forms may have contributed to the variability of genetic associations with cognitive empathy observed in the results of Chapter 4. Caution is recommended when interpreting the results of this test and its associations with other variables.

Comparison with Previous Research

Mental and physical health. Much previous research has focused on understanding the role of vasopressin and the *AVPR1A* gene in human social behaviour (Ebstein et al., 2012). The results of this project add new evidence indicating that the *AVPR1A RS3* variant is also associated with individual differences in women's mental and physical health. While, to the best of the author's knowledge, this is the first study to directly link *AVPR1A RS3*

variation to individual differences in mental health, it adds to a body of evidence suggesting that vasopressin has an important influence on mental health. There has been some interest in determining the role of vasopressin, and the closely related neuropeptide oxytocin, in anxiety, depression, and their treatment (Meyer-Lindenberg, Domes, Kirsch, & Heinrichs, 2011; Neumann & Landgraf, 2012). Vasopressin is anxiogenic and overexpression of the neuropeptide is well associated with anxiety- and depression-like behaviours in rodents (Caldwell et al., 2008; Frank & Landgraf, 2008). In humans, high plasma vasopressin levels are associated with high levels of anxiety and with major depressive disorder (Caldwell et al., 2008). Major depressive disorder is also linked to high hypothalamic levels of vasopressin (Caldwell et al., 2008). Additionally, the gene encoding the vasopressin 1b receptor (*AVPR1B*) is associated with depression and certain genotypes at this locus may be protective against the development of major depressive disorder (Caldwell et al., 2008; Israel et al., 2008). The results of the present project add to this evidence, suggesting that *AVPR1A RS3* genotype is associated with women's overall mental health, as measured by the SF-12 (Ware Jr et al., 1996).

A novel association was found between *AVPR1A RS3* genotype and women's physical health. As discussed in Chapter 3, vasopressin has multiple functional roles, including regulating water homeostasis, and influencing hypothalamic-pituitary-adrenal (HPA) axis function and the physiological response to chronic stress (Frank & Landgraf, 2008; Kormos & Gaszner, 2013; Thibonnier et al., 2000). Vasopressin's role in stress physiology is likely primarily mediated by the V1-type receptors, including V1a which is expressed in the smooth muscle of blood vessels, suggesting the *AVPR1A* gene may influence blood pressure (Kormos & Gaszner, 2013; Thibonnier et al., 2000). Although one previous study suggests that *AVPR1A* genotype is not associated with blood pressure (Thibonnier et al., 2000), the association between the *AVPR1A RS3* polymorphism and this phenotype was not specifically

tested. The present study provides evidence that *AVPR1A RS3* variation is directly related to self-reported physical and mental health in women. The mechanism for this association remains unclear, and evidence from the literature suggests that it will be beneficial to investigate potential physiological and social trait pathways in future.

The relationship between *AVPR1A RS3* and women's physical and mental health emphasises the importance of continuing to understand the role of this gene in human social behaviour. Evidence from the literature suggests that exogenous vasopressin can alter people's neural, emotional, physiological, and behavioural responses to social stimuli, including positive and negative social interaction (Brunnlieb et al., 2016; Rilling et al., 2014; Rilling et al., 2012). The neuropeptide may alter the salience of positive and negative social stimuli and modulate emotional responses and regulation after exposure to social stimuli (Feng, DeMarco, et al., 2015; Feng, Hackett, et al., 2015; Rilling et al., 2012). Vasopressin may play a particularly important role in responding to negative social stimuli and administration of the neuropeptide has been associated with increased autonomic responsiveness to social stressors, such as evaluative threat (Shalev et al., 2011; Thompson et al., 2006). Additionally, *AVPR1A* genotype may influence approach and avoidance behavioural traits, and has been associated with individual differences in response to threatening stimuli (Meyer-Lindenberg et al., 2009; Reuter et al., 2015; Vogel et al., 2012).

The results of the present project, reported in Chapter 3, suggest that *AVPR1A RS3* genotype may moderate the impact of social interaction (both positive and negative) on mental and physical health. Using a larger sample size, a replication of the study reported in Chapter 3 may clarify the pattern of moderation effects. A better understanding of the pattern of moderation effects may subsequently clarify the mechanism via which *AVPR1A RS3* genotype modulates the impact of positive and negative social interaction (e.g. by affecting the salience of social stimuli or by influencing how emotional responses to negative

interactions are regulated). The interpretation of moderation effects may be complicated by characteristics of the social interaction (e.g. source of interaction) and the individual (e.g. personality) which can also modulate the effect of social interaction on health (Lee & Szinovacz, 2016; Lyu & Agrigoroaei, 2017; Santini et al., 2016; Schuster et al., 1990; Seeman et al., 2014; Teo et al., 2013; Yang et al., 2014). Future research should seek to determine how *RS3* allele length affects the impact of social interaction on health, whether moderation effects differ depending on the type (positive or negative) or source (family, friend, spouse etc) of social interaction, and whether effects differ depending on participant personality traits, sex, or gender.

Relationship quality. In Chapter 2 it was suggested that variation in the *AVPR1A* gene may influence health and wellbeing via an impact on the quality of social relationships, particularly spousal relationships. Evidence from a small number of studies suggests that *AVPR1A* variation is associated with pair-bonding and spousal relationship satisfaction (Maher et al., 2011; Walum et al., 2008). For men, carrying one or two copies of the 334bp target allele of the *AVPR1A RS3* variant is associated with lower scores on a measure of partner bonding (Walum et al., 2008). In the present project, no significant direct associations between *AVPR1A RS3* genotype and dyadic adjustment, positive social interaction, or negative social interaction were found for male or female participants (Appendices A and C). These results do not support the hypothesis that *AVPR1A RS3* genotype affects the quality of social relationships. Thus, *AVPR1A RS3* genotype may be more likely to impact health via the other pathways discussed in this thesis (e.g. modulating reactivity to social stimuli or physiological pathways).

Empathy. Individual differences in cognitive and emotional empathy present a potential link between *AVPR1A RS3* and health. People with high levels of empathy tend to be sensitive to other people's emotions, including their distress, thus they may be more

sensitive to the impacts of positive and negative social interaction (Smith & Rose, 2011; Tone & Tully, 2014). In the present project, evidence was mixed for an association between *AVPR1A RS3* and individual differences in empathy. One previous study suggests that empathy, particularly cognitive empathy, is associated with variation at the *AVPR1A RS3* locus (Uzefovsky et al., 2015). This is supported by evidence linking *AVPR1A* variants with empathy-related behaviours such as altruism and maternal sensitivity (Avinun et al., 2012; Avinun et al., 2011; Bisceglia et al., 2012; Knafo, Israel, et al., 2008; Wang et al., 2016). Although the results of Chapter 4 were not conclusive, the significant associations found were in the expected direction, that is, shorter *RS3* allele lengths were associated with lower cognitive empathy test scores and lower levels of self-reported empathy. Combined with the results of Chapter 3, these results suggest that women carrying the Short *AVPR1A RS3* allele may have a lower tendency to empathise and may be less sensitive to the negative effects of social interaction. Previous evidence suggests that individual differences in empathy can alter the impact of social interaction on a person's mental health (Chow et al., 2013; Smith & Rose, 2011; Tone & Tully, 2014). Thus, continuing to investigate the relationship between *AVPR1A RS3* and empathy could improve understanding of how genetic variation at this locus influences social behaviour and health.

Future research seeking to improve understanding of this relationship wold benefit from the use of valid and reliable measures of cognitive and emotional empathy. Chapter 5 contributes to the evidence for the reliability and validity of two commonly used empathy measures: the EQ and the RMET (Baron-Cohen & Wheelwright, 2004; Baron-Cohen, Wheelwright, Hill, et al., 2001). As discussed above, the findings of Chapter 5 support existing evidence that the EQ is a valid and reliable measure of self-reported empathy (Lawrence et al., 2004; Muncer & Ling, 2006). The findings do not support the reliability and validity of the RMET short-forms used in this project: the RMET-17 and RMET-14. While

the psychometric properties of these short-form measures were very similar to those previously reported for the full-length test, the results presented here suggest that they are not suitable for research use (Olderbak et al., 2015; Ragsdale & Foley, 2011; Vellante et al., 2013). Together, these findings call the validity of the RMET into question. The test was originally designed to be a single-factor measure of cognitive empathy ability but three studies, including the present project, have failed to find a single-factor solution (Olderbak et al., 2015; Ragsdale & Foley, 2011). Previous literature has suggested that the test is not a single-factor measure of cognitive empathy and may instead be measuring multiple underlying constructs or rely on multiple cognitive abilities (Ragsdale & Foley, 2011). Further research is required to confirm what trait, or traits, the RMET is measuring.

Sex and gender differences. The findings of the project emphasise the importance of considering and accounting for the impact of participant sex and/or gender when studying the role of *AVPR1A* in social traits, social behaviour, and health. In Chapter 3, it was hypothesised that the effects of *AVPR1A RS3* genotype would be stronger for male participants. While some evidence of moderation effects was found for male participants, a direct relationship between *AVPR1A RS3* and health was only found for female participants. Vasopressin and the *AVPR1A* gene have previously been associated with social behaviours and traits in female participants (e.g. Bisceglia et al., 2012; Feng, Hackett, et al., 2015; Rilling et al., 2014). However, in the literature, vasopressin has often been explored in the context of male social behaviour, particularly in neuropeptide administration studies. In Chapter 4, sex differences were found when exploring the relationship between *AVPR1A RS3* variation and individual differences in empathy. In this case, significant associations were found for male and female participants. Although the direction of the relationships was the same, the statistical significance of the relationships differed between the sexes and between age groups. Sex and gender differences in the behavioural effects of vasopressin and the

phenotypic associations of *AVPR1A* highlight the importance of including both male and female participants in studies exploring vasopressin and its receptors, and the importance of testing for sex or gender differences in this area of research (Albers, 2012; Dumais & Veenema, 2016; Feng, Hackett, et al., 2015; McCall & Singer, 2012).

Limitations of the Project

The results of this project should be interpreted in the context of the following limitations. As discussed above, the RMET-17 and RMET-14 were found to have poor internal consistency and test-retest reliability. The small number of factor analyses that have been conducted, including Chapter 5, also call into question the validity of the claim that the RMET is a single-factor measure of cognitive empathy ability (Olderbak et al., 2015; Ragsdale & Foley, 2011). The results of work using this test, including the work reported in Chapter 4, should be interpreted cautiously until the underlying constructs measured by the RMET have been determined.

Splitting tandem repeat variation at the *AVPR1A RS3* locus into Long and Short repeat-length-based categories is a limited representation of the full range of length variation present in the population. This approach is necessitated by our current limited understanding of how alleles at the *RS3* locus interact, which prevents genotype from being validly modelled as a continuous length-based variable. Thus, binning alleles into a small number of categories is necessary to preserve power by reducing the number of genotype categories modelled in an analysis. As was done in this project, allele categories are often selected to produce relatively even genotype group sample sizes (e.g. Knafo, Israel, et al., 2008). Despite efforts to produce relatively even groups, analyses were limited by the size of the least common genotype group, Short/Short. In Chapter 3, the effective sample size limited the number of moderation effects that could be tested in a single multi-level model. In Chapter 4, the smaller overall sample

size limited the power available to detect small effects and conclusive evidence about the relationship between *AVPR1A RS3* and empathy was not obtained.

Future Research

Direct extensions of the project. While convergent evidence supports the direct relationship between *AVPR1A RS3* and women's physical and mental health, it is a novel finding (to the best of the author's knowledge). Thus, it is important that this finding is replicated in an independent sample of women (Tabor et al., 2002). Additionally, with an adequate sample size future research could confirm whether *AVPR1A RS3* genotype does moderate the impact of positive and negative social interaction on physical and mental health. To interpret gene-environment interactions of interest, all potential moderation effects should be tested in a single model.

In the present project, a direct association between *AVPR1A RS3* genotype and positive or negative social interaction was not found. Thus, social interaction could not be tested as a potential mediator of the relationship between *AVPR1A* and health. In Chapter 3, two potential pathways linking *AVPR1A* to mental and physical health were suggested: (1) *AVPR1A* could impact on health via its role in social behaviour and cognition, or (2) *AVPR1A* could impact on mental and physical health via its role in homeostatic maintenance and HPA-axis regulation. Chapter 4 explored the first potential pathway, aiming to determine if there was a relationship between *AVPR1A RS3* and empathy. The second pathway was not explored in this project. As vasopressin's role in the physiological response to stress is likely mediated by the V1-type receptors expressed in the smooth muscle of blood vessels (Kormos & Gaszner, 2013; Thibonnier et al., 2000), determining whether *AVPR1A* variation is associated with individual differences in blood pressure is suggested as a logical starting point. This association could be tested in a sample of women from the same age group as the

current study (20 to 30 years-old), and women in middle and older age, controlling for the use of blood pressure lowering medication and other relevant health factors. If a relationship between *AVPR1A RS3* and blood pressure is found, a mediation model of the pathway from genotype through blood pressure to health could subsequently be tested.

While the findings of Chapter 4 were inconclusive, evidence from the literature suggests the empathy pathway is also a productive area for further exploration (Avinun et al., 2012; Knafo, Israel, et al., 2008; Tone & Tully, 2014; Uzefovsky et al., 2015; Wang et al., 2016). The two key limitations of Chapter 4 were the small sample size and the empathy measures available in the data set. Addressing these limitations could greatly improve our understanding of the role of *AVPR1A RS3* in empathy. Future work exploring this relationship should endeavour to include valid measures of both self-reported empathy and objectively measured empathy ability or sensitivity, as the results of self-report and objective measures of empathy can differ. For example, sex differences are commonly found on self-report empathy measures, but not measures of ability (Maibom, 2012). It will also be important to use measures that allow for the clear separation of the cognitive and emotional domains of empathy, as defined in Chapter 2.

The present project and the work conducted by Uzefovsky and colleagues (2015) both contain measures of general tendency to empathise and cognitive empathy, but lack a clear and valid measurement of emotional empathy. Well-validated self-report measures able to separately tap the constructs of emotional and cognitive empathy are currently available (e.g. the Basic Empathy Scale; Jolliffe & Farrington, 2006). Emotional empathy sensitivity could be objectively tested by measuring physiological responsiveness to empathy-provoking stimuli, such as pictures of emotional expressions or situations. Physiological measures such as heart rate or electrodermal (skin) conductance are commonly used to indicate emotional arousal (Blair, 1999; Blair, Jones, Clark, & Smith, 1997; Eisenberg & Fabes, 1990; Eisenberg

et al., 1988). While the results of this project suggest that the RMET may not be a valid test of cognitive empathy, there are similar alternatives available. For example, the matching and labelling tasks described by Palermo and colleagues (2013) test the perception and identification of facial expressions for the six basic emotions (i.e. happiness, sadness, anger, disgust, fear and surprise). Like the RMET, these tasks are relatively easy to administer, and they can detect individual differences in expression recognition in the general population (Palermo et al., 2013). In addition to using valid measures of empathy, future research will need to gather sufficient sample sizes to detect the small effects of a single genetic variant, taking into account the relatively low frequency of the Short/Short *AVPR1A RS3* genotype.

Broader issues to be covered in future work. Additional lines of research have the potential to support future candidate gene investigations of the relationship between the *AVPR1A* gene and social traits, social behaviour, and physical and mental health. A major limitation of studies involving *AVPR1A RS3* is the large number of alleles at this locus (Prichard et al., 2007), which results in the need to bin alleles into broad length-based categories to provide adequate statistical power. Combining alleles to reduce the number of genotype categories may, however, mask the effects of specific alleles or genotypes because the relationship between alleles at the *RS3* locus is not currently understood. If allele interaction at this locus was better understood it may be possible to model *AVPR1A RS3* genotype as a continuous variable (e.g. if a codominant relationship exists, genotype could be modelled as the average length of two alleles in base pairs). A continuous representation of genotype could provide a more valid measure of the genetic variation present in the population, compared to the current category-based methods.

Understanding the functional consequences of *AVPR1A RS3* genotype is further complicated by the features of the *RS3* variant. *AVPR1A RS3* is a complex short tandem repeat variation; it has two dinucleotide repeats which can both vary in length (Prichard et al.,

2007). The allele categories used in this project and in previous research assume that it is the total length of the *RS3* variant which has functional consequences. The two dinucleotide repeats may instead have different impacts on gene function, or different strengths of influence. Currently, work to improve our understanding of the *AVPR1A RS3* locus is hampered by the limitations of sequencing technology (Bakhtiari et al., 2018). The identification and genotyping of repetitive sequences of DNA, including STRs, is computationally complex and presents multiple challenges for current variant and alignment calling tools (Bakhtiari et al., 2018). Individually accounting for the two dinucleotide repeats in the *AVPR1A RS3* variant may increase the number of allele and genotype categories that must be modelled, reducing group size and power to detect statistical effects. However, a better understanding of the functional consequences of genetic variation at this locus may improve our ability to accurately model and test the phenotypic correlates of *AVPR1A RS3*.

Conclusion

Vasopressin, the V1a receptor, and the *AVPR1A* gene are implicated in a variety of social behaviours and cognitive processes in humans and non-human mammals (Albers, 2012; Caldwell et al., 2008; Donaldson & Young, 2008; Hammock & Young, 2005; McCall & Singer, 2012). Positive social interaction and social support, and negative social interaction and social strain are well-associated with physical and mental health outcomes (Ditzen & Heinrichs, 2014; Lee & Szinovacz, 2016; Seeman et al., 2014). In the present project, genetic variation in the 5' promoter region of the *AVPR1A* gene was found to be directly associated with individual differences in women's mental and physical health. The results suggest that *AVPR1A RS3* genotype may moderate the effects of positive and negative social interaction on mental and physical health. *AVPR1A RS3* may also be associated with individual

differences in empathy, a trait that can influence a person's sensitivity to social stimuli (Smith & Rose, 2011; Tone & Tully, 2014). The dual role of *AVPR1A* in social traits and health suggests that this gene is a prime target for research seeking to understand the mechanisms via which social interaction impacts on health. The impact of social support and social strain varies between individuals (Ditzen & Heinrichs, 2014) and biological factors, including *AVPR1A RS3* genotype, may underlie some of this variation. There is already interest in the potential therapeutic applications of vasopressin for mental disorders involving social deficits (Bartz & Hollander, 2008; Finger, 2011; McGregor & Bowen, 2012; Modi & Young, 2012; Striepens et al., 2011). Pursuing a better understanding of the role of the *AVPR1A* gene in social behaviour and health may contribute to the future development of interventions for psychosocial difficulties tailored to individual differences in sensitivity to social interactions.

Appendices

Appendix A: Testing for a direct association between *AVPR1A RS3* genotype and Social Interaction variables.

Methods

Participants were from the 20+ cohort of the PATH data set (see Chapter 3; Anstey et al., 2012). A series of one-way ANOVAs were conducted to determine whether Positive Social Interaction or Negative Social Interaction scores differed between *AVPR1A RS3* genotype groups. Separate ANOVAs were performed for men and women, and for each wave of data collection. Positive Social Interaction and Negative Social Interaction were measured using the composite scores described in Chapter 3, calculated from the scales developed by Schuster and colleagues (1990).

Results

Table A1. Women: Descriptive statistics for Social Interaction variables, split by *AVPR1A RS3* genotype group and Wave.

	Wave 1			Wave 2			Wave 3		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD
Positive Social Interaction									
Short/Short	177	50.23	6.61	164	51.11	6.89	157	50.84	7.14
Short/Long	564	50.76	7.43	522	51.28	6.66	493	50.75	7.33
Long/Long	427	50.97	7.28	386	50.97	7.54	361	51.53	7.16
Negative Social Interaction									
Short/Short	177	52.21	6.84	164	50.12	6.96	157	48.89	7.34
Short/Long	564	51.52	7.93	522	49.84	8.02	493	48.78	7.33
Long/Long	427	51.94	8.32	386	50.13	8.10	361	48.62	7.27

Table A2. Men: Descriptive statistics for Social Interaction variables, split by *AVPR1A RS3* genotype group and Wave.

	Wave 1			Wave 2			Wave 3		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD
Positive Social Interaction									
Short/Short	180	49.91	7.36	159	49.31	7.50	149	47.74	8.96
Short/Long	536	48.29	9.10	467	48.85	8.26	419	48.31	8.05
Long/Long	386	49.12	8.58	340	49.52	7.62	310	48.97	7.52
Negative Social Interaction									
Short/Short	180	50.71	7.67	159	49.76	7.86	149	47.77	7.92
Short/Long	536	52.04	8.47	467	49.89	7.90	419	48.91	7.82
Long/Long	386	51.74	8.44	340	49.88	7.95	310	48.76	6.89

Table A3. Women: One-way ANOVA of the effect of *AVPR1A RS3* genotype groups on Social Interaction, conducted separately for each Wave.

Participant Group	Source of Variance	SS	*df*	MS	*F*	Partial η^2
Dependent Variable: Positive Social Interaction						
Wave 1	*AVPR1A RS3* genotype	68.269	2	34.134	0.648	0.001
	Error	61376.406	1165	52.684		
Wave 2	*AVPR1A RS3* genotype	21.092	2	10.546	.214	<0.001
	Error	52707.570	1069	49.305		
Wave 3	*AVPR1A RS3* genotype	134.462	2	67.231	1.282	0.003
	Error	52858.540	1008	52.439		
Dependent Variable: Negative Social Interaction						
Wave 1	*AVPR1A RS3* genotype	83.344	2	41.672	.664	0.001
	Error	73162.756	1165	62.801		
Wave 2	*AVPR1A RS3* genotype	21.439	2	10.720	.172	<0.001
	Error	66651.047	1069	62.349		
Wave 3	*AVPR1A RS3* genotype	9.296	2	4.648	.087	<0.001
	Error	53872.177	1008	53.445		

Note: Genotype groups are coded as Short/Short, Short/Long and Long/Long. The effect of genotype group did not approach significance at the $p < 0.05$ level in any model.

Table A4. Men: One-way ANOVA of the effect of *AVPR1A RS3* genotype groups on Social Interaction, conducted separately for each Wave.

Participant Group	Source of Variance	SS	*df*	MS	*F*	Partial η^2
Dependent Variable: Positive Social Interaction						
Wave 1	*AVPR1A RS3* genotype	394.188	2	197.094	2.631	0.005
	Error	82336.388	1099	74.919		
Wave 2	*AVPR1A RS3* genotype	93.405	2	46.703	.745	0.002
	Error	60339.291	963	62.658		
Wave 3	*AVPR1A RS3* genotype	166.168	2	83.084	1.287	0.003
	Error	56472.033	875	64.539		
Dependent Variable: Negative Social Interaction						
Wave 1	*AVPR1A RS3* genotype	240.056	2	120.028	1.728	0.003
	Error	76345.083	1099	69.468		
Wave 2	*AVPR1A RS3* genotype	2.223	2	1.111	.018	<0.001
	Error	60264.127	963	62.580		
Wave 3	*AVPR1A RS3* genotype	147.952	2	73.976	1.308	0.003
	Error	49491.528	875	56.562		

Note: Genotype groups are coded as Short/Short, Short/Long and Long/Long. The effect of genotype group did not approach significance at the $p < 0.05$ level.

Appendix B: Testing association between Behavioural Inhibition System (BIS) sensitivity and *AVPR1A RS3* genotype.

Methods

Participants were from the 20+ cohort of the PATH data set (see Chapter 3; Anstey et al., 2012). One-way ANOVAs were used to test whether BIS sensitivity differed between *AVPR1A RS3* genotype groups. BIS sensitivity was measured using the BIS/BAS scales developed by Carver and White (1994). Separate ANOVAs were performed for men and women.

Results

Table A5. Descriptive statistics for BIS sensitivity in Wave 1, split by *AVPR1A RS3* genotype group and gender.

	Women			Men		
Genotype	N	Mean	SD	N	Mean	SD
Short/Short	176	21.688	3.142	178	19.163	3.580
Short/Long	562	21.737	3.255	532	19.066	3.439
Long/Long	426	21.608	3.324	384	19.180	3.573

Note: BIS sensitivity was measured using the BIS/BAS scales developed by Carver and White (1994).

Table A6. One-way ANOVA of the effect of *AVPR1A RS3* genotype groups on Behavioural Inhibition System sensitivity

Participant Group	Source of Variance	SS	*df*	MS	*F*	Partial η^2
Men	*AVPR1A RS3* genotype	3.257	2	1.628	0.132	<0.001
	Error	13435.574	1091	12.315		
Women	*AVPR1A RS3* genotype	4.018	2	2.009	0.189	<0.001
	Error	12368.370	1161	10.653		

Note: Dependent variable is BIS sensitivity (Carver & White, 1994) measured in Wave 1. Genotype groups are coded as Short/Short, Short/Long and Long/Long. The effect of genotype group did not approach significance at the $p < 0.05$ level.

Appendix C: Testing association between *AVPR1A RS3* genotype and individual differences in Dyadic Adjustment

Methods

Participants were from the 20+ cohort of the PATH data set (see Chapter 3; Anstey et al., 2012). A one-way ANOVA was used to test whether Dyadic Adjustment differed between *AVPR1A RS3* genotype groups. Dyadic Adjustment was measured using the DAS-7 (Hunsley et al., 2001), which was administered in Wave 3 to participants currently in a romantic relationship.

Results

Table A7. Descriptive statistics for Dyadic Adjustment, split by *AVPR1A RS3* genotype group and gender.

	Women			Men		
Genotype	N	Mean	SD	N	Mean	SD
Short/Short	123	25.58	4.298	96	25.44	4.220
Short/Long	331	25.83	4.658	268	24.64	4.628
Long/Long	251	25.96	4.290	202	24.66	4.743

Note: Dyadic Adjustment was measured using the DAS-7 (Hunsley et al., 2001).

Table A8. One-way ANOVA of the effect of *AVPR1A RS3* genotype groups on Dyadic Adjustment

Participant Group	Source of Variance	SS	*df*	MS	*F*	Partial η^2
Men	*AVPR1A RS3* genotype	49.886	2	24.943	1.177	0.004
	Error	11932.947	563	21.195		
Women	*AVPR1A RS3* genotype	12.111	2	6.056	0.303	0.001
	Error	14014.802	702	19.964		

Note: Dependent variable is DAS-7 score (Hunsley et al., 2001). Genotype groups are coded as Short/Short, Short/Long and Long/Long. The DAS-17 was administered to participants currently in a romantic relationship in Wave 3. The effect of genotype group did not approach significance at the $p < 0.05$ level.

Appendix D: Testing the association between a target *AVPR1A RS3* allele (allele 20) and self-reported physical and mental health.

Methods

Participants were from the 20+ cohort of the PATH data set (see Chapter 3; Anstey et al., 2012). The series of nested multi-level linear models from Chapter 3 examining the repeated cross-sectional effects of *AVPR1A RS3* genotype, Positive Social Interaction, Negative Social Interaction and Time (IVs) on Physical Health and Mental Health (DVs) at each measurement occasion were repeated, using genotype categories based on a target allele (allele 20) in place of the length-based categories used in the original analysis. Participants with one or two copies of allele 20 were categorised as "*RS3* Allele 20 Present" and participants with no copies of allele 20 as "*RS3* Allele 20 Absent". The binary genotype categories were necessary to compensate for the small number of participants homozygous for allele 20. Separate analyses were conducted for each DV and for men and women using the following procedure.

Nested multi-level linear models were analysed using maximum likelihood estimation with a Gaussian distribution. Starting from an intercepts-only model, predictors were individually introduced and their effects analysed in subsequent models (Singer & Willet, 2003). All models included a fixed and random intercept. A 10000-sample stratified bootstrap resampling of individuals was conducted for each model to produce 95% percentile confidence intervals (CIs) around fixed parameter estimates. A Wald Z test was conducted to produce 95% percentile CIs around random parameter estimates. The nested models examined the cross-sectional association between the independent variable (IV) and the dependent variable (DV) across multiple periods of time. Model 1 was a random intercepts

(variance components) model with no predictors used to calculate the intra-class correlation coefficient, which represented the proportion of variance in the DV (Mental Health or Physical Health) due to variation between individuals. Model 2 included the fixed effect of *AVPR1A RS3* genotype. As *RS3* Allele 20 Absent was the reference category, the intercept of Model 2 represents average health of individuals in the *RS3* Allele 20 Absent group. Model 3 included the fixed effects of *AVPR1A RS3* and Time. Model 4 included the fixed effects of *AVPR1A RS3*, Time, Positive Social Interaction and Negative Social Interaction. To compare the nested models, three information criteria statistics are reported: -2 Log Likelihood (-2LL), Akaike's Information Criterion (AIC), and Schwarz's Bayesian Criterion (BIC). These abbreviations are used in the results section and information criteria are displayed in a smaller-is-better form.

Results

Table A9. *AVPR1A RS3* Allele 20: Female mental health models

	Parameter				95% Confidence Interval		Information Criteria		
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		45.967	<.001	45.758	46.147	24723.500	24729.500	24747.900
		Intercept	39.697	<.001	35.004	45.019			
		Residual	56.476	<.001	53.206	59.946			
Model 2	Intercept		46.121	<.001	45.847	46.355	23538.306	23546.306	23570.647
	RS3 Allele 20[a] Present		-0.243	.194	-0.629	0.178			
		Intercept	39.499	<.001	34.729	44.923			
		Residual	56.052	<.001	52.734	59.579			
Model 3	Intercept		45.688	<.001	45.302	46.035	23530.965	23542.965	23579.477
	RS3 Allele 20[a] Present		-0.246	.188	-0.633	0.176			
	Wave 2[b]		0.527	.052	0.029	1.049			
	Wave 3[b]		**0.871**	**.002**	**0.349**	**1.400**			
		Intercept	39.508	<.001	34.744	44.925			
		Residual	55.883	<.001	52.575	59.400			
Model 4	Intercept		46.379	<.001	43.409	49.348	23011.684	23027.684	23076.366
	RS3 Allele 20[a] Present		-0.236	.188	-0.640	0.145			

	Parameter				95% Confidence Interval			Information Criteria	
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Wave 2[b]		-0.150	.560	-0.596	0.358			
	Wave 3[b]		-0.126	.645	-0.585	0.427			
	Positive Social		**0.319**	**<.001**	**0.270**	**0.345**			
	Negative Social		**-0.326**	**<.001**	**-0.349**	**-0.281**			
		Intercept	24.603	<.001	21.118	28.663			
		Residual	52.017	<.001	48.940	55.287			

Note: a. Reference category is *RS3* Allele 20 Absent.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except fixed intercepts). Information criteria are displayed in smaller-is-better form.

Table A10. *AVPR1A RS3* Allele 20: Male mental health models

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		48.858	<.001	48.681	49.035	21782.697	21788.697	21806.793
		Intercept	40.593	<.001	36.018	45.749			
		Residual	43.751	<.001	41.070	46.607			
Model 2	Intercept		48.682	<.001	48.446	48.916	20794.733	20802.733	20826.673
	RS3 Allele 20[a] Present		0.348	.039	-0.017	0.703			
		Intercept	39.785	<.001	35.156	45.022			
		Residual	44.268	<.001	41.494	47.228			
Model 3	Intercept		48.983	<.001	48.646	49.304	20790.956	20802.956	20838.867
	RS3 Allele 20[a] Present		0.346	.041	-0.018	0.700			
	Wave 2[b]		-0.463	.069	-0.926	0.026			
	Wave 3[b]		**-0.539**	**.043**	**-1.050**	**-0.051**			
		Intercept	39.851	<.001	35.220	45.092			
		Residual	44.165	<.001	41.397	47.118			
Model 4	Intercept		49.538	<.001	46.984	52.104	20342.198	20358.198	20406.079
	RS3 Allele 20[a] Present		0.172	.293	-0.188	0.514			

	Parameter				95% Confidence Interval			Information Criteria	
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Wave 2[b]		**-1.067**	**<.001**	**-1.479**	**-0.575**			
	Wave 3[b]		**-1.232**	**<.001**	**-1.700**	**-0.726**			
	Positive Social		**0.287**	**<.001**	**0.241**	**0.307**			
	Negative Social		**-0.280**	**<.001**	**-0.300**	**-0.236**			
		Intercept	25.246	<.001	21.828	29.199			
		Residual	41.818	<.001	39.188	44.625			

Note: a. Reference category is *RS3* Allele 20 Absent.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except fixed intercepts). Information criteria are displayed in smaller-is-better form.

Table A11. *AVPR1A RS3* Allele 20: Female physical health models

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		50.588	<.001	50.255	50.921	23086.174	23092.174	23110.564
		Intercept	21.578	<.001	18.853	24.698			
		Residual	37.313	<.001	35.157	39.601			
Model 2	Intercept		50.807	<.001	50.593	50.995	21955.360	21963.360	21987.687
	RS3 Allele 20[a] Present		**-0.474**	**.003**	**-0.798**	**-0.117**			
		Intercept	21.298	<.001	18.549	24.453			
		Residual	36.835	<.001	34.659	39.147			
Model 3	Intercept		50.990	<.001	50.691	51.277	21951.197	21963.197	21999.688
	RS3 Allele 20[a] Present		**-0.471**	**.003**	**-0.795**	**-0.113**			
	Wave 2[b]		-0.089	.684	-0.492	0.311			
	Wave 3[b]		**-0.511**	**.026**	**-0.956**	**-0.109**			
		Intercept	21.351	<.001	18.600	24.508			
		Residual	36.748	<.001	34.577	39.055			
Model 4	Intercept		54.034	<.001	51.321	56.428	21816.577	21832.577	21881.231
	RS3 Allele 20[a] Present		**-0.453**	**.004**	**-0.794**	**-0.107**			

	Parameter				95% Confidence Interval			Information Criteria	
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Wave 2[b]		-0.404	.058	-0.771	0.017			
	Wave 3[b]		**-0.990**	**<.001**	**-1.398**	**-0.537**			
	Positive Social		**0.103**	**<.001**	**0.059**	**0.126**			
	Negative Social		**-0.160**	**<.001**	**-0.175**	**-0.118**			
		Intercept	17.612	<.001	15.137	20.491			
		Residual	36.713	<.001	34.540	39.024			

Note: a. Reference category is *RS3* Allele 20 Absent.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except fixed intercepts). Information criteria are displayed in smaller-is-better form.

Table A12. *AVPR1A RS3* Allele 20: Male physical health models

Model	Parameter		Estimate	Sig.	95% Confidence Interval		Information Criteria		
	Fixed	Random			Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 1	Intercept		52.169	<.001	52.018	52.288	19832.168	19838.168	19856.259
		Intercept	15.171	<.001	13.138	17.519			
		Residual	26.344	<.001	24.728	28.065			
Model 2	Intercept		52.213	<.001	50.593	50.995	18872.621	18880.621	18904.554
	RS3 Allele 20[a] Present		0.009	.949	-0.247	0.323			
		Intercept	15.100	<.001	13.039	17.488			
		Residual	25.811	<.001	24.190	27.542			
Model 3	Intercept		52.334	<.001	52.062	52.562	18865.896	18877.896	18913.797
	RS3 Allele 20[a] Present		0.009	.947	-0.248	0.322			
	Wave 2[b]		0.070	.708	-0.301	0.412			
	Wave 3[b]		**-0.503**	**.011**	**-0.900**	**-0.157**			
		Intercept	15.182	<.001	13.117	17.573			
		Residual	25.695	<.001	24.080	27.418			
Model 4	Intercept		54.553	<.001	52.696	56.385	18764.863	18780.863	18828.731
	RS3 Allele 20[a] Present		-0.055	.676	-0.313	0.260			

	Parameter				95% Confidence Interval			Information Criteria	
Model	Fixed	Random	Estimate	Sig.	Lower Bound	Upper Bound	-2LL	AIC	BIC
Model 4	Wave 2[b]		-0.164	.383	-0.519	0.191			
	Wave 3[b]		**-0.816**	**<.001**	**-1.206**	**-0.442**			
	Positive Social		**0.079**	**<.001**	**0.049**	**0.093**			
	Negative Social		**-0.117**	**<.001**	**-0.133**	**-0.086**			
		Intercept	12.829	<.001	10.938	15.048			
		Residual	25.769	<.001	24.144	27.503			

Note: a. Reference category is *RS3* Allele 20 Absent.

b. Reference category is Wave 1

Bootstrapping used to produce 95% CIs around fixed effects, Wald Z statistics are presented for random effects. Significant fixed parameter estimates at the $p < 0.05$ level are bolded (except fixed intercepts). Information criteria are displayed in smaller-is-better form.

Appendix E – Item lists for the EQ-17, RMET-17 and RMET-14

Table A13. The Empathy Quotient-17 questionnaire

Factor	No.	Item
Cognitive Empathy	1	I can easily tell if someone else wants to enter a conversation.
	6	I can pick up quickly if someone says one thing but means another.
	8	I find it easy to put myself in somebody else's shoes.
	9	I am good at predicting how someone will feel.
	10	I am quick to spot when someone in a group is feeling awkward or uncomfortable.
	15	Friends usually talk to me about their problems as they say that I am very understanding.
	17	I can tune into how someone else feels rapidly and intuitively.
Emotional Reactivity	5	In a conversation, I tend to focus on my own thoughts rather than on what my listener might be thinking.
	7	It is hard for me to see why some things upset people so much.
	11	If I say something that someone else is offended by, I think that that's their problem, not mine.
	12	I can't always see why someone should have felt offended by a remark.
	13	Seeing people cry doesn't really upset me.
	16	If I see a stranger in a group, I think that it is up to them to make an effort to join in.
Social Skills	2	I find it hard to know what to do in a social situation.
	3	Friendships and relationships are just too difficult, so I tend not to bother with them.
	4	I often find it difficult to judge if something is rude or polite.
	14	I don’t tend to find social situations confusing.

EQ-17 items presented in factor groups. Item numbers represent the order of items in the Twin Study (Hatemi et al., 2015) Wave 1 survey. Item 18 (“I don't consciously work out the rules of social situations.”) was excluded from the final model as it failed to load onto any factor.

Table A14. The Reading the Eyes in the Mind Test-17 and RMET-14 questionnaires

Item	Response Options	Correct Answer
1	Jealous, Panicked, Arrogant, Hateful	Panicked
2	Terrified, Upset, Arrogant, Annoyed	Upset
3	Joking, Insisting, Amused, Relaxed	Insisting
4	Aghast, Fantasising, Impatient, Alarmed	Fantasising
5	Irritated, Disappointed, Depressed, Accusing	Accusing
6	Irritated, Thoughtful, Encouraging, Sympathetic	Thoughtful
7	Decisive, Amused, Aghast, Bored	Decisive
8	Embarrassed, Fantasising, Confused, Panicked	Fantasising
9	Contented, Apologetic, Defiant, Curious	Defiant
10	Alarmed, Shy, Hostile, Anxious	Hostile
11	Joking, Cautious, Arrogant, Reassuring	Cautious
12	Interested, Insisting, Affectionate, Contented	Interested
13	Impatient, Aghast, Irritated, Reflective	Reflective
14	Serious, Ashamed, Bewildered, Alarmed	Serious
15	Aghast, Baffled, Distrustful, Terrified	Distrustful
16	Puzzled, Nervous, Insisting, Contemplative	Nervous
17	Ashamed, Nervous, Suspicious, Indecisive	Suspicious

Note: Items 1, 4, and 11 were not included in the RMET-14. The RMET-14 included an additional “Don’t know” response option for all items.

This page left intentionally blank.

Appendix F – Testing for an association between *AVPR1A RS3* and the EQ-17 and its subscales.

Methods

As was done in Chapter 4, an exploratory population-level analysis of the relationship between *AVPR1A RS3* genotype and the EQ-17 and its subscales was conducted in unrelated subsamples of participants. Participants were from a subset of the Brisbane Longitudinal Twin Study (see Chapter 4; Hatemi et al., 2015; Wright & Martin, 2004). Analyses were conducted separately for men and women as performance on empathy measures and the effects of *AVPR1A* genotype and vasopressin have been found to vary by sex (Albers, 2012; McCall & Singer, 2012; Shalev et al., 2011; Walum et al., 2008). Four linear regression models were tested to determine whether empathy scores differed significantly between *AVPR1A RS3* genotype groups. The independent variable for all three models was *AVPR1A RS3* genotype. Genotype categories (Short/Short, Short/Long and Long/Long) were dummy-coded with Long/Long as the reference category. The dependent variable differed between the regression models; each model tested the effect of genotype on the EQ-17 or one of its subscales (Cognitive Empathy, Emotional Reactivity, or Social Skills).

Twin pairs were split to produce two unrelated subsamples: Twin 1 and Twin 2. Regression analyses were conducted with the Twin 1 subsample. No significant results were found; thus, the analysis was not replicated in the Twin 2 subsample. Participants with missing genetic data were excluded from these analyses. Participants were also excluded from an analysis if they were missing the relevant phenotypic data for that analysis. Results are reported in Table A15.

Results

Table A15. Regression parameter estimates for association between *AVPR1A RS3* genotype and EQ-17, Cognitive Empathy, Emotional Reactivity, and Social Skills score.

					95% Confidence Interval	
Scale	Sex	Parameter	*B*	*p*-value	Upper Bound	Lower Bound
EQ-17	Female	Intercept	66.204	<0.001	64.064	68.345
		RS3 Short/Short	-1.130	0.535	-4.721	2.461
		RS3 Short/Long	-1.897	0.163	-4.568	0.774
	Male	Intercept	59.789	<0.001	56.023	63.556
		RS3 Short/Short	-1.335	0.669	-7.555	4.885
		RS3 Short/Long	0.089	0.970	-4.638	4.817
Cognitive Empathy	Female	Intercept	27.408	<0.001	26.389	28.427
		RS3 Short/Short	-1.001	0.249	-2.710	0.709
		RS3 Short/Long	-0.954	0.140	-2.225	0.318
	Male	Intercept	24.895	<0.001	22.963	26.826
		RS3 Short/Short	-0.986	0.539	-4.175	2.204
		RS3 Short/Long	-0.137	0.910	-2.562	2.287
Emotional Reactivity	Female	Intercept	23.020	<0.001	22.025	24.016
		RS3 Short/Short	0.165	0.846	-1.506	1.835
		RS3 Short/Long	-0.634	0.315	-1.877	0.608
	Male	Intercept	20.158	<0.001	18.541	21.774
		RS3 Short/Short	-0.885	0.510	-3.555	1.785
		RS3 Short/Long	-0.552	0.588	-2.581	1.477

Scale	Sex	Parameter	*B*	*p*-value	95% Confidence Interval	
					Upper Bound	Lower Bound
Social Skills	Female	Intercept	15.776	<0.001	14.965	16.586
		RS3 Short/Short	-0.294	0.670	-1.654	1.066
		RS3 Short/Long	-0.310	0.546	-1.321	0.702
	Male	Intercept	14.737	<0.001	13.388	16.085
		RS3 Short/Short	0.536	0.632	-1.691	2.763
		RS3 Short/Long	0.778	0.361	-0.914	2.471

Note: The reference category for all models is *AVPR1A RS3* Long/Long genotype, the intercept represents the mean of the Long/Long genotype group.

References

Adams, R. B., Rule, N. O., Franklin, R. G., Wang, E., Stevenson, M. T., Yoshikawa, S., . . . Ambady, N. (2010). Cross-cultural reading the mind in the eyes: An fMRI investigation. *Journal of cognitive neuroscience, 22*(1), 97-108. doi:10.1162/jocn.2009.21187

Ainsworth, M. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Oxford: Lawrence Erlbaum.

Albers, H. E. (2012). The regulation of social recognition, social communication and aggression: Vasopressin in the social behavior neural network. *Hormones and Behavior, 61*(3), 283-292. doi:10.1016/j.yhbeh.2011.10.007

Allemand, M., Steiger, A. E., & Fend, H. A. (2015). Empathy Development in Adolescence Predicts Social Competencies in Adulthood. *Journal of Personality, 83*(2), 229-241. doi:10.1111/jopy.12098

Anstey, K. J., Christensen, H., Butterworth, P., Easteal, S., Mackinnon, A., Jacomb, T., . . . Jorm, A. F. (2012). Cohort Profile: The PATH through life project. *International Journal of Epidemiology, 41*(4), 951-960. doi:10.1093/ije/dyr025

Avinun, R., Ebstein, R. P., & Knafo, A. (2012). Human maternal behaviour is associated with arginine vasopressin receptor 1A gene. *Biology Letters, 8*(5), 894-896. doi:10.1098/rsbl.2012.0492

Avinun, R., Israel, S., Shalev, I., Gritsenko, I., Bornstein, G., Ebstein, R. P., & Knafo, A. (2011). AVPR1A Variant Associated with Preschooler's Lower Altruistic Behavior. *PLoS ONE, 6*(9), e25274. doi:10.1371/journal.pone.0025274.t002

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2014). A sociability gene? Meta-analysis of oxytocin receptor genotype effects in humans. *Psychiatric Genetics, 24*(2), 45-51. doi:10.1097/YPG.0b013e3283643684

Bakhtiari, M., Shleizer-Burko, S., Gymrek, M., Bansal, V., & Bafna, V. (2018). Targeted genotyping of variable number tandem repeats with adVNTR. *Genome Research, 28*(11), 1709-1719.

Baron-Cohen, S., Cassidy, S., Auyeung, B., Allison, C., Achoukhi, M., Robertson, S., . . . Lai, M.-C. (2014). Attenuation of Typical Sex Differences in 800 Adults with Autism vs. 3,900 Controls. *PLoS ONE, 9*(7), e102251. doi:10.1371/journal.pone.0102251

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high- functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241-251.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17. doi:10.1023/A:1005653411471

Bartz, J. A., & Hollander, E. (2008). Oxytocin and experimental therapeutics in autism spectrum disorders. *Progress in Brain Research, 170*, 451-462. doi:10.1016/S0079-6123(08)00435-4

Bisceglia, R., Jenkins, J., Barr, C. L., Wigg, K. G., & Schmidt, L. A. (2014). Arginine Vasopressin Gene Variation and Behavioural Inhibition in Children: An Exploratory Study. *Infant and Child Development, 23*(3), 249-258. doi:10.1002/icd.1866

Bisceglia, R., Jenkins, J. M., Wigg, K. G., O'Connor, T. G., Moran, G., & Barr, C. L. (2012). Arginine vasopressin 1a receptor gene and maternal behavior: evidence of association and moderation. *Genes, Brain and Behavior, 11*(3), 262-268. doi:10.1111/j.1601-183X.2012.00769.x

Blair, R. J. (1999). Psychophysiological responsiveness to the distress of others in children with autism. *Personality and Individual Differences, 26*(3), 477-485.

Blair, R. J., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology, 34*(2), 192-198.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., . . . Fox, J. (2011). OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika, 76*(2), 306-317. doi:10.1007/s11336-010-9200-6

Bons, D., van den Broek, E., Scheepers, F., Herpers, P., Rommelse, N., & Buitelaaar, J. K. (2013). Motor, emotional, and cognitive empathy in children and adolescents with autism spectrum disorder and conduct disorder. *Journal of abnormal child psychology, 41*(3), 425-443.

Brunnlieb, C., Nave, G., Camerer, C. F., Schosser, S., Vogt, B., Münte, T. F., & Heldmann, M. (2016). Vasopressin increases human risky cooperative behavior. *Proceedings of the National Academy of Sciences, 113*(8), 2051-2056. doi:10.1073/pnas.1518825113

Caldwell, H. K., Lee, H.-J., Macbeth, A. H., & Young Iii, W. S. (2008). Vasopressin: Behavioral roles of an "original" neuropeptide. *Progress in Neurobiology, 84*(1), 1-24. doi:http://dx.doi.org/10.1016/j.pneurobio.2007.10.007

Caltabiano, M. L., Sarafino, E. P., & Byrne, D. (2008). *Health Psychology: Biopyschosocial Interactions* (2nd Australasion ed.). Milton QLD: John Wiley & Sons Australia, Ltd.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology, 67*, 319-333.

Chen, X., Hackett, P. D., DeMarco, A. C., Feng, C., Stair, S., Haroon, E., . . . Rilling, J. K. (2016). Effects of oxytocin and vasopressin on the neural response to unreciprocated cooperation within brain regions involved in stress and anxiety in men and women. *Brain imaging and behavior, 10*(2), 581-593. doi:10.1007/s11682-015-9411-7

Chow, C. M., Ruhl, H., & Buhrmester, D. (2013). The mediating role of interpersonal competence between adolescents' empathy and friendship quality: A dyadic approach. *Journal of Adolescence, 36*(1), 191-200. doi:https://doi.org/10.1016/j.adolescence.2012.10.004

Cramer, D., & Jowett, S. (2010). Perceived empathy, accurate empathy and relationship satisfaction in heterosexual couples. *Journal of Social and Personal Relationships, 27*(3), 327-349.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalogue of Selected Documents in Psychology*(10), 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113.

Davis, M. H., & Oathout, H. A. (1987). Maintenance of satisfaction in romantic relationships: Empathy and relational competence. *Journal of Personality and Social Psychology, 53*(2), 397.

de Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences, 10*(10), 435-441. doi:10.1016/j.tics.2006.08.008

de Waal, F. B. M. (2008). Putting the Altruism Back into Altruism: The Evolution of Empathy. *Annual Review of Psychology, 59*(1), 279-300. doi:10.1146/annurev.psych.59.103006.093625

Decety, J., & Jackson, P. L. (2004). The Functional Architecture of Human Empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71-100. doi:10.1177/1534582304267187

Dehning, S., Girma, E., Gasperi, S., Meyer, S., Tesfaye, M., & Siebeck, M. (2012). Comparative cross-sectional study of empathy among first year and final year medical students in Jimma University, Ethiopia: Steady state of the heart and opening of the eyes. *BMC medical education, 12*, 34. doi:10.1186/1472-6920-12-34

Ditzen, B., & Heinrichs, M. (2014). Psychobiology of social support: The social dimension of stress buffering. *Restor Neurol Neurosci, 32*(1), 149-162. doi:10.3233/rnn-139008

Ditzen, B., Neumann, I. D., Bodenmann, G., von Dawans, B., Turner, R. A., Ehlert, U., & Heinrichs, M. (2007). Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women. *Psychoneuroendocrinology, 32*(5), 565-574. doi:https://doi.org/10.1016/j.psyneuen.2007.03.011

Donaldson, Z. R., & Young, L. J. (2008). Oxytocin, Vasopressin, and the Neurogenetics of Sociality. *Science, 322*(5903), 900-904. doi:10.1126/science.1158668

Donaldson, Z. R., & Young, L. J. (2013). The relative contribution of proximal 5' flanking sequence and microsatellite variation on brain vasopressin 1a receptor (Avpr1a) gene expression and behavior. *PLoS Genet, 9*(8), e1003729. doi:10.1371/journal.pgen.1003729

Dumais, K. M., & Veenema, A. H. (2016). Vasopressin and oxytocin receptor systems in the brain: Sex differences and sex-specific regulation of social behavior. *Frontiers in Neuroendocrinology, 40*, 1-23. doi:http://dx.doi.org/10.1016/j.yfrne.2015.04.003

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of Cognitive and Emotional Empathy in Adults with Asperger Syndrome Using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders, 38*(3), 464-473. doi:10.1007/s10803-007-0486-x

Ebstein, R. P., Knafo, A., Mankuta, D., Chew, S. H., & Lai, P. S. (2012). The contributions of oxytocin and vasopressin pathway genes to human behavior. *Hormones and Behavior, 61*(3), 359-379. doi:10.1016/j.yhbeh.2011.12.014

Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion, 14*(2), 131-149.

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94*(1), 100-131. doi:10.1037/0033-2909.94.1.100

Eisenberg, N., Schaller, M., Fabes, R. A., Bustamante, D., Mathy, R. M., Shell, R., & Rhodes, K. (1988). Differentiation of Personal Distress and Sympathy in Children and Adults. *Developmental Psychology, 24*(6), 766-775.

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

Elfenbein, H. A., & Ambady, N. (2003). Universals and cultural differences in recognizing emotions. *Current Directions in Psychological Science, 12*(5), 159-164.

Engels, W. R. (2009). Exact tests for Hardy-Weinberg proportions. *Genetics, 183*(4), 1431-1441. doi:10.1534/genetics.109.108977

Evans, D. M., Frazer, I. H., & Martin, N. G. (1999). Genetic and environmental causes of variation in basal levels of blood cells. *Twin Research, 2*(4), 250-257.

Evans, D. M., Gillespie, N. A., & Martin, N. G. (2002). Biometrical genetics. *Biological Psychology, 61*(1-2), 33-51.

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews, 35*(3), 903-911. doi:10.1016/j.neubiorev.2010.10.009

Feldman, R., Vengrober, A., & Ebstein, R. P. (2014). Affiliation buffers stress: Cumulative genetic risk in oxytocin–vasopressin genes combines with early caregiving to predict PTSD in war-exposed young children. *Translational Psychiatry, 4*(3), e370. doi:10.1038/tp.2014.6

Feng, C., DeMarco, A. C., Haroon, E., & Rilling, J. K. (2015). Neuroticism modulates the effects of intranasal vasopressin treatment on the neural response to positive and negative social interactions. *Neuropsychologia, 73*(0), 108-115. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2015.05.004

Feng, C., Hackett, P. D., DeMarco, A. C., Chen, X., Stair, S., Haroon, E., . . . Rilling, J. K. (2015). Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans. *Brain imaging and behavior, 9*(4), 754-764. doi:10.1007/s11682-014-9333-9

Fernández-Abascal, E. G., Cabello, R., Fernández-Berrocal, P., & Baron-Cohen, S. (2013). Test-retest reliability of the 'Reading the Mind in the Eyes' test: A one-year follow-up study. *Molecular Autism, 4*(1), 33. doi:10.1186/2040-2392-4-33

Finger, E. C. (2011). New Potential Therapeutic Approaches in Frontotemporal Dementia: Oxytocin, Vasopressin, and Social Cognition. *Journal of Molecular Neuroscience, 45*(3), 696-701. doi:10.1007/s12031-011-9550-2

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior, 6*(3), 347-369.

Frank, E., & Landgraf, R. (2008). The vasopressin system — From antidiuresis to psychopathology. *European Journal of Pharmacology, 583*(2–3), 226-242. doi:http://dx.doi.org/10.1016/j.ejphar.2007.11.063

Freeman, S. M. (2016). Comparative Perspectives on Oxytocin and Vasopressin Receptor Research in Rodents and Primates: Translational Implications. *Journal of neuroendocrinology, 28*(4). doi:10.1111/jne.12382

Golimbet, V., Alfimova, M., Abramova, L., Kaleda, V., & Gritsenko, I. (2015). Arginine vasopressin 1a receptor RS3 promoter microsatellites in schizophrenia: A study of the effect of the "risk" allele on clinical symptoms and facial affect recognition. *Psychiatry Research, 225*(3), 739-740. doi:http://dx.doi.org/10.1016/j.psychres.2014.11.043

Gozzi, M., Dashow, E. M., Thurm, A., Swedo, S. E., & Zink, C. F. (2017). Effects of Oxytocin and Vasopressin on Preferential Brain Responses to Negative Social Feedback. *Neuropsychopharmacology, 42*(7), 1409-1419. doi:10.1038/npp.2016.248

Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., & Doebley, J. (2012). *Introduction to Genetic Analysis* (10 ed.). New York: W. H. Freeman and Company.

Guastella, A. J., Kenyon, A. R., Alvares, G. A., Carson, D. S., & Hickie, I. B. (2010). Intranasal Arginine Vasopressin Enhances the Encoding of Happy and Angry Faces in Humans. *BPS, 67*(12), 1220-1222. doi:10.1016/j.biopsych.2010.03.014

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

Hammock, E. A. D., & Young, L. J. (2002). Variation in the vasopressin V1a receptor promoter and expression: Implications for inter- and intraspecific variation in social behaviour. *European Journal of Neuroscience, 16*, 399-402.

Hammock, E. A. D., & Young, L. J. (2005). Microsatellite Instability Generates Diversity in Brain and Sociobehavioral Traits. *Science, 308*(5728), 1630-1634. doi:10.1126/science.1111427

Hannan, A. J. (2009). Tandem repeat polymorphisms: Modulators of disease suscptibility and candidates for "missing heritability". *Trends in Genetics, 26*(2), 59-65.

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

Harkness, K. L., Jacobson, J. A., Duong, D., & Sabbagh, M. A. (2010). Mental state decoding in past major depression: Effect of sad versus happy mood induction. *Cognition and Emotion, 24*(3), 497-513. doi:10.1080/02699930902750249

Harkness, K. L., Sabbagh, M., Jacobson, J., Chowdrey, N., & Chen, T. (2005). Enhanced accuracy of mental state decoding in dysphoric college students. *Cognition and Emotion, 19*(7), 999-1025. doi:10.1080/02699930541000110

Hatemi, P. K., Smith, K., Alford, J. R., Martin, N. G., & Hibbing, J. R. (2015). The Genetic and Environmental Foundations of Political, Psychological, Social, and Economic Behaviors: A Panel Study of Twins and Families. *Twin Research and Human Genetics, 18*(3), 243-255. doi:10.1017/thg.2015.13

Heinrichs, M., Baumgartner, T., Kirschbaum, C., & Ehlert, U. (2003). Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biological Psychiatry, 54*(12), 1389-1398. doi:https://doi.org/10.1016/S0006-3223(03)00465-7

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social Relationships and Mortality Risk: A Meta-analytic Review. *PLOS Medicine, 7*(7), e1000316. doi:10.1371/journal.pmed.1000316

Hopkins, W. D., Donaldson, Z. R., & Young, L. J. (2012). A polymorphic indel containing the RS3 microsatellite in the 5′ flanking region of the vasopressin V1a receptor gene is associated with chimpanzee (Pan troglodytes) personality. *Genes, Brain and Behavior, 11*(5), 552-558. doi:10.1111/j.1601-183X.2012.00799.x

Hunsley, J., Best, M., Lefebvre, M., & Vito, D. (2001). The Seven-Item Short Form of the Dyadic Adjustment Scale: Further Evidence for Construct Validity. *The American Journal of Family Therapy, 29*(4), 325-335. doi:10.1080/01926180126501

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., . . . Kendrick, K. M. (2010). Oxytocin Enhances Amygdala-Dependent, Socially Reinforced Learning and Emotional Empathy in Humans. *Journal of Neuroscience, 30*(14), 4999-5007. doi:10.1523/JNEUROSCI.5538-09.2010

Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships, 7*(1), 95-109. doi:10.1111/j.1475-6811.2000.tb00006.x

Inoue, T., Kimura, T., Azuma, C., Inazawa, J., Takemura, M., Kikuchi, T., . . . Saji, F. (1994). Structural organization of the human oxytocin receptor gene. *Journal of Biological Chemistry, 269*(51), 32451-32456.

Insel, T. R. (2010). The Challenge of Translation in Social Neuroscience: A Review of Oxytocin, Vasopressin, and Affiliative Behavior. *Neuron, 65*(6), 768-779. doi:10.1016/j.neuron.2010.03.005

Insel, T. R., & Shapiro, L. E. (1992). Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proceedings of the National Academy of Sciences of the United States of America, 89*, 5981-5985.

Insel, T. R., Wang, Z. X., & Ferris, C. F. (1994). Patterns of brain vasopressin receptor distribution associated with social organization in microtine rodents. *The Journal of Neuroscience, 14*(9), 5381-5392.

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Reibold, M., Bachnermelman, R., . . . Yirmiya, N. (2008). Molecular genetic studies of the arginine vasopressin 1a receptor (AVPR1a) and the oxytocin receptor (OXTR) in human behaviour: from autism to altruism with some notes in between. *Progress in Brain Research, 170*, 435-449. doi:10.1016/S0079-6123(08)00434-2

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Riebold, M., Laiba, E., . . . Ebstein, R. P. (2009). The Oxytocin Receptor (OXTR) Contributes to Prosocial Fund Allocations in the Dictator Game and the Social Value Orientations Task. *PLoS ONE, 4*(5), e5535. doi:10.1371/journal.pone.0005535.t003

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*(4), 589-611. doi:10.1016/j.adolescence.2005.08.010

Kantojärvi, K., Oikkonen, J., Kotala, I., Kallela, J., Vanhala, R., Onkamo, P., & Järvelä, I. (2015). Association and Promoter Analysis of AVPR1A in Finnish Autism Families. *Autism research, 8*(5), 634-639. doi:10.1002/aur.1473

Khorashad, B. S., Baron-Cohen, S., Roshan, G. M., Kazemian, M., Khazai, L., Aghili, Z., . . . Afkhamizadeh, M. (2015). The "Reading the Mind in the Eyes" Test: Investigation of Psychometric Properties and Test-Retest Reliability of the Persian Version. *J Autism Dev Disord, 45*(9), 2651-2666. doi:10.1007/s10803-015-2427-4

Kiely, K. M., & Butterworth, P. (2015). Validation of four measures of mental health against depression and generalized anxiety in a community based sample. *Psychiatry Research, 225*(3), 291-298. doi:http://dx.doi.org/10.1016/j.psychres.2014.12.023

Kim, S.-J., Young, L. J., Gonen, D., Veenstra-VanderWeele, J., Courchesne, R., Courchesne, E., . . . Insel, T. R. (2002). Transmission disequilibrium testing of arginine vasopressin receptor 1A (AVPR1A) polymorphisms in autism. *Molecular Psychiatry, 7*(5), 503-507. doi:10.1038/sj.mp.4001125

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81.

Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L., . . . Ebstein, R. P. (2008). Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. *Genes, Brain and Behavior, 7*(3), 266-275. doi:10.1111/j.1601-183X.2007.00341.x

Knafo, A., Zahn-Waxler, C., Davidov, M., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2009). Empathy in Early Childhood. *Annals of the New York Academy of Sciences, 1167*(1), 103-114. doi:10.1111/j.1749-6632.2009.04540.x

Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: Genetic and environmental contributions. *Emotion, 8*(6), 737-752. doi:10.1037/a0014179

Konrath, S., Corneille, O., Bushman, B. J., & Luminet, O. (2013). The Relationship Between Narcissistic Exploitativeness, Dispositional Empathy, and Emotion Recognition Abilities. *Journal of Nonverbal Behavior, 38*(1), 129-143. doi:10.1007/s10919-013-0164-y

Kormos, V., & Gaszner, B. (2013). Role of neuropeptides in anxiety, stress and depression: From animals to humans. *Neuropeptides, 47*(6), 401-419.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychological Medicine, 34*(5), 911-924. doi:10.1017/S0033291703001624

Lee, H. J., & Szinovacz, M. E. (2016). Positive, Negative, and Ambivalent Interactions With Family and Friends: Associations With Well-being. *Journal of Marriage and Family, 78*(3), 660-679. doi:10.1111/jomf.12302

Leerkes, E. M., Su, J., Calkins, S., Henrich, V. C., & Smolen, A. (2017). Variation in mothers' arginine vasopressin receptor 1a and dopamine receptor D4 genes predicts maternal sensitivity via social cognition AVPR1a, DRD4 and sensitivity. *Genes, Brain and Behavior, 16*(2), 233-240. doi:10.1111/gbb.12326

Levin, R., Heresco-Levy, U., Bachner-Melman, R., Israel, S., Shalev, I., & Ebstein, R. P. (2009). Association between arginine vasopressin 1a receptor (AVPR1a) promoter region polymorphisms and prepulse inhibition. *Psychoneuroendocrinology, 34*(6), 901-908. doi:10.1016/j.psyneuen.2008.12.014

Loup, F., Tribollet, E., Dubois-Dauphin, M., & Dreifuss, J. J. (1991). Localization of high-affinity binding sites for oxytocin and vasopressin in the human brain. An autoradiographic study. *Brain Research, 555*(2), 220-232.

Lyu, J., & Agrigoroaei, S. (2017). Childhood Misfortune and Adult Health in a National Study. *The International Journal of Aging and Human Development, 84*(3), 213-230. doi:10.1177/0091415016669147

Maher, B. S., Vladimirov, V. I., Latendresse, S. J., Thiselton, D. L., McNamee, R., Kang, M., . . . Vanyukov, M. M. (2011). The AVPR1A Gene and Substance Use Disorders: Association, Replication, and Functional Evidence. *Biological Psychiatry, 70*(6), 519-527. doi:http://dx.doi.org/10.1016/j.biopsych.2011.02.023

Mahovetz, L. M., Young, L. J., & Hopkins, W. D. (2016). The influence of *AVPR1A* genotype on individual differences in behaviors during a mirror self-recognition task in chimpanzees (*Pan troglodytes*). *Genes, Brain and Behavior, 15*(5), 445-452. doi:10.1111/gbb.12291

Maibom, H. L. (2012). The many faces of empathy and their relation to prosocial action and aggression inhibition. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(2), 253-263. doi:10.1002/wcs.1165

McCall, C., & Singer, T. (2012). The animal and human neuroendocrinology of social cognition, motivation and behavior. *Nature Neuroscience, 15*(5), 681-688. doi:10.1038/nn.3084

McGregor, I. S., & Bowen, M. T. (2012). Breaking the loop: Oxytocin as a potential treatment for drug addiction. *Hormones and Behavior, 61*(3), 331-339. doi:10.1016/j.yhbeh.2011.12.001

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*(4), 525–543.

Melchers, M., Montag, C., Markett, S., & Reuter, M. (2015). Assessment of empathy via self-report and behavioural paradigms: data on convergent and discriminant validity. *Cogn Neuropsychiatry, 20*(2), 157-171. doi:10.1080/13546805.2014.991781

Members of the Complex Trait Consortium. (2003). The nature and identification of quantitative trait loci: A community's view. *Nature Reviews Genetics, 4*(11), 911-916.

Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature Reviews Neuroscience, 12*, 1-15. doi:10.1038/nrn3044

Meyer-Lindenberg, A., Kolachana, B., Gold, B., Olsh, A., Nicodemus, K. K., Mattay, V., . . . Weinberger, D. R. (2009). Genetic variants in AVPR1A linked to autism predict amygdala activation and personality traits in healthy humans. *Mol Psychiatry, 14*(10), 968-975.

Miller, G. E., Rohleder, N., & Cole, S. W. (2009). Chronic Interpersonal Stress Predicts Activation of Pro- and Anti-Inflammatory Signaling Pathways 6 Months Later. *Psychosomatic Medicine, 71*(1), 57-62.

Modi, M. E., & Young, L. J. (2012). The oxytocin system in drug discovery for autism: Animal models and novel therapeutic strategies. *Hormones and Behavior, 61*(3), 340-350. doi:10.1016/j.yhbeh.2011.12.010

Moons, W. G., Way, B. M., & Taylor, S. E. (2014). Oxytocin and vasopressin receptor polymorphisms interact with circulating neuropeptides to predict human emotional reactions to stress. *Emotion, 14*(3), 562-572. doi:10.1037/a0035503

Muncer, S. J., & Ling, J. (2006). Psychometric analysis of the empathy quotient (EQ) scale. *Personality and Individual Differences, 40*(6), 1111-1119. doi:10.1016/j.paid.2005.09.020

Neale, M. C., de Knijff, P., Havekes, L. M., & Boomsma, D. I. (2000). ApoE Polymorphism Accounts for Only Part of the Genetic Variation in Quantitative ApoE Levels. *Genetic Epidemiology, 18*, 331-340.

Neumann, I. D., & Landgraf, R. (2012). Balance of brain oxytocin and vasopressin: implications for anxiety, depression, and social behaviors. *Trends in Neurosciences, 35*(11), 649-659. doi:10.1016/j.tins.2012.08.004

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Front Psychol, 6*, 1503. doi:10.3389/fpsyg.2015.01503

Palermo, R., O'Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PLoS ONE, 8*(6), e68126. doi:10.1371/journal.pone.0068126

Pare, P., Paixão-Côrtes, V. R., Tovo-Rodrigues, L., Vargas-Pinilla, P., Viscardi, L. H., Salzano, F. M., . . . Bortolini, M. C. (2016). Oxytocin and arginine vasopressin receptor evolution: implications for adaptive novelties in placental mammals. *Genetics and molecular biology, 39*(4), 646-657. doi:10.1590/1678-4685-gmb-2015-0323

Parens, E., Chapman, A. R., & Press, N. (Eds.). (2006). *Wrestling with Behavioural Genetics: Science, Ethics, and Public Conversation*. Baltimore, Maryland: The John Hopkins University Press.

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

Prichard, Z. M., & Easteal, S. (2006). Characterization of simple sequence repeat variants linked to candidate genes for behavioral phenotypes. *Human Mutation, 27*(1), 120. doi:10.1002/humu.9394

Prichard, Z. M., Mackinnon, A. J., Jorm, A. F., & Easteal, S. (2007). AVPR1A and OXTR polymorphisms are associated with sexual and reproductive behavioral phenotypes in humans. *Human Mutation, 28*(11), 1150-1150. doi:10.1002/humu.9510

Procyshyn, T. L., Hurd, P. L., & Crespi, B. J. (2017). Association testing of vasopressin receptor 1a microsatellite polymorphisms in non-clinical autism spectrum phenotypes Avpr1a polymorphisms and non-clinical autism. *Autism research, 10*(5), 750-756. doi:10.1002/aur.1716

Ragsdale, G., & Foley, R. A. (2011). A maternal influence on reading the mind in the eyes mediated by executive function: Differential parental influences on full and half-siblings. *PLoS ONE, 6*(8), e23236. doi:10.1371/journal.pone.0023236.t006

Reuter, M., Cooper, A. J., Smillie, L. D., Markett, S., & Montag, C. (2015). A new measurement for the revised reinforcement sensitivity theory: psychometric criteria and genetic validation. *Frontiers in Systems Neuroscience, 9*. doi:10.3389/fnsys.2015.00038

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Chen, X., Gautam, P., Stair, S., . . . Pagnoni, G. (2014). Sex differences in the neural and behavioral response to intranasal oxytocin

and vasopressin during human social interaction. *Psychoneuroendocrinology, 39*, 237-248. doi:10.1016/j.psyneuen.2013.09.022

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Thompson, R., Ditzen, B., Patel, R., & Pagnoni, G. (2012). Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology, 37*(4), 447-461. doi:10.1016/j.psyneuen.2011.07.013

Santini, Z. I., Fiori, K. L., Feeney, J., Tyrovolas, S., Haro, J. M., & Koyanagi, A. (2016). Social relationships, loneliness, and mental health among older men and women in Ireland: A prospective community-based study. *Journal of Affective Disorders, 204*, 59-69. doi:http://doi.org/10.1016/j.jad.2016.06.032

Schuster, T. L., Kessler, R. C., & Aseltine, R. H. J. (1990). Supportive Interactions, Negative Interactions, and Depressed Mood *American Journal of Community Psychology, 18*(3), 423-437.

Seeman, T. E., Gruenewald, T. L., Cohen, S., Williams, D. R., & Matthews, K. A. (2014). Social relationships and their biological correlates: Coronary Artery Risk Development in Young Adults (CARDIA) study. *Psychoneuroendocrinology, 43*, 126-138.

Shalev, I., Israel, S., Uzefovsky, F., Gritsenko, I., Kaitz, M., & Ebstein, R. P. (2011). Vasopressin needs an audience: Neuropeptide elicited stress responses are contingent upon perceived social evaluative threats. *Hormones and Behavior, 60*(1), 121-127. doi:10.1016/j.yhbeh.2011.04.005

Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis: Modelling change and event occurence*. New York: Oxford University Press.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews, 30*(6), 855-863. doi:10.1016/j.neubiorev.2006.06.011

Smith, R. L., & Rose, A. J. (2011). The "Cost of Caring" in Youths' Friendships: Considering Associations Among Social Perspective Taking, Co-Rumination, and Empathetic Distress. *Developmental Psychology, 47*(6), 1792-1803.

Söderstrand, P., & Almkvist, O. (2012). Psychometric data on the Eyes Test, the Faux Pas Test, and the Dewey Social Stories Test in a population-based Swedish adult sample. *Nordic Psychology, 64*(1), 30-43. doi:10.1080/19012276.2012.693729

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38*, 15–28.

Spek, A. A., Scholte, E. M., & Berckelaer-Onnes, I. A. (2010). Theory of Mind in adults with HFA and Asperger Syndrome. *Journal of Autism and Developmental Disorders, 40*(3), 280-289. doi:10.1007/s10803-009-0860-y

Staes, N., Koski, S. E., Helsen, P., Fransen, E., Eens, M., & Stevens, J. M. G. (2015). Chimpanzee sociability is associated with vasopressin (Avpr1a) but not oxytocin receptor gene (OXTR) variation. *Hormones and Behavior, 75*, 84-90. doi:http://dx.doi.org/10.1016/j.yhbeh.2015.08.006

Staes, N., Weiss, A., Helsen, P., Korody, M., Eens, M., & Stevens, J. M. G. (2016). Bonobo personality traits are heritable and associated with vasopressin receptor gene 1a variation. *Scientific Reports, 6*, 38193. doi:10.1038/srep38193

Striepens, N., Kendrick, K. M., Maier, W., & Hurlemann, R. (2011). Prosocial effects of oxytocin and clinical evidence for its therapeutic potential. *Frontiers in Neuroendocrinology, 32*(4), 426-450. doi:10.1016/j.yfrne.2011.07.001

Tabor, H., Risch, N. J., & Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: Practical considerations. *Nature Reviews Genetics, 3*, 1-7.

Tansey, K. E., Hill, M. J., Cochrane, L. E., Gill, M., Anney, R. J. L., & Gallagher, L. (2011). Functionality of promoter microsatellites of arginine vasopressin receptor 1A (AVPR1A): implications for autism. *Molecular Autism, 2*(3).

Teo, A. R., Choi, H., & Valenstein, M. (2013). Social Relationships and Depression: Ten-Year Follow-Up from a Nationally Representative Study. *PLoS ONE, 8*(4), e62396. doi:10.1371/journal.pone.0062396

Thibonnier, M., Graves, M. K., Wagner, M. S., Auzan, C., Clauser, E., & Willard, H. F. (1996). Structure, sequence, expression, and chromosomal localization of the human V1a vasopressin receptor gene. *Genomics, 31*(3), 327-334. doi:10.1006/geno.1996.0055

Thibonnier, M., Graves, M. K., Wagner, M. S., Chatelain, N., Soubrier, F., Corvol, P., . . . Jeunemaitre, X. (2000). Study of V1-vascular Vasopressin Receptor Gene Microsatellite Polymorphisms in Human Essential Hypertension. *Journal of Molecular and Cellular Cardiology, 32*(4), 557-564. doi:10.1006/jmcc.2000.1108

Thompson, R. R., George, K., Walton, J. C., Orr, S. P., & Benson, J. (2006). Sex-specific influences of vasopressin on human social communication. *Proceedings of the National Academy of Sciences of the United States of America, 103*(20), 7889-7894. doi:10.1073/pnas.0600406103

Tone, E. B., & Tully, E. C. (2014). Empathy as a "risky strength": A multilevel examination of empathy and risk for internalizing disorders. *Development and Psychopathology, 26*(4), 1547-1565. doi:10.1017/S0954579414001199

Twisk, J. W. R. (2003). *Applied longitudinal data analysis for epidemiology: A practical guide*. Cambridge: Cambridege University Press.

Uzefovsky, F., Shalev, I., Israel, S., Edelman, S., Raz, Y., Mankuta, D., . . . Ebstein, R. P. (2015). Oxytocin receptor and vasopressin receptor 1a genes are respectively associated with emotional and cognitive empathy. *Hormones and Behavior, 67*, 60-65. doi:http://dx.doi.org/10.1016/j.yhbeh.2014.11.007

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

Vogel, F., Wagner, S., Baskaya, Ö., Leuenberger, B., Mobascher, A., Dahmen, N., . . . Tadic, A. (2012). Variable number of tandem repeat polymorphisms of the arginine vasopressin receptor 1A gene and impulsive aggression in patients with borderline personality disorder. *Psychiatric Genetics, 22*(2), 105-106. doi:10.1097/YPG.0b013e32834accad

Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D:4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Personality and Individual Differences, 41*(8), 1481-1491. doi:10.1016/j.paid.2006.06.009

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., . . . Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences, 41*(5), 929-940. doi:10.1016/j.paid.2006.03.017

Walum, H., Westberg, L., Henningsson, S., Neiderhiser, J. M., Reiss, D., Igl, W., . . . Lichtenstein, P. (2008). Genetic variation in the vasopressin receptor 1a gene (AVPR1A) associates with pair-bonding behavior in humans. *Proceedings of the National Academy of Sciences, 105*(37), 14153-14156. doi:10.1073/pnas.0803081105

Wang, J., Qin, W., Liu, F., Liu, B., Zhou, Y., Jiang, T., & Yu, C. (2016). Sex-Specific Mediation Effect of the Right Fusiform Face Area Volume on the Association Between Variants in Repeat Length of AVPR1A RS3 and Altruistic Behavior in Healthy Adults. *Human brain mapping, 37*(7), 2700-2709. doi:10.1002/hbm.23203

Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care, 34*(3), 220-233.

Wassink, T. H., Piven, J., Vieland, V. J., Pietila, J., Goedken, R. J., Folstein, S. E., & Sheffield, V. C. (2004). Examination of AVPR1a as an autism susceptibility gene. *Molecular Psychiatry, 9*(10), 968-972. doi:10.1038/sj.mp.4001503

Wong, S. T., Wu, A., Gregorich, S., & Perez-Stable, E. J. (2014). What Type of Social Support Influences Self-Reported Physical and Mental Health Among Older Women? *Journal of Aging and Health, 26*(4), 663-678.

World Health Organization. (1993). *The ICD-10 Classification of Mental and Behavioural Disorders: clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Wright, M. J., & Martin, N. G. (2004). Brisbane Adolescent Twin Study: Outline of study methods and research projects. *Australian Journal of Psychology, 56*(2), 65-78. doi:10.1080/00049530410001734865

Yang, S. Y., Cho, S.-C., Yoo, H. J., Cho, I. H., Park, M., Yoe, J., & Kim, S. A. (2010). Family-based association study of microsatellites in the 5' flanking region of *AVPR1A* with autism spectrum disorder in the Korean population. *Psychiatry Research, 178*(1), 199-201. doi:10.1016/j.psychres.2009.11.007

Yang, Y. C., Schorpp, K., & Harris, K. M. (2014). Social support, social strain and inflammation: Evidence from a national longitudinal study of U.S. adults. *Social Science & Medicine, 107*, 124-135.

Yıldırım, E. A., Kaşar, M., Güdük, M., Ateş, E., Küçükparlak, I., & Ozalmete, E. O. (2011). Investigation of the reliability of the "Reading the Mind in the Eyes test" in a Turkish population. *Turkish Journal of Psychiatry, 22*(3), 177-186.

Yirmiya, N., Rosenberg, C., Levi, S., Salomon, S., Shulman, C., Nemanov, L., . . . Ebstein, R. P. (2006). Association between the arginine vasopressin 1a receptor (AVPR1a) gene and autism in a family-based study: mediation by socialization skills. *Molecular Psychiatry, 11*(5), 488-494. doi:10.1038/sj.mp.4001812

Young, L. J., Nilsen, R., Waymire, K. G., MacGregor, G. R., & Insel, T. R. (1999). Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. *Nature, 400*(6746), 766-768.

Zhu, M., & Zhao, S. (2007). Candidate gene identification approach: Progress and challenges. *International Journal of Biological Sciences, 3*(7), 420-427.

Zink, C. F., & Meyer-Lindenberg, A. (2012). Human neuroimaging of oxytocin and vasopressin in social cognition. *Hormones and Behavior, 61*(3), 400-409. doi:http://dx.doi.org/10.1016/j.yhbeh.2012.01.016

Zink, C. F., Stein, J. L., Kempf, L., Hakimi, S., & Meyer-Lindenberg, A. (2010). Vasopressin Modulates Medial Prefrontal Cortex–Amygdala Circuitry during Emotion Processing

in Humans. *The Journal of Neuroscience, 30*(20), 7017-7022.

doi:10.1523/jneurosci.4899-09.2010