Agreeableness is Related to Social-cognitive, but Not Social-perceptual, Theory of Mind

DANIEL NETTLE* and BETHANY LIDDLE

Centre for Behaviour and Evolution, Henry Wellcome Building, Newcastle University, Newcastle, UK

Abstract

We hypothesise on a number of grounds that the personality dimension of Agreeableness may be associated with inter-individual differences in theory of mind (ToM) functioning. However, it is important to distinguish social-perceptual from social-cognitive ToM. Previous findings on ToM in psychopathic individuals, sex differences in ToM and the associations between ToM and social relationships, all suggest that social-cognitive ToM is more likely than social-perceptual ToM to relate to Agreeableness. In separate empirical studies, we find that Agreeableness is substantially correlated with social-cognitive ToM performance, but uncorrelated with social-perceptual ToM performance. We suggest that the propensity or motivation to attend to the mental states of others may be central to the personality dimension of Agreeableness. Copyright © 2008 John Wiley & Sons, Ltd.

Key words: theory of mind; empathy; empathising; Agreeableness; five-factor model

INTRODUCTION

The five-factor model of personality (Costa & McCrae, 1992; Goldberg, 1990) posits Agreeableness as a major dimension of inter-individual dispositional variation. High Agreeableness is associated with warmth, friendliness, altruism and compliance to the needs of others (Digman & Takemoto-Chock, 1981; Graziano & Eisenberg, 1997), is uniquely predictive of social support and harmonious relationships (Asendorpf & Wilpers, 1998; Soldz & Vaillant, 1999), and is negatively associated with anger, aggression and interpersonal arguments (Jensen-Campbell & Graziano, 2001; Meier & Robinson, 2004).

Personality researchers are no longer content merely to identify stable personality dimensions. Instead, they have begun to turn their attention to researching the cognitive and

*Correspondence to: Daniel Nettle, Centre for Behaviour and Evolution, Henry Wellcome Building, Newcastle University, Framlington Place, Newcastle NE2 4HH, UK. E-mail: daniel.nettle@ncl.ac.uk

Copyright © 2008 John Wiley & Sons, Ltd.

neural mechanisms which underlie them (see Whittle, Allen, Lubman, & Yücel, 2006 for a recent review). Thus, for example, the dimension of Extraversion has been linked to variation in dopamine-using mid-brain reward systems (Depue & Collins, 1999), whilst Neuroticism has been linked to variation in the activity of serotonin-mediated negative emotion systems, particularly the circuit involving the amygdala (see Whittle et al., 2006). Whittle et al. (2006) also review evidence suggesting a key role for response inhibition mechanisms in the dorsolateral prefrontal cortex in Conscientiousness, whilst DeYoung, Peterson, & Higgins (2005) have suggested that Openness reflects variation in prefrontal higher cognitive mechanisms more generally. Of all the big five personality dimensions, it is Agreeableness for which research into the underlying cognitive or neural mechanisms is least developed.

An attractive hypothesis is that Agreeableness relates to variation in theory of mind (ToM) functioning. ToM is the capacity to infer and reason about the mental states of others. ToM has been most intensively studied developmentally (see Wellman, Cross, & Watson, 2001 for a review), and in relation to autism, where it is markedly impaired (Baron-Cohen, Leslie, & Frith, 1985). There is also some evidence concerning the neural structures involved, with a circuit including right medial temporal and orbitofrontal areas, and left medial frontal areas, being especially implicated (Sabbagh, 2004; Siegal & Varley, 2002). Recent research using several different ToM tests has demonstrated that there is considerable variation in ToM performance amongst 'normal' adults (e.g. Davis, 1983; Paal & Bereczkei, 2007; Stiller & Dunbar, 2007). Could this variation be related to individual differences in Agreeableness?

The grounds for suspecting that this might be the case are several. First, the content of the questionnaire items used to measure Agreeableness largely concerns the mental states of others (e.g. 'I sympathise with others' feelings, or, negatively scored, 'I am not interested in other people's problems'), and the behaviours specifically associated with Agreeableness are those for which consideration of the needs and perspectives of others are key. Second, the sex difference in Agreeableness (at least half a standard deviation, favouring females; Costa, Terraciano, & McCrae, 2001), closely mirrors that found on various measures of ToM (Stiller & Dunbar, 2007; see also Nettle, 2007). Third, Stiller and Dunbar (2007) find that a measure of ToM predicts the size of people's social networks, Paal and Bereczkei (2007) find ToM performance to be closely correlated with social cooperation, and Mohr, Howells, Gerace, Day, & Wharton (2007) find that ToM is associated with angry responses in interpersonal interaction. These are all very similar correlates to those which have been found for Agreeableness (Asendorpf & Wilpers, 1998; Jensen-Campbell & Graziano, 2001; Meier & Robinson, 2004). Fourth, primary psychopathic traits, as measured by the psychopathy checklist, load strongly on big five Agreeableness (Jakobwitz & Egan, 2006). Psychopathy represents callousness and indifference to the suffering of others, and abnormalities of ToM processing appear to be implicated (Dolan & Fullam, 2004).

Finally, the strongest reason for suspecting a relationship between ToM and Agreeableness is that the 'empathy quotient' (Baron-Cohen & Wheelwright, 2004), a self-report measure designed to tap ToM differences in normal adults, correlates with five-factor Agreeableness at above 0.7 (Nettle, 2007). This correlation is sufficiently high to suspect that the two scales are measuring closely related traits. However, the empathy quotient is a self-report scale very like a personality inventory. The evidence would be much more compelling if, instead of the empathy quotient, one of the measures based on objective performance on a cognitive ToM task were used instead.

The aim of this paper is therefore to examine the relationship between Agreeableness, as measured by a standard self-report personality questionnaire, and ToM as measured by objective cognitive tasks. Our general prediction is that there will be a relationship between ToM functioning and Agreeableness. However, ToM appears to consist of several dissociable components. In particular, there are grounds for distinguishing between a *social-perceptual* component and a *social-cognitive* component (Sabbagh, 2004; Tager-Flusberg & Sullivan, 2000). The social-perceptual component of ToM is the ability to detect the mental states of others using immediately available cues such as facial expressions and bodily movements. The classic task assessing this capacity is the 'Reading the Mind in the Eyes test' (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), in which the participant has to judge the best description of an actor's mental state from a photograph of the eye region. The social-cognitive component of ToM is the ability to reason about the content of another's mental state, and use such reasoning to predict or explain their actions. Tasks tapping the social-cognitive component of ToM involve hearing stories or scenarios, and making correct inferences about what the individuals involved know or believe. The social-cognitive component of ToM is later developing than the social-perceptual component, and more closely related to the linguistic capacities.

The two components of ToM seem to involve distinct neural structures, as evidenced by ERP, imaging and lesion evidence (Sabbagh, 2004). The social-perceptual component is especially related to right hemisphere medial temporal and orbitofrontal areas, whilst the social-cognitive component relies on a separate left hemisphere circuit involving medial frontal areas (Siegal & Varley, 2002; Sabbagh, 2004), and the temporoparietal junction (Samson, Apperly, Chiavarino, & Humphreys, 2004). The two components of ToM can also be dissociated neuropsychologically. Williams syndrome is a developmental disorder involving general retardation, but which leaves the social-perceptual, but not the social-cognitive component, of ToM relatively spared. This is detectable by good performance by Williams syndrome children on a reading the mind in the eyes task but not in story- or scenario-based tests of reasoning about mental states (Tager-Flusberg & Sullivan, 2000).

In terms of a prediction about which component of ToM will be most relevant to Agreeableness, suggestive evidence comes psychopathy, and from sex differences. As mentioned above, primary psychopathy is strongly related to Agreeableness. Studies of psychopathic individuals have found no general impairment in the social-perceptual component of ToM, as measured by the Reading the Mind in the Eyes test (Dolan & Fullam, 2004; Richell, Mitchell, Newman, Leonard, Baron-Cohen, & Blair, 2003). However, when social-cognitive ToM tasks of sufficient difficulty to avoid ceiling effects are administered, psychopathic individuals are clearly impaired relative to controls (Dolan & Fullam, 2004).

As for sex differences, it is social-cognitive ToM tasks that are strongly sexually dimorphic (Stiller & Dunbar, 2007), whereas in the social-perceptual eyes task, the sex difference is absent or very small (Baron-Cohen et al., 2001). Since Agreeableness is clearly sexually dimorphic (Costa et al., 2001), it appears to pattern more with social-cognitive than social-perceptual ToM.

Our prediction is thus that any relationship between Agreeableness and ToM will be less likely to appear in a social-perceptual ToM task, and more likely to appear in a social-cognitive task, provided that the task is of sufficient difficulty to produce a range of variation in high-functioning adults. The two studies that follow use, respectively, a social-perceptual and a difficult social-cognitive ToM task, in order to test this prediction.

STUDY 1

Introduction

In Study 1, we relate self-reported Agreeableness to the standard social-perceptual ToM task, the Reading the Mind in the Eyes test (revised; Baron-Cohen et al., 2001). Though a correlation with Agreeableness is possible, prior findings concerning sex differences and psychopathy (see above) tend to argue against finding one.

Methods

Participants
Participants were an opportunity sample of 96 individuals (48 males and 48 females), the majority of whom were undergraduate students at Newcastle University, England.

Five-factor personality measure
The 50-item instrument of the International Personality Item Pool (Goldberg, 1999) was used to measure the big five personality factors (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism). Participants were asked to rate on a five-point Likert scale how accurate each statement was of themselves.

Theory of mind task
Participants were tested on the revised version of the adult Eyes Test (Baron-Cohen et al., 2001). They are presented with a booklet of 36 photographs of the eye region of the face (18 photographs of different male actors and 18 photographs of different female actors). For each eye stimulus, participants are asked to choose which word best describes what the person in the photograph is thinking or feeling, from four forced-choice response options. The three foil words have roughly the same emotional valence as the target word. Participants were provided with a glossary of all the mental state terms which they were told they could consult at any time if they were unsure of a word meaning. The test stimuli contain only complex mental states. The ToM score was simply the number of correct responses given out of 36.

Procedure
Both tasks were administered individually in a quiet room in Newcastle University. On completion of the first task participants moved on to the second task. Task order was randomised.

Results

Personality scores
Cronbach's α for the all the personality scales were high (0.77–0.89). Females had higher scores than males for Agreeableness ($d = 0.60$, $p < 0.01$) and Neuroticism ($d = 0.60$, $p < 0.01$), whilst males had higher scores than females for Openness ($d = 0.45$, $p < 0.05$).

Theory of mind task
Performance on the ToM task showed an overall mean score of 27.32 (SD 3.41; median 28, modal score 27). The distribution of scores was approximately normal with the lowest (17)

well above chance guessing (which would be 9 on this task), and the highest (33) below the perfect performance of 36. Reliability was assess using the Spearman–Brown corrected correlation between scores on items 1–18 and items 19–36, which yielded a reliability coefficient of 0.75. ToM scores were slightly higher in males (mean 27.81, SD 3.59) than females (mean 26.83, SD 3.18). This difference was, however, not significant ($t = -1.41$, $p = 0.161$).

There was no significant correlation between ToM score and Agreeableness, or indeed any of the five personality factors (Table 1).

Discussion

Scores on the Eyes Test were close to those reported by Baron-Cohen et al. (2001) for a general population adult group (their mean 26.2, SD 3.6, our mean 27.32, SD 3.41). Like them, we find no significant sex difference on this task. Thus, we confirm the generalisation that the social-perceptual component of ToM is not sexually dimorphic.

Personality scores were also similar to previous studies using this instrument, and the sex differences (higher female scores on Agreeableness and Neuroticism, higher male scores on Openness) agreed with those previously reported in the literature (Costa et al., 2001; Nettle, 2007).

There were no significant relationships between ToM score and Agreeableness, or indeed any of the five personality factors. This despite the fact that, with 96 participants, our power to detect a correlation of even 0.2 was over 99%. Thus, the results of Study 1 suggest that if there is any relationship between Agreeableness and ToM, it is not with the social-perceptual component of ToM.

STUDY 2

Introduction

The second study used a challenging social-cognitive ToM task instead of the social-perceptual one. The task is derived from Stiller and Dunbar (2007), and was in turn derived from Kinderman, Dunbar, and Bentall (1998). It is based on the participant hearing multi-character stories, and afterwards asking questions, some of which necessitate reasoning about the mental states of the characters. In particular, the questions tap nested ToM judgements (such as the belief of character A about the belief of character B). Previous research suggests that normal adults perform successfully on this task about to

Table 1. Correlations between ToM and personality factors, and amongst the personality factors, Study 1

	ToM	O	C	E	A
Openness	−0.04				
Conscientiousness	−0.02	0.13			
Extraversion	−0.06	0.19	−0.03		
Agreeableness	0.11	0.02	0.16	0.15	
Neuroticism	−0.09	−0.27*	−0.11	−0.37*	−0.04

*$p < 0.05$.

about four levels of embedding, beyond which performance deteriorates rapidly (Kinderman, Dunbar, & Bentall, 1998; Stiller & Dunbar, 2007). However, there are considerable individual differences in performance, with a sex difference favouring females (Stiller & Dunbar, 2007). As discussed in the Introduction, we have a greater expectation of a relationship with Agreeableness for this task than the task of Study 1.

Methods

Participants

Participants were an opportunity sample of 100 undergraduate students (50 males and 50 females), mostly from universities in Northern England. None of the participants had taken part in Study 1.

Five-factor personality measure

This personality measure was the same as for Study 1.

Theory of mind task

Participants hear a series of seven stories (five from Kinderman et al., 1998, with two more added by Stiller & Dunbar, 2007), each with a number of interacting characters, and answers several two-alternative forced-choice questions about each story. The questions are a combination of 30 factual memory items designed to check the participant has been attending, and 30 ToM questions, which probe their understanding of characters' mental states. The ToM questions are at varying levels of embedding. In level 2 items, the correct answer describes the mental state of a character (e.g. 'Emma wanted more money'), whereas in level 3 items, the correct answer describes the mental state of a character concerning the mental state of another character (e.g. 'Jenny thought the boss would believe Emma's story'), and so on with higher levels of embedding. The items cover all levels up to and including level 9. Over the set of stories, level 2 was tested seven times, level 3 five times, level 4 six times, level 5 five times, levels 6, 7 and 8 were each tested twice and level 9 was tested once (Stiller and Dunbar's study tested level 9 twice, but we deleted one of the items as we judged that the correct answer was indeterminate in the context of the story).

The ToM score was simply the number of correct answers given across the set of 30 ToM questions. Though there is an expectation of 15 correct answers by guessing, and Stiller and Dunbar use a more complex weighted mean to allow for the fact that there are different numbers of questions at each level of embedding, the simple mean is perfectly adequate for the purposes of comparing across participants for overall ToM performance.

Stories were digitally recorded and played back to ensure consistency in delivery. Participants, who were tested in groups of four to six, had the questions for each story on a separate sheet of paper, which they revealed only after they had heard the story in question. After all participants had completed all items, the next story was played, and then the next set of questions revealed.

Results

Personality scores

Cronbach's α statistics for the five personality scales were high (0.79–0.87). As in Study 1, females scored significantly more highly than males on Agreeableness ($d = 0.92, p < 0.01$)

and Neuroticism ($d = 0.74$, $p < 0.01$), whereas males scored more highly on Openness ($d = 0.52$, $p < 0.01$).

Theory of mind task

Participants answered the 30 memory questions in ToM task overwhelmingly correctly (mean 27.34, SD 1.66), indicating that they had attended to the stories. The proportion of ToM questions answered correctly declined with increasing level of embedding of item ($r_s = -0.91$, $p < 0.01$), from 95% for level 2 to 55% for level 8. The overall ToM score was reasonably normally distributed (Figure 1), with a range from just above chance guessing (which would produce a score of 15) to a perfect score (2 out of 100 participants). We computed split-half reliability, constructing the two halves to each span across the different levels of item difficulty (one half consisted of items at levels 3, 5, 7 and 9, and the others, items at levels 2, 4, 6 and 8). The resulting Spearman–Brown corrected reliability coefficient was 0.55. This is a lower reliability coefficient than the other tests reported in this paper. One reason for the low reliability may be the near-ceiling performance at the easy levels of embedding, and the near-chance performance at the hardest levels. This suspicion is confirmed by the fact that the item-total correlations for level 2 items (0.35) and level 9 items (0.38) are lower than for the mid-difficulty level 7 items (0.61).

The mean ToM score was 25.24 (SD 2.44), higher in females (mean 26.00, SD 1.75) than males (mean 24.80, SD 2.80). The sex difference was significant ($t = 3.25$, $p < 0.01$, $d = 0.51$).

Correlations of the five personality factors with each other and with ToM scores are shown in Table 2. As the table shows, Agreeableness and Neuroticism are significantly correlated with ToM score.
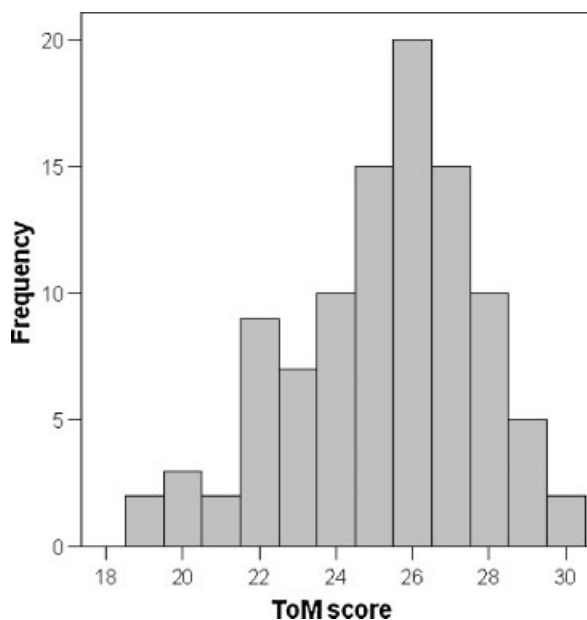

Figure 1.   Distribution of ToM scores, Study 2.

Table 2.   Correlations between ToM score and personality factors, and amongst the personality factors, Study 2

	ToM	O	C	E	A
Openness	0.16				
Conscientiousness	0.11	0.33*			
Extraversion	−0.11	−0.02	−0.25*		
Agreeableness	0.48*	0.11	0.20*	−0.05	
Neuroticism	0.40*	0.05	0.15	−0.11	0.23*

*$p < 0.05$.

To establish whether the effects of Agreeableness and Neuroticism are independent of one another, we ran regressions with ToM score as the independent variable, and Agreeableness, Neuroticism or both as the independent variables. R^2 values were 0.16 (Neuroticism alone), 0.23 (Agreeableness alone) and 0.31 (both predictors). The best model has both Neuroticism ($\beta = 0.30$) and Agreeableness ($\beta = 0.41$) as significant predictors (overall model: $F_{(2,97)} = 22.88$, $p < 0.001$).

It is possible that the observed relationship is an artefact produced by comparing across the two sexes, given the observed sex differences in ToM, and also in Agreeableness and Neuroticism. In order to test this, the regression above was rerun with sex as an additional independent variable. Sex did not emerge as a significant predictor ($t = 0.47$, n.s.), whilst the significance of Neuroticism and Agreeableness was undiminished.

Discussion

Performance on the ToM stories task seems to have been comparable to that found by Stiller and Dunbar (2007), though their use of a weighted mean makes comparison of overall scores unstraightforward. However, like them, we find near-perfect performance on the factual memory questions, very good performance on ToM questions up to level 5, and a decline in performance at higher levels than this. We also replicate the sex difference, favouring females, that they observe on this task.

The results show that, as predicted, Agreeableness and performance on this difficult social-cognitive ToM task covary fairly strongly (a correlation of 0.48 corresponds to 0.70 when attenuated for the reliability of both the Agreeableness and ToM measures). Conscientiousness, Extraversion and Openness are unrelated to ToM score. There was however a positive loading of ToM on Neuroticism. This latter relationship corroborates some previous findings in the literature. Neuroticism is strongly associated with depression and negative mood. Although severe depression is associated with ToM deficits (Lee, Harkness, Sabbagh, & Jacobson, 2005), mild depression or dysphoria actually enhances ToM performance (Gleicher & Weary, 1991; Harkness, Sabbagh, Jacobson, Chowdrey, & Chen, 2005). This can be related to the general hypothesis that (mild) negative mood serves to focus people on protecting the social connections and resources that they have by up-rating their social cognition (Allen & Badcock, 2003). It must be noted however, that one of the previous studies showing a ToM enhancement in mild depression used the social-perceptual eyes task (Harkness et al., 2005), whereas we found no Neuroticism loading using the eyes task in Study 1, but do find it in Study 2 using a social-cognitive stories task.

GENERAL DISCUSSION

In two separate studies, we find that the five-factor personality dimension of Agreeableness is unrelated to performance on a social-perceptual ToM task (the 'Reading the Mind in the Eyes' test), but strongly related to performance on a challenging social-cognitive ToM task involving reasoning about the mental states of characters in stories. This finding is congruent with several of findings from the literature, such us the impairment of psychopaths, who may be considered as the extreme of low Agreeableness, on social-cognitive but not social-perceptual ToM, and the correlations observed between social-cognitive ToM and measures of social support, social network size and social cooperation.

Some limitations of the study should be noted. First, ideally we would have tested both types of ToM in the same individuals. This was unfortunately not done as the two studies were done in successive years as part of an ongoing programme of personality research. Nonetheless, with sizable samples in both studies, a strong correlation in Study 2, and no correlation at all in Study 1, we feel justified in arguing that social-cognitive and social-perceptual ToM behave differently. Second, our assessment of the big five personality dimensions relies on a publicly available, relatively short self-report measure whose resolution is only at the broad domain, not the narrower facet level. However, recent research suggests that instruments like the one we used, and even much shorter ones, have good external validity and capture much of the variation measured by longer proprietary questionnaires (Gow, Whiteman, Pattie, & Deary, 2005; Gosling, Rentfrow, & Swann, 2003).

Third, although the Kinderman–Stiller stories task has been extensively used and validated as a measure of ToM (Kinderman et al., 1998; Paal & Bereczkei, 2007; Stiller & Dunbar, 2007), there are potential methodological issues surrounding it. The task is unlikely to be a completely pure ToM measure, since the questions vary not just in their level of ToM embedding, but also in their syntactic complexity. This is hard to avoid, since higher levels of embedding require more complex syntax, but it may mean that additional variation related, for example, to working memory is introduced. This raises the possibility that performance may be affected by intelligence. However, it should be noted firstly that the student participants in this study come from a programme that is academically highly selective, and thus come from a very narrow slice of the academic spectrum, which should keep intelligence variance to a minimum, and secondly, that IQ differences could not account for the observed sex differences in social-cognitive ToM performance, since there are no comparable sex differences in overall IQ.

The estimated reliability of the stories task is also quite low. There are probably several reasons for this. A substantial number of the items (levels 2 and 3) are very easy for this student population, and performance is near ceiling, whilst performance on the hardest item is near chance. This leaves a quite small number of items with good variation, and because of the two-alternative, forced-choice format, the proportion of correct answers expected by guessing is one half. A future possibility might be to revise the task with more of the items at the intermediate levels of difficulty and with more response options to mitigate the guessing problem.

We do not feel that the relatively low reliability of the social-cognitive ToM measure undermines our findings here. The low reliability should tend to obscure relationships that actually exist between social-cognitive ToM and other constructs, and thus the fact that we were able to detect such relationships despite it suggests that the relationships are strong.

Indeed, the dissociation between the results of studies 1 and 2 tends to strengthen our argument that social-cognitive ToM is related to Agreeableness whilst social-perceptual ToM is not; in Study 1, we found no relationship despite having a relatively reliable measure, and in Study 2, we found a relationship despite having a relatively unreliable one.

We believe that the results reported here help elucidate the cognitive basis of individual differences in self-reported Agreeableness. There might be several hypotheses about the nature of the causal relationship. One might be in terms of capacity. That is, high levels of cognitive ToM capacity allow individuals to be agreeable because they are better able to anticipate and infer the mental states of those around them. An alternative account would be in terms of motivation. We suspect that the motivational account is more likely to be true. Klein and Hodges (2001) show that the higher female performance on an empathy task disappears when the participants are given a motivational incentive, in the form of money. In related manner, we suggest that high-Agreeableness individuals simply have higher baseline motivation to attend to the mental states of others, and this affects what they recall from the stories task. This does not mean that low-Agreeableness individuals could not attend to and reason about mental states if directly incentivised to do so. This interpretation also links to the finding that Neuroticism is linked to ToM performance in Study 2. The anxiety that is potentiated by Neuroticism could serve as an additional motivator to attend to socially significant cues such as mental states. These hypotheses are directly testable, in that they predict that the provisions of incentives or the induction of anxiety should have measurable effects on social-cognitive ToM performance.

The results of this study are useful for several reasons. First, it is often argued, perhaps with some justice, that studies in the five-factor tradition have a certain circularity, since factors are derived inductively from ratings of self-reported behaviour, and then used to explain further types of self-reported behaviour, with the psychological mechanisms underlying the extracted factors never being independently identified (Block, 1995). This criticism is mitigated when five-factor theorists can show that objective measures of cognition coming from other research areas of psychology vary from individual to individual in ways that relate systematically to five-factor scores. The present study suggests a relationship between a broad cognitive mechanism—social-cognitive ToM— and a rating-derived personality trait, potentially paving the way for a greater understanding of the cognitive underpinnings of that trait.

Second, and on a related point, this cognitive study opens the way for a search for neural correlates of Agreeableness. In other cases where hypotheses have been developed about the mechanisms underlying personality traits, this has lead to brain imaging studies which find personality-related differences in the metabolic activity and even size of specific brain structures (see Whittle et al., 2006, for a review). The neural basis of ToM is a well-researched area (Sabbagh, 2004; Samson et al., 2004; Siegal & Varley, 2002). The hypothesis that Agreeableness relates to inter-individual variation in the activity of neural circuits responsible for social-cognitive ToM thus paves the way for a unification of these currently disparate literatures.

Third, this study helps further establish that Agreeableness is a distinct trait, quite separate from Conscientiousness and other constructs. Eysenck's (1992) contention that the Agreeableness and Conscientiousness dimensions have not been shown to be importantly distinct except in rating data cannot be upheld any longer. It would seem that Agreeableness is specially related to the propensity to consider the mental states of others.

Finally, the study further elucidates the sexually dimorphic nature of Agreeableness in humans. Women's higher Agreeableness is related to a spontaneously higher motivation to

attend to the mental states of others. Increasing Agreeableness, through increased ToM processing, enhances social harmony, but often at the expense of pursuing one's own status interests and goals (Nettle, 2006); for example, occupational success in large organisations is negatively predicted by Agreeableness (Boudreau, Boswell, & Judge, 2001). The sex difference in Agreeableness/ToM suggests that, over evolutionary time, women have on average gained more in fitness terms from harmonious social relationships relative to personal status than men have. There are two non-mutually exclusive reasons why this might be the case. First, men have high variance in reproductive success, and their reproductive success depends strongly on their status (see e.g. Pollet & Nettle, 2008). Thus, men potentially gain more by increasing status than women do, even at the expense of social harmony. Second, women invest more and for longer in their offspring after birth than men do. They often draw on networks of social support in order to do so successfully. It is thus more critical for them than for men to avoid risks and remain well-integrated into their local social network (Campbell, 1999). Because of these differing balances of costs and benefits, women may have evolved to be the more agreeable and hence more empathetic of the two sexes.

REFERENCES

Allen, N. B., & Badcock, P. B. T. (2003). The social risk hypothesis of depression: Evolutionary, psychosocial and neurobiological perspectives. *Psychological Bulletin*, *129*, 887–913.

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology*, *74*, 1531–1544.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, *21*, 37–46.

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*, 163–175.

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 of high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*, 241–251.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187–215.

Boudreau, J. W., Boswell, W. R., & Judge, T. A. (2001). Effects of personality on executive career success in the United States and Europe. *Journal of Vocational Behaviour*, *58*, 53–81.

Campbell, A. (1999). Staying alive: Evolution, culture and women's intra-sexual aggression. *Behavioral and Brain Sciences*, *22*, 203–252.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *135*, 653–665.

Costa, P. T., Terraciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*, 322–331.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*, 491–520.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of Openness/Intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*, 825–858.

Digman, J., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison and interpretation of six major studies. *Multivariate Behavioral Research*, *16*, 149–170.

Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine*, *34*, 1093–1102.

Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, *13*, 667–673.

Gleicher, F., & Weary, G. (1991). The effect of depression on the quality and quantity of social inferences. *Journal of Personality and Social Psychology*, *51*, 140–148.

Goldberg, L. R. (1990). An alternative 'description of personality': The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.

Goldberg, L. R. (1999). A broad-bandwith, public-domain, personality inventory measuring the lower levels of several five-factor models. In F. Ostendorf (Ed.), *Personality psychology in Europe* (Volume 7, pp. 7–28). Tilburg: Tilburg University Press.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the big five personality domains. *Journal of Research in Personality*, *37*, 504–528.

Gow, A. J., Whiteman, M. C., Pattie, A., & Dearie, I. J. (2005). Goldberg's IPIP Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences*, *39*, 317–329.

Graziano, W. G., & Eisenberg, N. H. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 795–825). San Diego: Academic Press.

Harkness, K. L., Sabbagh, M. A., Jacobson, J., Chowdrey, N., & Chen, T. (2005). Sensitivity to subtle social information in dysphoric college students: Evidence for an enhanced theory of mind. *Cognition and Emotion*, *19*, 999–1025.

Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual Differences*, *40*, 331–339.

Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of Personality*, *69*, 323–361.

Kinderman, P., Dunbar, R. I. M., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology*, *89*, 191–204.

Klein, K. J., & Hodges, S. D. (2001). Gender differences, motivation, and empathic accuracy. *Personality and Social Psychology Bulletin*, *27*, 720–730.

Lee, L., Harkness, K. L., Sabbagh, M. A., & Jacobson, J. A. (2005). Mental state decoding abilities in clinical depression. *Journal of Affective Disorders*, *86*, 247–258.

Meier, B. P., & Robinson, M. D. (2004). Does quick to blame mean quick to anger? The role of Agreeableness in dissociating blame and anger. *Personality and Social Psychology Bulletin*, *30*, 856–867.

Mohr, P., Howells, K., Gerace, A., Day, A., & Wharton, M. (2007) The role of perspective taking in anger arousal. *Personality and Individual Differences*, *43*, 507–517.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*, 622–631.

Nettle, D. (2007). Empathizing and systemizing: What are they, and what do they contribute to our understanding of psychological sex differences? *British Journal of Psychology*, *98*, 237–255.

Paal, T., & Bereczkei, T. (2007). Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences*, *43*, 541–551.

Pollet, T. V., & Nettle, D. (2008). Driving a hard bargain: Sex ratio and male marriage success in a historical U.S. population. *Biology Letters*, *4*, 31–33.

Richell, R. A., Mitchell, D. G. V., Newman, C., Leonard, A., Baron-Cohen, S., & Blair, R. J. R. (2003). Theory of mind and psychopathy: Can psychopathic individuals read the 'language of the eyes'? *Neuropsychologia*, *41*, 523–526.

Sabbagh, M. A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism. *Brain and Cognition*, *55*, 209–219.

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience*, *7*, 499–500.

Siegal, M., & Varley, R. (2002). Neural systems involved in 'theory of mind'. *Nature Neuroscience*, *3*, 463–471.

Soldz, S., & Vaillant, G. E. (1999). The big five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personality*, *33*, 208–232.

Stiller, J., & Dunbar, R. I. M. (2007). Perspective-taking and memory capacity predict social network size. *Social Networks*, *29*, 93–104.

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition*, *76*, 59–89.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655–684.

Whittle, S., Allen, N. B., Lubman, D. I., & Yücel, M. (2006). The neurobiological basis of temperament: Towards a better understanding of psychopathology. *Neuroscience and Biobehavioral Reviews*, *30*, 511–525.