HHS Public Access
Author manuscript
Eur J Pers. Author manuscript; available in PMC 2018 December 14.

Published in final edited form as:
Eur J Pers. 2017 ; 31(6): 599–613. doi:10.1002/per.2133.

Personality and Neural Correlates of Mentalizing Ability

Timothy A. Allen[a], **Amanda R. Rueter**[b], **Samantha V. Abram**[b], **James S. Brown**[c], and **Colin G. DeYoung**[b]

[a]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health

[b]Department of Psychology, University of Minnesota

[c]Department of Psychology, Morehouse College

Abstract

Theory of mind, or *mentalizing*, defined as the ability to reason about another's mental states, is a crucial psychological function that is disrupted in some forms of psychopathology, but little is known about how individual differences in this ability relate to personality or brain function. One previous study linked mentalizing ability to individual differences in the personality trait Agreeableness. Agreeableness encompasses two major subdimensions: Compassion reflects tendencies toward empathy, prosocial behavior, and interpersonal concern, whereas Politeness captures tendencies to suppress aggressive and exploitative impulses. We hypothesized that Compassion but not Politeness would be associated with better mentalizing ability. This hypothesis was confirmed in Study 1 ($N = 329$) using a theory of mind task that required reasoning about the beliefs of fictional characters. *Post hoc* analyses indicated that the honesty facet of Agreeableness was *negatively* associated with mentalizing. In Study 2 ($N = 217$), we examined whether individual differences in mentalizing and related traits were associated with patterns of resting-state functional connectivity in the brain. Performance on the theory of mind task was significantly associated with patterns of connectivity between the dorsal medial and core subsystems of the default network, consistent with evidence implicating these regions in mentalization.

Keywords

theory of mind; mentalizing; Agreeableness; default network; personality

---

Human beings must navigate many complex social and interpersonal situations. To drive a car safely from one place to another, we must coordinate with other drivers, who likewise expect safe passage. To get a raise at work, we appease our boss, whose own livelihood depends on our productivity. To maintain a positive romantic relationship, we attempt to accommodate our partner's career aspirations, financial goals, recreational interests, etc. The extent to which individuals are successful in navigating these dilemmas is dictated in part by

---

Please address correspondence concerning this article to: Timothy A. Allen, Centre for Addiction and Mental Health, 33 Russell Street Suite 1028, Toronto, ON, Canada, M5S 2S1 timothy.allen@camh.ca.

Conflict of Interest
The authors declare no conflict of interest.

individual differences in the ability to interpret social cues and understand the goals, feelings, and beliefs of others. This capacity to understand the minds of other people is referred to as “mentalizing” or “theory of mind” (Premack & Woodruff, 1978). Mentalizing has been studied extensively in terms of the underlying processes and neural systems that support it (Frith & Frith, 2006), its development in children (Carlson, Koenig, & Harms, 2013; Frith & Frith, 2003), and its role in clinical conditions, especially autism (Baron-Cohen, Leslie, & Frith, 1986). However, relatively little research has examined individual differences in mentalizing ability among healthy adults.

The paucity of research on variation in theory of mind in adults may be a function of the tasks that have been used to measure the construct. The original theory of mind tasks were designed for children (e.g., the false belief task, Wimmer & Perner, 1983), and their results tended to yield only categorical information (theory of mind is intact, or not). The need to assess more nuanced variation in mentalizing capabilities led to the development of new tasks designed for older children and adults. Studies employing these tasks led researchers to observe two apparently distinct, though correlated, components of mentalizing (Sabbagh, 2004; Tager-Flusberg & Sullivan, 2000). The *perceptual* component involves the use of facial cues and nonverbal behaviors to infer an actor’s mental state. In adults, this component has typically been measured using the ‘Reading the Mind in the Eyes test’ (RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), in which a participant infers an actor’s mental state from their eye gaze. In contrast, *cognitive* theory of mind taps the ability to reason about another person’s mental state, and is typically measured by asking the participant to identify an actor’s belief or intention after listening to a complex social story or scenario. Studies supporting the two components of mentalizing suggest that they may have distinct neural underpinnings, with the cognitive component relying on a circuit involving the medial prefrontal cortex and temporoparietal junction and the social-perceptual component relying on an orbitofrontal/medial temporal circuit that includes the amygdala (Sabbagh, 2004). Other evidence supporting the two components comes from studies of children with Williams syndrome, who appear to have intact perceptual theory of mind but impaired cognitive theory of mind (Tager-Flusberg & Sullivan, 2000).

In the present research, we investigated how cognitive theory of mind is related to personality and to patterns of functional connectivity in the adult brain. Some evidence already exists to link cognitive theory of mind to individual differences in normative personality traits, particularly Agreeableness. Agreeableness is one of the so-called *Big Five*, the major dimensions of covariation in any sufficiently comprehensive set of personality descriptors (John, Naumann, & Soto, 2008; Markon, Krueger, & Watson, 2005). It reflects individual differences in the capacity to coordinate one’s own goals with those of others, and encompasses a variety of traits related to altruism and cooperation (DeYoung, 2015). Empirical research on Agreeableness has repeatedly demonstrated its relevance for social and interpersonal functioning. Individuals high on Agreeableness tend to be more prosocial (Graziano & Habashi, 2010; Habashi, Graziano, & Hoover, 2016), better at suppressing aggressive impulses (Meier, Robinson, & Wilkowski, 2006), and less prejudicial toward others (Sibley & Duckitt, 2008). In contrast, low Agreeableness has been associated with tendencies toward psychopathy (Lynam et al., 2005), antisocial behavior (Miller & Lynam, 2001), and an increased likelihood both of being socially ostracized and of ostracizing others

(Hales, Kassner, Williams, Graziano, 2016). Unsurprisingly, these tendencies are especially evident in relationships, as those high on Agreeableness self-report higher levels of relationship satisfaction with intimate partners (Malouff et al., 2010), greater peer acceptance and friendship (Jensen-Campbell et al., 2002), and more perceived support from family members (Branje, van Lieshout, & van Aken, 2004).

One reason individuals high on Agreeableness are socially adept may be that they have advanced mentalizing capabilities. This would be consistent with empirical studies showing that mentalizing is positively correlated with many of same interpersonal outcomes as Agreeableness, including aggression (Mohr, Howells, Gerace, Day, & Wharton, 2007; Meier et al., 2006), size of one's support network (Asendorpf & Wilpers, 1998; Stiller & Dunbar, 2007), and social competence (Liddle & Nettle, 2006; Jensen-Campbell et al., 2002). Some evidence suggests that the link between mentalizing and Agreeableness is specific to cognitive theory of mind. For example, females tend to score more highly than males on both cognitive theory of mind tasks (Stiller & Dunbar, 2007; Baron-Cohen et al., 2001) and questionnaires assessing Agreeableness (Weisberg, DeYoung, & Hirsh, 2011), though only very small sex differences are noted on social-perceptual tasks (Kirkland, Peterson, Baker, Miller, & Pulos, 2013). Likewise psychopathy, which involves an extreme lack of empathy and concern for others, is negatively correlated with both Agreeableness and cognitive theory of mind, though it is unrelated to RMET performance (Dolan & Fullam, 2004; Jakobwitz & Egan, 2006).

The most direct test of the hypothesis that Agreeableness is linked specifically to cognitive theory of mind involved two studies in which participants were assessed using a self-report measure of the Big Five prior to completing a theory of mind task (Nettle & Liddle, 2008). In one sample, 96 participants completed the RMET to assess social-perceptual theory of mind. In a second sample, 100 participants completed a cognitive theory of mind task that required participants to track the beliefs and mental states of characters in complex narratives (Kinderman, Dunbar, & Bentall, 1998; Stiller & Dunbar, 2007). As predicted, Agreeableness was positively associated with performance on the cognitive, but not the perceptual, theory of mind task.

Nonetheless, several limitations are apparent in Nettle and Liddle's (2008) report. First, cognitive measures of theory of mind may conflate mentalizing ability with general intelligence. With each additional level of social-cognitive complexity in the narrative (e.g., X believes that Y believes that X feels… etc.), participants are required to attend not only to additional complexity in characters' mental states, but also to increasing complexity of the story as a whole, which is likely to require other cognitive skills in addition to perspective-taking (e.g., attention, working memory, verbal comprehension). In order to demonstrate that the relation between Agreeableness and task performance is due to improved mentalizing capabilities, it is therefore important to control for intelligence, which Nettle and Liddle (2008) did not do.

A second limitation is that the measure of personality included in their study assessed only the Big Five, and did not include subscales to assess more narrowly-defined traits. This is important because personality is structured hierarchically, such that traits residing at the top of the hierarchy reflect covariation among a broad range of behavior and experience, whereas those at lower levels of the hierarchy reflect more narrow and differentiated characteristics (Costa & McCrae, 1995; Jang, McCrae, Angleitner, Riemann, & Livesley, 1998; Markon et al., 2005). The Big Five constitute a relatively abstract level of the hierarchy, known as *domains* (Costa & McCrae, 1992). Both behavior-genetic and factor-analytic studies have shown that each of the Big Five domains can be decomposed into two correlated but distinct subfactors, known as *aspects*, which have meaningfully different correlates (Jang, Livesley, Angleitner, Riemann, & Vernon, 2002; DeYoung, Quilty, & Peterson, 2007).

The two aspects of Agreeableness are labeled *Compassion* and *Politeness* (DeYoung et al., 2007). Compassion reflects the tendency to experience concern and empathy for others, whereas Politeness reflects the tendency to refrain from being aggressive, exploitative, and socially disruptive. Note that trait Compassion differs from the compassion construct frequently employed in the social-cognitive literature, where it refers specifically to showing concern for others and is often considered distinct from empathy, with the latter typically referring to feeling others' emotions (Singer & Klimecki, 2014). (Notably, "empathy" is sometimes used in the social-cognitive literature as a broader term additionally encompassing theory of mind.) In contrast, trait Compassion is a broader construct that encompasses individual differences in both empathy and compassion in social-cognitive terms. For the remainder of this manuscript, when we use the term "Compassion" we are referring specifically to trait Compassion.

Previous research has shown that Compassion and Politeness differentially predict political ideology and moral values, prosocial decision-making, and various manifestations of psychopathology (Hirsh, DeYoung, Xu, & Peterson, 2010; Zhao, Ferguson, & Smillie, 2016; DeYoung, Carey, Krueger, & Ross, 2016), and there is reason to hypothesize that the two aspects are also differentially associated with theory of mind. For instance, questionnaires measuring individual differences in empathy tend to be highly related to theory of mind (Baron-Cohen & Wheelwright, 2004) and such measures are indicators of Compassion rather than Politeness (DeYoung et al., 2016). A structural neuroimaging study found that the neural substrates of empathy overlap with those of Compassion but not Politeness (Xin et al., in press). Finally, the Agreeableness scale from the measure used by Nettle and Liddle (2008; a 50-item Big Five measure from the International Personality Item Pool; IPIP-50; Goldberg, 1999) contains items that primarily reflect Compassion (e.g. "Feel others' emotions") rather than Politeness, and previously unpublished correlations in the Eugene Springfield community sample (Goldberg, 1999) show that IPIP-50 Agreeableness is much more strongly related to Compassion ($r = .89$) than to Politeness ($r = .47$). Taken together, this evidence suggests that the relation between personality traits and theory of mind may be best explained at the aspect-level, as opposed to the domain-level. The present study, therefore, aims to test the hypothesis that individual differences in Compassion, rather than Politeness, are associated with mentalizing ability.

When testing our primary hypothesis regarding the relation between Compassion and mentalizing, we encountered an unexpected finding that led us to develop a secondary hypothesis. Specifically, we noticed that Politeness was actually *negatively* related to mentalizing ability (though the effect was not significant once we controlled for IQ). This finding led us to hypothesize that a subfactor of Politeness relating to honesty and a lack of manipulativeness (in other words, the opposite of Machiavellianism), might be negatively related to mentalizing ability.

Theorists have long speculated that an intact theory of mind may be necessary for socially exploitative individuals to successfully perceive the mental states of others that they in turn manipulate (e.g., McIlwain, 2003), but little research has formally tested this hypothesis. One study in children found the ability to tell an effective lie was positively associated with theory of mind (Talwar, Gordon, & Lee, 2007). In another recent study, preschoolers who were initially unable to lie began to engage in deception following eleven days of theory of mind training (Ding, Wellman, Wang, Fu, & Lee, 2015).

In adults, performance on the RMET and other social-perceptual theory of mind tasks are negatively related to Machiavellianism (Ali & Chamorro-Premuzic, 2010). However, this association has rarely been investigated using cognitive theory of mind tasks. In what little research does exist on the topic, one study found no relation between Machiavellianism and theory of mind performance (Paal & Bereczkei, 2007), whereas another found that Machiavellianism was negatively related to theory of mind (Lyons, Caldwell, & Shultz, 2010). Clearly, the literature on Machiavellianism, dishonesty, and theory of mind is inconsistent and underdeveloped at present. We therefore seized on the unexpected Politeness finding that emerged from our primary analysis and explored the secondary hypothesis that socially exploitative, rather than aggressive, forms of low Politeness might be positively related to mentalizing ability.

Study 1

In our first study, we build on the findings of Nettle and Liddle (2008) by examining whether the relation between Agreeableness and theory of mind generalizes to the lower-order aspects of Agreeableness, controlling for intelligence. We hypothesized that individual differences in Compassion, but not Politeness, would be positively associated with behavioral performance on a cognitive theory of mind task, even when controlling for IQ. On an exploratory basis, we also examined the hypothesis that dishonesty is positively related to mentalizing ability, making use of additional personality questionnaires that more specifically assess maladaptive forms of Agreeableness. Although these additional questionnaires were designed to measure traits relevant to psychopathology, they were nonetheless designed to be valid in the general population, and evidence is accumulating that most such measures of personality psychopathology assess the same latent variables as the Big Five (Markon et al., 2005; Suzuki, Samuel, Pahlen, & Krueger, 2015; Stepp et al., 2012; Widiger & Trull, 2007).

Method

Participants—Right-handed participants were recruited via craigslist.com from the Minneapolis-Saint Paul metropolitan area, to participate in a study of the neural correlates of personality and cognition. The research protocol was approved by the University of Minnesota Institutional Review Board. All participants gave informed consent for behavioral and neuroimaging data collection.[1] Exclusion criteria included MRI safety issues, current use of psychotropic medications, history of neurological or psychiatric disorders (other than Attention Deficit/Hyperactivity Disorder), and current problems with alcohol or drug use. Participants (*N* = 329; 161 female) were selected for the present study if they had complete data on the Big Five and the theory of mind task. All participants were between the ages of 20 and 40 (*M* = 26.21, *SD* = 4.98). Participants identified as 72.64% White, 6.08% Black, .61% American Indian, 2.74% Hispanic/Latino, 3.65% Asian; 13.07% of participants reported that they were of mixed heritage; 1.22% chose not to report their ethnicity. The initial target sample size was 300, to provide adequate power to detect small-to-moderate effects even with some loss of data due to movement in the MRI scanner. However, some participants completed the behavioral portion of the study but opted out of the MRI portion, leading to a slightly larger total sample than originally anticipated for the behavioral portion.

Measures

Big Five Aspect Scales (BFAS): The BFAS (DeYoung et al., 2007) is a 100-item questionnaire assessment of the Big Five and their ten aspects. Participants use a five point scale to rate the extent to which they disagree or agree with each item, with higher scores reflecting greater agreement. Each of the ten aspect scales is composed of 10 items, and scores from each pair of aspects can be averaged to derive domain scores for the Big Five. Here, we focused on the subscales for the two aspects of Agreeableness, Compassion (α = .86) and Politeness (α = .76).

Externalizing Spectrum Inventory (ESI-BF): The ESI-BF (Patrick, Kramer, Krueger, & Markon, 2013) comprises160 items that constitute a brief version of the original 415-item ESI (Krueger, Markon, Patrick, Benning, & Kramer, 2007; Venables & Patrick, 2012). The ESI-BF maintains the organization of the original ESI, assessing three higher-order domains (general disinhibition, substance abuse, and callous aggression) and 23 lower-order facets. Participants rate each item on a 4-point scale, with higher scores corresponding to greater agreement with the item. The present study used the nine lower-order facets most strongly correlated with BFAS Agreeableness, including relational aggression, fraud, destructive aggression, irresponsibility, physical aggression, impatient urgency, boredom proneness, empathy, and honesty (α ranged from .72 – .88).

Personality Inventory for DSM-5 (PID-5): The PID-5 (Krueger, Derringer, Markon, Watson & Skodol, 2012) is a 220-item self-report questionnaire developed to measure the dimensional trait model of personality pathology proposed in Section III of the *Diagnostic and Statistical Manual, Fifth Edition* (*DSM-5*; American Psychiatric Association, 2013). The measure assesses 25 maladaptive personality traits that can be grouped into five broad domains that are akin to the Big Five. Facet scales range from 4 to 14 items, and participants rate items on a four point scale ranging from 0 (very false or often false) to 4 (very true or often true). The PID-5 has been extensively validated in both healthy and clinical populations (Hopwood, Thomas, Markon, Wright, & Krueger, 2012; Wright et al., 2012;

[1]We are unable to provide open access to the data used in this study because consent forms assured participants that their data would not be shared. Analytical scripts, supplemental analyses, and a list of all procedures and measures included in this study are openly available at the Open Science Framework: https://osf.io/g65vn/

Quilty, Ayearst, Chmielewski, Pollock, & Bagby, 2013). The present study used four of the five facets belonging to the Antagonism domain, including callousness, deceitfulness, manipulativeness, and grandiosity (α ranged from .41 – .70; note however, that the PID-5 was constructed via item response theory (IRT), and the quality of scales developed using IRT is not dependent on high α, as selection of items to cover the full range of the latent variable can actually decrease α. The fifth facet, attention-seeking, was excluded due to its low correlation with Agreeableness (see below for scale inclusion criteria). The Antagonism domain represents the maladaptive variant of Agreeableness (Gore & Widiger, 2013). Four participants did not complete the PID-5 and therefore are excluded from our analysis of Agreeableness facets.

Intelligence: IQ was estimated based on four subtests of the Wechsler Adult Intelligence Scale— Fourth Edition (WAIS–IV): Block Design, Matrix Reasoning, Vocabulary, and Similarities (Wechsler, 2008). IQ was above average (M = 113.49, SD = 15.62).

Theory of Mind: Theory of Mind was assessed using a series of five short stories depicting social situations (Dunbar, personal communication, February 3, 2010). Each story describes a social interaction involving multiple characters. Participants are asked to read each story to themselves twice, after which they answer 10 theory of mind questions and 10 memory questions pertaining to the story. All questions are in true-false format. Memory questions are designed to measure the participants' ability to retain the factual contents of the story, and the number of facts that the participant must retain varies between two and six in each question. Performance on memory questions within the task is used as a covariate to ensure that effects are due to participants' mentalizing ability rather than their memory for the details of the story. Theory of mind questions require that the participant reason, or infer, a character's perspective in the story. Questions vary across five levels of difficulty, with each successive level requiring the participant to track an additional character or level of perspective within the same character. For example, in second level questions, participants must track their own mental state and the mental state of one story character (e.g., "John wanted to go home after work"). In fourth level questions, participants track the mental state of three characters or perspectives in addition to their own (e.g., "John thought that Penny knew what Sheila wanted to do"). In order to assess performance on the task, we adopted the procedure used by Nettle and Liddle (2008) and computed simple sums of correct responses to memory questions and theory of mind questions for each participant.

Analytic Procedures: Descriptive statistics indicated that several of the self-report scales demonstrated excessive skew, which can attenuate correlations. As a result, scales exhibiting skew greater than .75 were log-transformed prior to being used in the analysis. For negatively skewed subscales, scores were reverse keyed, log-transformed, and then flipped in sign to retain their original keying. Multiple regression was used to test our primary hypothesis.

We adopted a factor analytic approach to identifying facets when assessing our exploratory hypothesis. Scales from the BFAS, ESI, and PID-5 were selected for inclusion in factor analysis based on two criteria: 1) the scale had a correlation of greater than .30 in absolute value with BFAS Agreeableness, and 2) the scale had its primary factor loading on the

Agreeableness vs. Antagonism domain, as assessed by previous research. ESI scales were judged only on the first criteria, as no factor analytic research has documented how the subscales load on Agreeableness. The attention-seeking subscale of the PID-5 was excluded because of its low correlation with BFAS Agreeableness ($r = -.16$), despite its previous assignment to Agreeableness (DeYoung et al., 2016; Krueger et al., 2012). Exploratory factor analysis was conducted with the 15 subscales from the ESI, PID-5, and BFAS that met both selection criteria, using principal-axis factoring with direct oblimin rotation ($\delta = 0$). Factor scores from the EFA were saved using the regression method for use in subsequent regression analyses.

Results

Descriptive Statistics and Bivariate Correlations—Table 1 shows descriptive statistics and bivariate correlations for the major study variables. The mean score on the memory questions of the story task indicated that participants were attending to and recalling the scenarios presented to them in the task. For theory of mind scores, greater level of complexity was associated with less accuracy, with accuracy declining from 98.42% on level 2 to 75.68% on level 6. Compassion was significantly associated with theory of mind performance at the zero-order, while Politeness and total Agreeableness were not.

Relation between Aspects and Theory of Mind—In order to test our primary hypothesis that Compassion, but not Politeness, is related to individual differences in theory of mind performance, we conducted a multiple linear regression with age, sex, and memory score as relevant covariates (Table 2). To investigate whether controlling for IQ had any effect on the results, we entered IQ in a second regression block. Results of the first regression, without IQ, indicated that the overall model was significant ($R^2 = .14$, $F_{(5, 323)} = 10.36$, $p < .001$). As hypothesized, Compassion was positively associated with theory of mind. Unexpectedly, Politeness was negatively associated with theory of mind. Adding IQ as a covariate significantly increased the amount of variance explained and reduced but did not eliminate the significant effect of Compassion. The effect of Politeness was no longer significant.

Factor Analysis of Agreeableness Facets—In order to test the hypothesis that a dishonesty subfactor was responsible for the weak negative correlation of Politeness with theory of mind, we first conducted an exploratory factor analysis of the Agreeableness subscales of the ESI, PID-5, and BFAS. Consistent with the notion that Agreeableness can be separated into two correlated aspects, the MAP test (O'Connor, 2000) indicated the presence of two factors across the 15 scales in the factor analysis (Table 3). Correlations between the BFAS aspect scales and the extracted factors showed that the two factors strongly resembled Compassion ($r = .86$, $p < .001$) and Politeness ($r = .74$, $p < .001$).[2] However, because we were interested in parsing subfactors within Politeness, we chose to extract three factors (Table 3). Based on the markers of the Politeness factor in the two-factor solution, as well as the facets and items that marked Politeness in the study leading to

[2]As a test of the robustness of our primary results, we saved the Compassion and Politeness factor scores from this analysis and repeated our initial regression analysis using the factors as simultaneous predictors in a regression predicting cognitive theory of mind. Results of the analysis were substantively identical to the analysis reported above.

the creation of the BFAS (DeYoung et al., 2007), we speculated that, with the extraction of a third factor, the Politeness factor would split into one factor reflecting absence of aggression and another reflecting honesty versus Machiavellianism. Results were consistent with this expectation. The first factor largely encompassed scales measuring various forms of aggression and impulsive behavior. In further analyses, we reversed the sign of the scores on this factor and refer to it as “Non-aggression,” so that all of our factors consistently indicated higher Agreeableness. The second factor, which we labeled “Compassion,” included subscales measuring empathy and concern for the emotional state of others. The third factor, which we labeled “Honesty,” included subscales measuring deceptiveness and entitlement.

Relation between 3-factor Solution and Theory of Mind—In addition to the three factor scores, our regression model again included age, sex, IQ, and memory scores as covariates (Table 4). The model accounted for 24.54% of the variance in theory of mind scores ($F_{(7, 317)} = 14.73$, $p < .001$). Consistent with our exploratory hypothesis, Honesty was negatively related to theory of mind performance, suggesting that dishonesty may be associated with more skillful mentalizing. Non-aggression was positively related to theory of mind performance. Finally, Compassion was also positively associated with theory of mind in this model, though the effect only barely met the threshold for significance. Importantly, however, the effect for Compassion is attenuated in this model due to suppression caused by Compassion’s positive association with Non-Aggression. When Non-Aggression was removed from the model, Compassion became more strongly predictive, $\beta = .14$, $SE = .11$, $p = .01$. The same pattern of suppression can be seen in the other direction as well. When Compassion was removed from the model, Non-Aggression became more strongly predictive, $\beta = .24$, $SE = .15$, $p = < .001$. This suggests that the shared variance of Compassion and Non-Aggression predicts theory of mind.

Scale-Level Analysis: At the suggestion of a reviewer and the editor, we performed supplemental analyses using the 15 Agreeableness scales. Partial correlations among theory of mind and the 15 scales (controlling for age, sex, IQ, movement, and memory scores) are presented in supplementary Table S1. Scale-level associations are generally consistent with the pattern indicated by associations with the three factors, but weaker. It is perhaps telling that the strongest association at the scale level is a positive one with Manipulativeness ($r = .15$, $p = .01$).

Discussion

Study 1 confirmed our hypothesis that mentalizing ability, as measured by performance on a cognitive theory of mind task, would be positively associated with individual differences in the personality trait Compassion, one aspect of the broader Big Five Agreeableness domain. This association did not generalize to Politeness, the other aspect of Agreeableness, suggesting that Nettle and Liddle’s (2008) original finding linking Agreeableness to mentalizing may have accidentally capitalized on the fact that the instrument they used to assess Agreeableness was heavily weighted toward Compassion rather than Politeness. The effect was present even after controlling for IQ, eliminating a potential confound in Nettle and Liddle’s study and confirming that the relation between Compassion and mentalizing ability is independent of any effect of general intelligence. Taken together, these results show that mentalizing, independently of general intellectual ability, is associated with increased tendencies toward empathic concern and caring for other people.

Our initial results suggested that Politeness might be negatively associated with mentalizing ability (though the effect was not significant after controlling for IQ). In *post hoc* analyses, using a variety of Agreeableness-related facet scales, we found that this effect was driven by an Honesty subfactor of Politeness. The finding that Honesty was negatively related to theory of mind provides some initial evidence in adults to support longstanding speculation that effective dishonesty, deception, and manipulation rely, in part, on the ability to reason about the mental states of others (for which some support has been found in childhood; Ding et al., 2015; Talwar et al., 2007). Nonetheless, we caution that replication of this effect is especially important given the *post hoc* nature of our analysis. This analysis also showed that Non-Aggression was, like Compassion, positively associated with mentalizing ability, suggesting that the positive association of ToM with personality extends to all parts of Agreeableness except those encompassed by Honesty, rather than being specific to Compassion.

Study 2

In our first study, we examined the personality correlates of mentalizing ability. In our second study, we investigated neural correlates of mentalizing ability and then tested whether those correlates would also be related to Agreeableness. One important neural substrate of theory of mind is likely to be brain's default network (DN; Andrews-Hanna, Smallwood, & Spreng, 2014). The DN was first discovered (and labeled "default") because it exhibits heightened activity during rest compared to many cognitive tasks, but its activity can be elicited by some tasks, such as episodic memory tasks, and it is active across a variety of contexts in which individuals are required to generate their own mental content. This includes the simulation of experience, whether it be when remembering past events, imagining future experience, evaluating ourselves, or—most importantly for our hypothesis—considering another person's experience (Andrews-Hanna et al., 2014). The DN comprises a core subsystem and two additional subsystems that covary strongly with the core in their neural activity. The core system has two large midline hubs in medial prefrontal cortex (mPFC) and the posterior cingulate cortex (PCC) plus adjacent precuneus. It also has a node in the anterolateral parietal cortex (alPC). A medial temporal subsystem includes the hippocampus and associated cortex, and a dorsal medial subsystem includes dorsomedial PFC (dmPFC), temporoparietal junction (TPJ), temporal pole, and superior temporal gyrus and sulcus (STGS) (Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010; Yeo et al., 2011).

Large-scale meta-analyses using the online database Neurosynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) have been employed to characterize the functions of the three DN subsystems (Andrews-Hanna et al., 2014). Results of these meta-analyses indicated that the medial temporal subsystem is primarily involved in autobiographical memory and recall, whereas the dorsal medial subsystem is most often linked to mentalizing, social cognition, theory of mind, and story comprehension. The core network is hypothesized to be important for transferring information between the two subsystems and is most often linked to the same functions as both subsystems as well as to affective judgments about self and others (Andrews-Hanna et al., 2014).

Based on these meta-analytic findings, the dorsal medial subsystem and the core are the two components of the DN most heavily involved in theory of mind. Indeed, a separate meta-analysis of task-based fMRI studies has also found associations between brain regions implicated in the dorsal medial and core DN subsystems and theory of mind (Schurz, Radua, Aichhorn, Richlan, & Perner, 2014). TPJ, in particular, has been repeatedly linked to the ability to reason about the mental states of others (Saxe & Wexler, 2005; Young, Dodell-Feder, & Saxe, 2010). Given these findings, our second study examines the hypotheses that individual differences in mentalizing ability will be associated with patterns of functional connectivity—that is, synchronized neural activity—within and between the dorsal medial and core DN subsystems. We also examine whether the functional connectivity of these DN subsystems is associated with the facets of Agreeableness identified in Study 1.

In order to examine functional connectivity, we used a data-driven approach known as spatial independent components analysis (ICA). ICA is similar to factor analysis, using brain voxels as variables. It identifies groups of voxels that tend to vary together in their activity levels over time. Compared to alternative connectivity methods (e.g., seed-based approaches), ICA is advantageous in that components reflecting artifactual patterns of covariance due to movement or boundary effects can be identified and excluded. For the present study, we identified intrinsic connectivity networks (ICNs) that corresponded to parts of the DN and then divided them into those that corresponded to the dorsal medial subsystem versus the core.

Method

Participants—The sample for study 2 consisted of a subset of participants from study 1 (N =289) with available resting state functional connectivity data. The final MRI sample consisted of 217 (111 female) participants following exclusions due to poor data quality (N = 9), not completing behavioral measures (N = 6) or excessive movement during the scan (N = 57; defined as mean absolute displacement above .5 mm). The mean age of the final sample was 25.77 (SD = 4.56). Participants identified as White (74.19%), Black (5.99%), American Indian (.46%), Hispanic/Latino (2.30%), Asian (4.15%); and mixed heritage (11.52%); 1.38% of participants elected not to report their ethnicity. Measures of personality and theory of mind were all identical to those in study 1.

Imaging Analysis

Image Acquisition and Preprocessing: Collection of neuroimaging data for this study and procedure for the ICA analysis have been outlined in more detail by Abram et al. (2015). Briefly, participants were asked to complete a minimal task designed to ensure wakefulness during the resting state fMRI scan. (The task consisted of pressing a button when a fixation cross turned from white to gray, five times in five minutes.) Images were acquired on a 3T Siemens Trio scanner. Sequence parameters for the rest scan included: gradient-echo echo-planar imaging of 150 volumes; repetition time (TR) 2 s; echo time (TE) 28 ms; flip angle 80°; voxel size 3.5 × 3.5 × 3.5 mm. A high-resolution T1-weighted structural scan was collected for registration. Standard preprocessing was completed using FMRIB Software Library (FSL 4.1.9) that included brain extraction, motion correction, grand mean intensity normalization of the 4D dataset, high-pass temporal filtering (at a filtering threshold of 0.1 Hz), and registration of functional images to high-resolution T1-weighted structural images (Wisner, Atluri, Lim, & MacDonald, 2013a; Wisner, Patzelt, Lim, & MacDonald, 2013b). Motion regression was completed as the final step.

Independent Components Analysis (ICA): ICNs were extracted using a meta-ICA procedure designed to optimize the reliability of the resulting network variables (Poppe et al., 2013). Twenty-five group-level probabilistic ICAs were completed using the MELODIC (Multivariate Exploratory Linear Optimized Decomposition into Independent Components) function in FSL (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC). Each ICA used a randomly generated subject order containing 80 of the participants as inputs; all of the subjects were not included in each ICA due to computational/hardware limitations and to reduce the likelihood of overfitting. We extracted 60 components based on reliability analyses conducted by Poppe et al. (2013). The 60 components from each ICA were concatenated into a single file, which was then used as the input to a meta-level MELODIC analysis (meta-ICA) to derive the 60 most consistent group-level components.

Component Selection: Procedures outlined by Kelly and colleagues (2010) were used to identify 31 artifactual components, including those which appeared to reflect cardiac function, respiration, nonneural fluctuations, white matter tracts, or movement. We next identified which of the remaining 29 nonartifactual components corresponded most closely to the DN as defined by Andrews-Hanna et al. using the 17-network parcellated resting functional connectivity maps derived from 1000 participants by Yeo et al. (2011). To compare our networks with their 17-network parcellation (https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation_Yeo2011), we calculated the percentage of cortical overlap of our components with their masks for the core and dorsal medial DN subsystems. The components with the highest overlap were then visually inspected to verify that they corresponded to the DN. This procedure yielded three components corresponding to the core subsystem—one located primarily in mPFC, one in PCC and precuneus, and one in precuneus and alPC—and two corresponding to the dorsal medial subsystem—one located primarily in dmPFC and one in TPJ and STGS. All five ICA-derived components and their overlap with the core and dorsal medial DN subsystems derived by Yeo and colleagues can be seen in Figure 1.

Computation of Intercomponent Connectivity Metrics: In order to assess connectivity *within* the dorsal medial and core subsystems, we derived intercomponent connectivity metrics (Abram et al., 2016; Wisner et al., 2013b) for each subject using the five DN components. Correlations between the time series of each component pair were computed and Fisher z-transformed. This process yielded one parameter for the two dorsal medial components and three for the core (three pairs among three variables). The latter three parameters were averaged together to create an index of connectivity within the core subsystem.

To examine patterns of interconnectivity *between* the dorsomedial and core subsystems, we computed intercomponent connectivity metrics between both dorsal medial components and the three core components, yielding six additional interconnectivity parameters. To reduce the dimensionality of these data for tractable hypothesis testing, all six interconnectivity parameters were entered into an exploratory factor analysis using principal-axis factoring with direct oblimin rotation (δ = 0). Using all six parameters as predictors in regression would create a high likelihood of suppression among the predictors that shared variance due to similar connectivity values. This likelihood is enhanced because each component was involved in computing multiple connectivity variables. Identifying a smaller number of factors underlying the covariance pattern among the connectivity variables allowed us to use predictors that were more likely to have discriminant validity and therefore not to suppress each other.

Velicer's MAP test (O'Connor, 2000) identified two underlying factors. The first factor was marked by the parameters reflecting interconnectivity between the three core components and the dmPFC component of the dorsal medial subsystem, whereas the second factor was marked by the parameters reflecting interconnectivity of the core components with the STGS/TPJ component of the dorsal medial subsystem (Table 5). Based on these factors, two interconnectivity scores were created by averaging scores according to their highest factor loadings, such that one score reflected connectivity between the core subsystem and the STG/TPJ component (STGS/TPJ–Core) and one reflected connectivity between the core and the dmPFC component (dmPFC–Core). (Using factor scores instead of averages did not substantively change the results.)

Results

Relation between Theory of Mind and the DN—Multiple linear regression was used to examine whether connectivity within and between the dorsal medial and core subsystems of the DN predicted performance on the cognitive theory of mind task. Age, sex, IQ, movement within the scanner (measured as root-mean-square head position change), and memory scores from the theory of mind task were entered as covariates. (None of our results changed substantively if we did not control for movement.) The two interconnectivity factors reflecting connectivity within the core subsystem and dorsomedial subsystem were entered as predictors, as were the two between-subsystem connectivity factors: STG/TPJ–Core and dmPFC–Core. Results of the regression are presented in Table 6. The overall model was significant (R^2 = .25, $F_{(9, 207)}$ = 7.46, $p < .001$), as was the connectivity variable indexing connectivity between the STGS/TPJ and the core subsystem.

Relation between Agreeableness Traits and the DN—We next examined whether connectivity between the core and dorsomedial subsystems of the DN was associated with individual differences in the three subfactors of Agreeableness identified in Study 1: Compassion, Honesty, and Non-Aggression. Each factor served as the dependent variable in a separate linear regression, controlling for age, sex, head motion, and the other two factors. Again four interconnectivity variables were entered, two representing connectivity within the DN subsystems and two representing connectivity between them (Table 7). Neither the within system nor the between system connectivity factors predicted individual differences in Compassion or Non-Aggression. However, the STG/TPJ–Core variable was significantly negatively associated with individual differences in the Honesty facet ($\Delta R^2 = .01$), though the *p*-value was so close to .05 that it would not survive Bonferroni correction for the three regressions.

Scale-Level Analysis—At the suggestion of a reviewer and the editor, we performed supplemental analyses using the 15 Agreeableness scales and the eight individual connectivity variables (rather than grouping the six between-network connectivity variables according to the two-factor solution). Partial correlations among theory of mind, the 15 scales and the connectivity variables (controlling for age, sex, IQ, movement, and memory scores) are presented in supplementary Tables S1–S2. The patterns for individual scales are generally consistent with the results we found using factor scores. It is notable that, as with theory of mind, Manipulativeness showed the strongest associations with neural variables; Deceitfulness and Honesty show similar, though slightly weaker, associations.

Supplemental Analyses—As an alternative method to dealing with multicollinearity among the connectivity variables, we conducted a set of stepwise regressions, in which the two within-system connectivity variables and the six between-system connectivity variables were entered as predictors of theory of mind and the three Agreeableness subfactors, respectively. Results indicated that connectivity between the STGS/TPJ region of the dorsal medial subsystem and the precuneus/alPC region of the core subsystem was positively associated with both theory of mind and the Compassion subfactor (supplementary Tables S3–S4). Theory of mind was also negatively predicted by connectivity between the dmPFC region of the dorsal medial subsystem and the PCC/precuneus component of the core network. None of the within- or between-system connectivity metrics were associated with scores on the Non-Aggression factor (supplementary Table S5). Connectivity between the STGS/TPJ component of the dorsal medial subsystem and PCC/precuneus component of the core subsystem was negatively associated with the Honesty subfactor (supplementary Table S6).

Discussion

Study 2 examined whether individual differences in mentalizing ability were associated with resting connectivity within and between subsystems of the DN. Connectivity between the core subsystem and part of the dorsal medial subsystem was significantly associated with mentalizing ability as assessed by a standard cognitive theory of mind task. These results are consistent with meta-analytic findings showing that these two subsystems are consistently activated by fMRI tasks requiring mentalizing (Andrews-Hanna et al., 2014; Schurz et al., 2014). Our results indicated that the link between mentalizing and functional connectivity within the DN is driven by communication between the core of the DN and the component of the dorsal medial subsystem that included the TPJ and STGS. This is consistent with previous social cognition research that has consistently found the TPJ and STGS to be activated during tasks that require one to reason about the beliefs and cognitions of others (Carrington & Bailey, 2009; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005).

Author Manuscript
Author Manuscript
Author Manuscript

We also examined whether connectivity within and between the dorsal medial and core subsystems was associated with individual differences in the personality traits related to mentalizing ability—namely, Compassion, Honesty, and Non-Aggression. The only significant association was a negative association between trait Honesty and connectivity between the STGS/TPJ region of the dorsal medial subsystem and the core network, and this did not survive correction for multiple tests. After reanalyzing our data using our six interconnectivity metrics (rather than the two aggregate variables based on factor analysis) as independent predictors of theory of mind and the three Agreeableness subfactors, we found some preliminary evidence that connectivity between the STGS/TPJ region of the dorsal medial subsystem and posterior components of the core subsystem was associated not only with theory of mind but also with Compassion and Honesty. Given the exploratory nature of these analyses however, more research will be necessary to determine whether connectivity between the dorsal medial and core subsystems, or any particular regions within the two subsytems, can account for the relations between Agreeableness and theory of mind.

In our supplemental analyses, we also found some evidence that the posterior portion of the core subsystem, involving the precuneus, may be more important for mentalizing than the anterior portion, as mentalizing was associated with connectivity between the STGS/TPJ component of the dorsal medial subsystem and a posterior but not anterior component. Interestingly, we also found an unpredicted negative association of mentalizing with connectivity between the dmPFC component of the dorsal medial subsystem and the posterior portion of the core subsystem. This suggests that different parts of the dorsal medial subsystem might compete in their influence on the core subsystem, in terms of their effects on mentalizing ability, but, again, this was an exploratory analysis, and replication would be necessary before giving this finding much weight.

General Discussion

The results of these two studies provide insight into the nature of individual differences in mentalizing ability. Study 1 showed that individual differences in Compassion—one of the two major subfactors or "aspects" of the Big Five Agreeableness domain—positively predicted mentalizing ability. This finding refines the results of a previous study, which found a similar link between mentalizing and Agreeableness when using an Agreeableness measure heavily weighted toward Compassion rather than Politeness (Nettle & Liddle, 2008). This finding also complements research showing that both mentalizing and Compassion are associated with more effective social and interpersonal functioning (Cassidy, Werner, Rourke, Zubernis, & Balaraman, 2003; Devine, White, Ensor, & Hughes, 2016; Sun, Kaufman, & Smillie, 2017). It is possible that the ability to accurately infer and reason about the mental state of others is a key psychological mechanism allowing individuals high in Compassion to behave in ways that foster greater cooperation and harmony within their relationships. Future research that formally tests this hypothesis, by examining whether mentalizing ability mediates the relation of Compassion and interpersonal behavior over time during development, would be an important next step toward confirming this causal pathway.

An unexpected finding was that Non-Aggression was also positively associated with mentalizing ability, suggesting that the positive association of mentalizing with subfactors of Agreeableness is not limited to Compassion. In fact, it was only the Honesty subfactor that was not positively related. The results of Study 1 also provided initial evidence for our secondary, *post hoc*, hypothesis that the Honesty subfactor of Politeness is negatively related to mentalizing ability. This finding is consistent with recent studies of children showing that the ability to effectively lie is positively associated with theory of mind capabilities (Talwar et al., 2007) and that theory of mind training can actually induce lying in previously honest preschoolers (Ding et al., 2015). Another study in children found that self-reported behaviors corresponding to our Honesty factor (such as lying and cheating) were positively related to performance on a theory of mind task that required children to reason about the beliefs, intentions, and thoughts of story characters (Lonigro, Laghi, Baiocco, & Baumgartner 2014). Consistent with the loading of manipulativeness on the Honesty factor in our study, mentalizing abilities have also been positively associated with persuasiveness in children (Slaughter, Peterson, & Moore, 2013). Our results extend these earlier findings in children to adults for the first time.

However, our results also stand in apparent contrast to adult studies that have found a negative relation between Machiavellianism and mentalizing abilities (Ali & Chamorro-Premuzic, 2010; Lyons, Caldwell, & Schultz, 2010). A likely explanation for this inconsistency is that these studies relied on the Mach-IV (Christie & Geis, 1970), an instrument that conflates dishonesty and manipulation with a variety of other constructs reflecting disagreeableness (e.g., immorality, distrust, cynicism). Our findings suggest that carefully parsing the different subfactors of Agreeableness is crucial for understanding the personality correlates of mentalizing ability, given that we found both Compassion and Non-Aggression to be positively associated with mentalizing ability and only Honesty to be negatively associated. Greater mentalizing ability, therefore, may facilitate most traits in the Agreeableness domain, even while leading to a specific increase in the tendency to lie and manipulate others—an interesting paradox. Given the post-hoc nature of our hypothesis about Honesty and the fact that we had no hypothesis about Non-Aggression, this conclusion should be replicated before one is confident in it.

In Study 2, we used a subset of our initial sample to examine the neural correlates of mentalizing abilities. Meta-analytic research has previously identified the DN and specifically its dorsal medial and core subsystems as important substrates of mentalizing in within-subjects analysis (Andrews-Hanna et al., 2014). Our results showed that these associations extend to individual differences. Mentalizing ability was positively associated with connectivity between the core subsystem and a component of the dorsal medial subsystem that includes the TPJ and the STGS. Interestingly, connectivity *within* the two subsystems was unrelated to mentalizing ability, suggesting that reasoning about the mental state of others may rely most heavily upon communication *between* the functional hubs of the core DN network and the STGS/TPJ. It is noteworthy that the connectivity effect we observed was specific to the STGS/TPJ component of the dorsal medial network, as previous fMRI studies have found increased activation in the bilateral TPJ and PCC (the latter being part of the core network) when participants read about the mental state of story characters (Saxe & Powell, 2006; Young et al., 2010).

In a secondary analysis in Study 2, we also examined whether connectivity of the dorsal medial and core subsystems was associated with individual differences in the three subfactors of Agreeableness— Compassion, Honesty, and Non-Aggression. Our initial analysis yielded no links between Compassion or Non-Aggression and connectivity within or between the core and dorsal medial DN subsystems. This is in contrast to a recent study of similar size linking trait empathy to functional connectivity within regions of the default network (Takeuchi et al., 2014). One possibility is that Compassion is linked to other parameters of the DN that were not investigated here; another is that the association is too weak to be detected reliably even in a sample as large as ours. A third possibility was supported by supplementary analyses we performed at the suggestion of the editor and a reviewer: it may be that the association between Compassion and DN connectivity is specific to links between TPJ/STGS and the posterior portion of the core subnetwork. Finally, in both our initial analyses and the supplementary analyses, we found some evidence for a possible association between trait Honesty and connectivity between the STGS/TPJ component and the core subsystem. Future studies should continue to investigate these possible associations.

A number of limitations are worthy of mention. First, our results rely on a measure of mentalizing that requires participants to simulate social narratives. Studies that that employ more diverse measures of cognitive theory of mind, including tasks that are better-suited to capturing real-time mentalizing abilities, will be essential to evaluating the generalizability of our results to real-world social behavior. Second, resting state scans in Study 2 were preceded by a series of task-based functional scans, and there is some recent evidence suggesting that task engagement can affect the subsequent coherence of networks at rest (Rosazza & Minati, 2011). Despite our relatively large sample size, it is also possible we lacked the necessary power to detect small effects between subdimensions of Agreeableness and the neural variables. Given the strength of these associations (Table 7 and Supplemental Tables S4–S6), future research on these effects should employ larger samples, as our observed (post-hoc) power to detect these effects was not much above 50%. Finally, our sample was largely restricted to healthy volunteers. Additional research will be needed to determine whether our findings are relevant to individuals diagnosed with clinical conditions that have been linked to altered mentalizing abilities (e.g., psychopathy, schizophrenia).

In sum, the results of this investigation provide a more extensive characterization of the psychological and neural correlates of normal individual differences in mentalizing ability than any previously published research. Deficits of theory of mind are already known to be involved in a range of clinical conditions, including autism (Baron-Cohen et al., 1986) and schizophrenia (Abram et al., 2016). Our results suggest that they may be involved broadly in disorders of the externalizing spectrum as well, and specifically the Antagonism (low Agreeableness) component of this spectrum, from which most of the Agreeableness-related facets in our assessment were drawn (Kotov et al., 2017; Krueger et al., 2007). However, this involvement is complicated by the fact that we found mentalizing ability to be associated in opposite directions with different subdimensions of Agreeableness. Having high levels of mentalizing ability may incline people to be compassionate and nonaggressive, but they may also incline people to be deceitful and manipulative. Additional research will be necessary to figure out when and why mentalizing ability leads to behavioral patterns at opposite ends of

Agreeableness. Such research is likely to be strengthened by the inclusion of personality measures that can isolate the unique effects of the various subdimensions of Agreeableness. Indeed, ignoring such effects may lead researchers to underestimate associations between Agreeableness and mentalizing, as opposing effects of different subdimensions may cancel one another out. We hope the present study provides the impetus for research leading to a more detailed understanding of the behavioral consequences and neural substrates of mentalizing ability, an understanding that may have important consequences for treating various disorders.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants to Colin G. DeYoung from the National Institute on Drug Abuse (R03 DA029177-01A1) and from the National Science Foundation (SES-1061817).

References

Abram SV, Wisner KM, Fox JM, Barch DM, Wang L, Csernansky JG, … Smith MJ. Fronto-temporal connectivity predicts cognitive empathy deficits and experiential negative symptoms in schizophrenia. Human Brain Mapping. 2016; 38:1111–1124. http://dx.doi.org/10.1002/hbm.23439. [PubMed: 27774734]

Abram SV, Wisner KM, Grazioplene RG, Krueger RF, MacDonald AW III, DeYoung CG. Functional coherence of insula networks is associated with externalizing behavior. Journal of Abnormal Psychology. 2015; 124:1079–1091. http://dx.doi.org/10.1037/abn0000078. [PubMed: 26301974]

Ali F, Chamorro-Premuzic T. Investigating theory of mind deficits in nonclinical psychopathy and Machiavellianism. Personality and Individual Differences. 2010; 49:169–174. http://dx.doi.org/10.1016/j.paid.2010.03.027.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5. Washington, DC: Author; 2013.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL. Functional-anatomic fractionation of the brain's default network. Neuron. 2010; 65:550–562. http://dx.doi.org/10.1016/j.neuron.2010.02.005. [PubMed: 20188659]

Andrews-Hanna JR, Smallwood J, Spreng RN. The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. Annals of the New York Academy of Sciences. 2014; 1316:29–52. http://dx.doi.org/10.1111/nyas.12360. [PubMed: 24502540]

Asendorpf JB, Wilpers S. Personality effects on social relationships. Journal of Personality and Social Psychology. 1998; 74:1531–1544. http://dx.doi.org/10.1037/0022-3514.74.6.1531.

Baron-Cohen S, Leslie AM, Frith U. Mechanical, behavioural and intentional understanding of picture stories in autistic children. British Journal of Developmental Psychology. 1986; 4:113–125. http://dx.doi.org/10.1111/j.2044-835x.1986.tb01003.x.

Baron-Cohen S, Wheelwright S. The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. Journal of Autism and Developmental Disorders. 2004; 34:163–175. http://dx.doi.org/10.1023/b:jadd.0000022607.19833.00. [PubMed: 15162935]

Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. 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. 2001; 42:241–251. http://dx.doi.org/10.1111/1469-7610.00715. [PubMed: 11280420]

Branje SJ, van Lieshout CF, van Aken MA. Relations between big five personality characteristics and perceived support in adolescents' families. Journal of Personality and Social Psychology. 2004; 86:615–628. http://dx.doi.org/10.1037/0022-3514.86.4.615. [PubMed: 15053709]

Carlson SM, Koenig MA, Harms MB. Theory of mind. Wiley Interdisciplinary Reviews: Cognitive Science. 2013; 4:391–402. http://dx.doi.org/10.1002/wcs.1232. [PubMed: 26304226]

Carrington SJ, Bailey AJ. Are there theory of mind regions in the brain? A review of the neuroimaging literature. Human Brain Mapping. 2009; 30:2313–2335. http://dx.doi.org/10.1002/hbm.20671. [PubMed: 19034900]

Cassidy KW, Werner RS, Rourke M, Zubernis LS, Balaraman G. The relationship between psychological understanding and positive social behaviors. Social Development. 2003; 12:198–221. http://dx.doi.org/10.1111/1467-9507.00229.

Christie, R., Geis, F. Scale construction. In: Christie, R., Geis, F., editors. Studies in Machiavellianism. New York: Academic Press; 1970. p. 388-400.

Costa, PT., McCrae, RR. Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. Odessa (FL): Psychological Assessment Resources; 1992.

Costa PT, McCrae RR. Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. Journal of Personality Assessment. 1995; 64:21–50. http://dx.doi.org/10.1207/s15327752jpa6401_2. [PubMed: 16367732]

Devine RT, White N, Ensor R, Hughes C. Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. Developmental Psychology. 2016; 52:758–771. http://dx.doi.org/10.1037/dev0000105. [PubMed: 26914214]

DeYoung CG. Cybernetic big five theory. Journal of Research in Personality. 2015; 56:33–58. http://dx.doi.org/10.1016/j.jrp.2014.07.004.

DeYoung CG, Carey BE, Krueger RF, Ross SR. Ten aspects of the big five in the Personality Inventory for DSM–5. Personality Disorders: Theory, Research, and Treatment. 2016; 7:113–123. http://dx.doi.org/10.1037/per0000170.

DeYoung CG, Quilty LC, Peterson JB. Between facets and domains: 10 aspects of the big five. Journal of Personality and Social Psychology. 2007; 93:880–896. http://dx.doi.org/10.1037/0022-3514.93.5.880. [PubMed: 17983306]

Ding XP, Wellman HM, Wang Y, Fu G, Lee K. Theory-of-mind training causes honest young children to lie. Psychological Science. 2015; 26:1812–1821. https://doi.org/10.1177/0956797615604628. [PubMed: 26431737]

Dolan M, Fullam R. Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. Psychological Medicine. 2004; 34:1093–1102. http://dx.doi.org/10.1017/S0033291704002028. [PubMed: 15554579]

Frith U, Frith CD. Development and neurophysiology of mentalizing. Philosophical Transactions of the Royal Society B: Biological Sciences. 2003; 358:459–473. https://doi.org/10.1098/rstb.2002.1218.

Frith CD, Frith U. The neural basis of mentalizing. Neuron. 2006; 50:531–534. https://doi.org/10.1016/j.neuron.2006.05.001. [PubMed: 16701204]

Goldberg, L. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In: Mervielde, I.Deary, IJ.De Fruyt, F., Ostendorf, F., editors. Personality psychology in Europe. Vol. 7. Tilburg, the Netherlands: Tilburg University Press; 1999. p. 7-28.

Gore WL, Widiger TA. The DSM-5 dimensional trait model and five-factor models of general personality. Journal of Abnormal Psychology. 2013; 122:816–821. http://dx.doi.org/10.1037/a0032822. [PubMed: 23815395]

Graziano WG, Habashi MM. Motivational processes underlying both prejudice and helping. Personality and Social Psychology Review. 2010; 14:313–331. http://dx.doi.org/10.1177/1088868310361239. [PubMed: 20435802]

Habashi MM, Graziano WG, Hoover AE. Searching for the prosocial personality: A big five approach to linking personality and prosocial behavior. Personality and Social Psychology Bulletin. 2016; 42:1177–1192. http://dx.doi.org/10.1177/0146167216652859. [PubMed: 27401167]

Hales AH, Kassner MP, Williams KD, Graziano WG. Disagreeableness as a cause and consequence of ostracism. Personality and Social Psychology Bulletin. 2016; 42:782–797. http://dx.doi.org/10.1177/0146167216643933. [PubMed: 27044246]

Hirsh JB, DeYoung CG, Xu X, Peterson JB. Compassionate liberals and polite conservatives: Associations of agreeableness with political ideology and moral values. Personality and Social Psychology Bulletin. 2010; 36:655–664. http://dx.doi.org/10.1177/0146167210366854. [PubMed: 20371797]

Hopwood CJ, Thomas KM, Markon KE, Wright AG, Krueger RF. DSM-5 personality traits and DSM–IV personality disorders. Journal of Abnormal Psychology. 2012; 121:424–432. http://dx.doi.org/10.1037/a0026656. [PubMed: 22250660]

Jakobwitz S, Egan V. The dark triad and normal personality traits. Personality and Individual Differences. 2006; 40:331–339. http://dx.doi.org/10.1016/j.paid.2005.07.006.

Jang KL, Livesley WJ, Angleitner A, Riemann R, Vernon PA. Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. Personality and Individual Differences. 2002; 33:83–101. http://dx.doi.org/10.1016/s0191-8869(01)00137-4.

Jang KL, McCrae RR, Angleitner A, Riemann R, Livesley WJ. Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of personality. Journal of Personality and Social Psychology. 1998; 74:1556–1565. http://dx.doi.org/10.1037/0022-3514.74.6.1556. [PubMed: 9654759]

Jensen-Campbell LA, Adams R, Perry DG, Workman KA, Furdella JQ, Egan SK. Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression. Journal of Research in Personality. 2002; 36:224–251. http://dx.doi.org/10.1006/jrpe.2002.2348.

John, O., Naumann, L., Soto, C. Paradigm shift to the integrative big five trait taxonomy. In: John, OP.Robins, RW., Pervin, LA., editors. Handbook of personality: Theory and research. New York: Guilford Press; 2008. p. 114-158.

Kelly RE, Alexopoulos GS, Wang Z, Gunning FM, Murphy CF, Morimoto SS, … Hoptman MJ. Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. Journal of Neuroscience Methods. 2010; 189:233–245. https://doi.org/10.1016/j.jneumeth.2010.03.028. [PubMed: 20381530]

Kinderman P, Dunbar R, Bentall RP. Theory-of-mind deficits and causal attributions. British Journal of Psychology. 1998; 89:191–204. http://dx.doi.org/10.1111/j.2044-8295.1998.tb02680.x.

Kirkland RA, Peterson E, Baker CA, Miller S, Pulos S. Meta-analysis Reveals Adult Female Superiority in "Reading the Mind in the Eyes Test". North American Journal of Psychology. 2013; 15:121.

Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, … Zimmerman M. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. Journal of Abnormal Psychology. 2017; 126:454–477. http://dx.doi.org/10.1037/abn0000258. [PubMed: 28333488]

Krueger RF, Derringer J, Markon KE, Watson D, Skodol AE. Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychological Medicine. 2012; 42:1879–1890. http://dx.doi.org/10.1017/S0033291711002674. [PubMed: 22153017]

Krueger RF, Markon KE, Patrick CJ, Benning SD, Kramer MD. Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. Journal of Abnormal Psychology. 2007; 116:645–666. http://dx.doi.org/10.1037/0021-843X.116.4.645. [PubMed: 18020714]

Liddle B, Nettle D. Higher-order theory of mind and social competence in school-age children. Journal of Cultural and Evolutionary Psychology. 2006; 4:231–244. http://dx.doi.org/10.1556/jcep.4.2006.3-4.3.

Lonigro A, Laghi F, Baiocco R, Baumgartner E. Mind reading skills and empathy: Evidence for nice and nasty ToM behaviours in school-aged children. Journal of Child and Family Studies. 2014; 23:581–590. http://dx.doi.org/10.1007/s10826-013-9722-5.

Lynam DR, Caspi A, Moffitt TE, Raine A, Loeber R, Stouthamer-Loeber M. Adolescent psychopathy and the big five: Results from two samples. Journal of Abnormal Child Psychology. 2005; 33:431–443. http://dx.doi.org/10.1007/s10648-005-5724-0. [PubMed: 16118990]

Lyons M, Caldwell T, Shultz S. Mind-reading and manipulation—Is Machiavellianism related to theory of mind? Journal of Evolutionary Psychology. 2010; 8:261–274. http://dx.doi.org/10.1556/jep.8.2010.3.7.

Malouff JM, Thorsteinsson EB, Schutte NS, Bhullar N, Rooke SE. The five-factor model of personality and relationship satisfaction of intimate partners: A meta-analysis. Journal of Research in Personality. 2010; 44:124–127. http://dx.doi.org/10.1016/j.jrp.2009.09.004.

Markon KE, Krueger RF, Watson D. Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. Journal of Personality and Social Psychology. 2005; 88:139–157. http://dx.doi.org/10.1037/0022-3514.88.1.139. [PubMed: 15631580]

McIlwain, D. Bypassing empathy: A Machiavellian theory of mind and sneaky power. In: Repacholi, B., Slaughter, V., editors. Individual differences in theory of mind: Implication for typical and atypical development (Macquarie Monographs in Cognitive Science). Hove, UK: Psychology Press; 2003. p. 39-66.

Meier BP, Robinson MD, Wilkowski BM. Turning the other cheek agreeableness and the regulation of aggression-related primes. Psychological Science. 2006; 17:136–142. http://dx.doi.org/10.1111/j.1467-9280.2006.01676.x. [PubMed: 16466421]

Miller JD, Lynam D. Structural models of personality and their relation to antisocial behavior: A meta-analytic review. Criminology. 2001; 39:765–798. http://dx.doi.org/10.1111/j.1745-9125.2001.tb00940.x.

Mohr P, Howells K, Gerace A, Day A, Wharton M. The role of perspective taking in anger arousal. Personality and Individual Differences. 2007; 43:507–517. https://doi.org/10.1016/j.paid.2006.12.019.

Nettle D, Liddle B. Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. European Journal of Personality. 2008; 22:323–335. http://dx.doi.org/10.1002/per.672.

O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behavior Research Methods. 2000; 32:396–402. http://dx.doi.org/10.3758/BF03200807.

Paal T, Bereczkei T. Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. Personality and Individual Differences. 2007; 43:541–551. http://dx.doi.org/10.1016/j.paid.2006.12.021.

Patrick CJ, Kramer MD, Krueger RF, Markon KE. Optimizing efficiency of psychopathology assessment through quantitative modeling: Development of a brief form of the Externalizing Spectrum Inventory. Psychological Assessment. 2013; 25:1332–1348. http://dx.doi.org/10.1037/a0034864. [PubMed: 24320765]

Poppe AB, Wisner K, Atluri G, Lim KO, Kumar V, MacDonald AW III. Toward a neurometric foundation for probabilistic independent component analysis of fMRI data. Cognitive, Affective, & Behavioral Neuroscience. 2013; 13:641–659. http://dx.doi.org/10.3758/s13415-013-0180-8.

Premack D, Woodruff G. Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences. 1978; 1:515–526. http://dx.doi.org/10.1017/s0140525x00076512.

Quilty LC, Ayearst L, Chmielewski M, Pollock BG, Bagby RM. The psychometric properties of the Personality Inventory for DSM-5 in an APA DSM-5 field trial sample. Assessment. 2013; 20:362–369. http://dx.doi.org/10.1177/1073191113486183. [PubMed: 23588687]

Rosazza C, Minati L. Resting-state brain networks: Literature review and clinical applications. Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology. 2011; 32:773–785. http://dx.doi.org/10.1007/s10072-011-0636-y.

Sabbagh MA. Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism. Brain and Cognition. 2004; 55:209–219. http://dx.doi.org/10.1016/j.bandc.2003.04.002. [PubMed: 15134854]

Saxe R, Kanwisher N. People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind". Neuroimage. 2003; 19:1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1. [PubMed: 12948738]

Saxe R, Powell LJ. It's the thought that counts: specific brain regions for one component of theory of mind. Psychological Science. 2006; 17:692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x. [PubMed: 16913952]

Saxe R, Wexler A. Making sense of another mind: the role of the right temporo-parietal junction. Neuropsychologia. 2005; 43:1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013. [PubMed: 15936784]

Schurz M, Radua J, Aichhorn M, Richlan F, Perner J. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. Neuroscience & Biobehavioral Reviews. 2014; 42:9–34. http://dx.doi.org/10.1016/j.neubiorev.2014.01.009. [PubMed: 24486722]

Sibley CG, Duckitt J. Personality and prejudice: A meta-analysis and theoretical review. Personality and Social Psychology Review. 2008; 12:248–279. http://dx.doi.org/10.1177/1088868308319226. [PubMed: 18641385]

Singer T, Klimecki OM. Empathy and compassion. Current Biology. 2014; 24:R875–R878. https://doi.org/10.1016/j.cub.2014.06.054. [PubMed: 25247366]

Slaughter V, Peterson CC, Moore C. I can talk you into it: Theory of mind and persuasion behavior in young children. Developmental Psychology. 2013; 49:227–231. http://dx.doi.org/10.1037/a0028280. [PubMed: 22545834]

Stepp SD, Yu L, Miller JD, Hallquist MN, Trull TJ, Pilkonis PA. Integrating competing dimensional models of personality: Linking the SNAP, TCI, and NEO using Item Response Theory. Personality Disorders: Theory, Research, and Treatment. 2012; 3:107–126. http://dx.doi.org/10.1037/a0025905.

Stiller J, Dunbar RI. Perspective-taking and memory capacity predict social network size. Social Networks. 2007; 29:93–104. http://dx.doi.org/10.1016/j.socnet.2006.04.001.

Sun, J., Kaufman, SB., Smillie, LD. Unique associations between big five personality aspects and multiple dimensions of well-being. Journal of Personality. 2017. http://dx.doi.org/10.1111/jopy.12301

Suzuki T, Samuel DB, Pahlen S, Krueger RF. DSM-5 alternative personality disorder model traits as maladaptive extreme variants of the five-factor model: An item-response theory analysis. Journal of Abnormal Psychology. 2015; 124:343–354. http://dx.doi.org/10.1037/abn0000035. [PubMed: 25665165]

Tager-Flusberg H, Sullivan K. A componential view of theory of mind: Evidence from Williams syndrome. Cognition. 2000; 76:59–90. http://dx.doi.org/10.1016/s0010-0277(00)00069-x. [PubMed: 10822043]

Takeuchi H, Taki Y, Nouchi R, Sekiguchi A, Hashizume H, Sassa Y, … Nakagawa S. Association between resting-state functional connectivity and empathizing/systemizing. NeuroImage. 2014; 99:312–322. https://doi.org/10.1016/j.neuroimage.2014.05.031. [PubMed: 24844739]

Talwar V, Gordon HM, Lee K. Lying in the elementary school years: verbal deception and its relation to second-order belief understanding. Developmental Psychology. 2007; 43:804–810. http://dx.doi.org/10.1037/0012-1649.43.3.804. [PubMed: 17484589]

Venables NC, Patrick CJ. Validity of the Externalizing Spectrum Inventory in a criminal offender sample: Relations with disinhibitory psychopathology, personality, and psychopathic features. Psychological Assessment. 2012; 24:88–100. http://dx.doi.org/10.1037/a0024703. [PubMed: 21787091]

Wechsler, D. Wechsler Adult Intelligence Scale —Fourth Edition. San Antonio, TX: Psychological Corporation; 2008.

Weisberg YJ, DeYoung CG, Hirsh JB. Gender differences in personality across the ten aspects of the big five. Frontiers in Psychology. 2011; 2:178. http://dx.doi.org/10.3389/fpsyg.2011.0017822. [PubMed: 21866227]

Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. American Psychologist. 2007; 62:71–83. http://dx.doi.org/10.1037/0003-066X.62.2.71. [PubMed: 17324033]

Wimmer H, Perner J. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition. 1983; 13:103–128. https://doi.org/10.1016/0010-0277(83)90004-5. [PubMed: 6681741]

Wisner KM, Atluri G, Lim KO, MacDonald AW. Neurometrics of intrinsic connectivity networks at rest using fMRI: retest reliability and cross-validation using a meta-level method. NeuroImage. 2013; 76:236–251. https://doi.org/10.1016/j.neuroimage.2013.02.066. [PubMed: 23507379]

Wisner KM, Patzelt EH, Lim KO, MacDonald AW III. An intrinsic connectivity network approach to insula-derived dysfunctions among cocaine users. The American Journal of Drug and Alcohol Abuse. 2013b; 39:403–413. http://dx.doi.org/10.3109/00952990.2013.848211. [PubMed: 24200210]

Wright AG, Thomas KM, Hopwood CJ, Markon KE, Pincus AL, Krueger RF. The hierarchical structure of DSM-5 pathological personality traits. Journal of Abnormal Psychology. 2012; 121:951–957. http://dx.doi.org/10.1037/a0027669. [PubMed: 22448740]

Xin H, Allen TA, Wei D, Huang H, Wang K, DeYoung CG, Qui Jiang. Trait compassion is associated with the neural substrate of empathy. Cognitive, Affective, and Behavioral Neuroscience. in press.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. Nature Methods. 2011; 8:665–670. http://doi.org/10.1038/nmeth.1635. [PubMed: 21706013]

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, … Fischl B. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology. 2011; 106:1125–1165. http://dx.doi.org/10.1152/jn.00338.2011. [PubMed: 21653723]

Young L, Dodell-Feder D, Saxe R. What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. Neuropsychologia. 2010; 48:2658–2664. https://doi.org/10.1016/j.neuropsychologia.2010.05.012. [PubMed: 20470808]

Zhao, K., Ferguson, E., Smillie, LD. Individual differences in good manners rather than compassion predict fair allocations of wealth in the dictator game; Journal of Personality. 2016. p. 1-13.http://doi.org/10.1111/jopy.12237

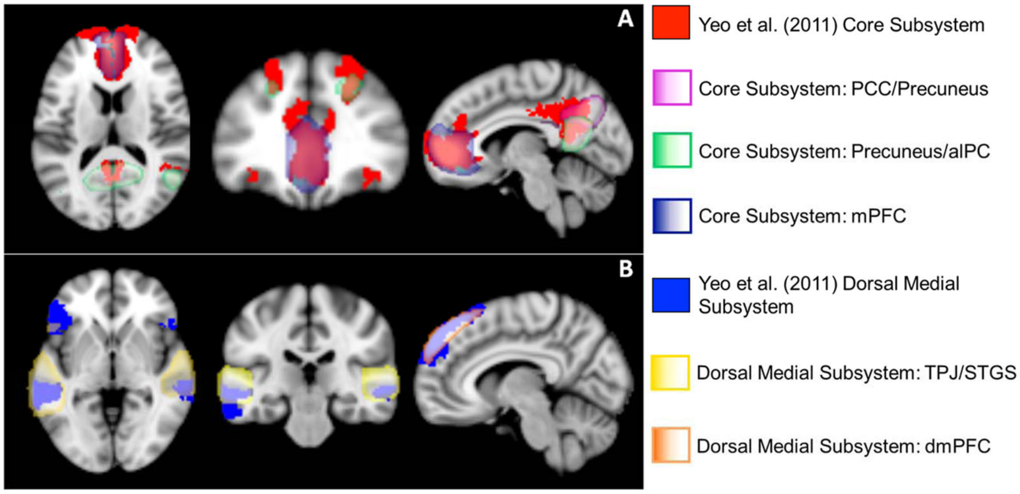

Figure 1.
Comparison of ICA-derived intrinsic connectivity networks with core and dorsal medial default network subsystems derived by Yeo and colleagues (2011). **A)** Three ICA-derived networks corresponded to the core subsystem. **B)** Two ICA-derived networks corresponded to the dorsal medial subsystem. Images are in neurological convention (i.e., left is left). Abbreviations: PCC = posterior cingulate cortex; mPFC = medial prefrontal cortex; alPC = anterolateral parietal cortex; TPJ = temporoparietal junction; STGS = superior temporal gyrus and sulcus; dmPFC = dorsomedial prefrontal cortex.

Table 1

Descriptive Statistics and Bivariate Correlations between BFAS Agreeableness scales and Theory of Mind Test.

Scale	*M (SD)*	1	2	3	4	5	6	7
1. Sex	---	---						
2. IQ	113.49 (15.61)	.04	---					
3. BFAS-Agreeableness	3.99 (.50)	−.32**	.03	---				
4. BFAS-Compassion	4.23 (.55)	−.27**	.11*	.82**	---			
5. BFAS-Politeness	3.77 (.63)	−.27**	−.05	.86**	.42**	---		
6. Memory Score	21.96 (1.30)	−.12*	.22**	.15**	.16**	.10	---	
7. Theory of Mind	20.91 (1.92)	−.08	.37**	.10	.18**	−.01	.35**	---

Notes. $N = 329$,

* $p < .05$,

** $p < .01$.

Table 2

Multiple Regression Predicting Theory of Mind from Trait Compassion and Politeness

	Without IQ						*With IQ*					
Measure	***β***	***B***	***SE***	***B 95% CI***	***p***	**ΔR^2**	***β***	***B***	***SE***	***B 95% CI***	***p***	**ΔR^2**
1. Sex (0 = F, 1 = M)	−.04	−.17	.21	[−.58, .24]	.42		−.06	−.24	.20	[−.63, .15]	.23	
2. Age	−.10	−.04	.02	[−.08, .00]	.05		−.09	−.03	.02	[−.07, .00]	.08	
3. Memory Score	.30	4.28	.76	[2.78, 5.77]	<.001		.23	3.34	.74	[1.88, 4.80]	<.001	
4. BFAS-Politeness	−.12	−.36	.18	[−.70, −.01]	.04	.01	−.08	−.25	.17	[−.58, .08]	.14	
5. BFAS-Compassion	.17	2.40	.83	[.76, 4.03]	.004	.03	.12	1.75	.80	[.18, 3.32]	.03	.01
6. IQ							.30	.04	.01	[.03, .05]	<.001	.08

Note. $N = 329$.

Table 3

Exploratory Factor Analysis of Agreeableness-Related Traits from the ESI, PID-5, and BFAS

		2-Factor Solution		3-Factor Solution		
	***r* with BFAS Agreeableness**	**Compassion**	**Politeness**	**Aggression**	**Compassion**	**Honesty**
ESI: Relational Aggression	−.52	−.39	**−.75**	**.74**	−.40	−.57
ESI: Fraud	−.33	−.18	**−.69**	**.72**	−.18	−.51
ESI: Destructive Aggression	−.35	−.28	**−.60**	**.70**	−.28	−.37
ESI: Irresponsibility	−.30	−.16	**−.63**	**.69**	−.15	−.44
ESI: Physical Aggression	−.35	−.34	**−.48**	**.61**	−.36	−.24
ESI: Impatient Urgency	−.35	−.22	**−.52**	**.47**	−.22	−.44
ESI: Boredom Proneness	−.35	−.28	**−.44**	**.41**	−.29	−.36
ESI: Empathy	.72	**.92**	.43	−.38	**.91**	.35
BFAS: Compassion	.81	**.80**	.35	−.30	**.80**	.29
PID-5: Callousness	−.72	−.71	**−.73**	.67	**−.71**	−.59
PID-5: Deceitfulness	−.49	−.32	**−.74**	.50	−.31	**−.88**
PID-5: Manipulativeness	−.37	−.14	**−.68**	.47	−.13	**−.77**
BFAS: Politeness	.86	.50	**.70**	−.55	.51	**.67**
ESI: Honesty	.44	.33	**.53**	−.37	.34	**.58**
PID-5 Grandiosity	−.39	−.24	**−.48**	.31	−.24	**−.55**

Notes. $N = 325$. Numbers in bold indicate the highest loading in each row for each analysis. Factor loadings are from structure matrices. All correlations in first column significant at $p < .01$.

Table 4

Multiple Regression Predicting Theory of Mind from three Agreeableness Factors.

Measure	β	*B*	*SE*	*B 95% CI*	*p*	ΔR^2
1. Sex (0 = F, 1 = M)	−.03	−.10	.20	[−.51, .30]	.61	
2. Age	−.09	−.03	.02	[−.07, .00]	.08	
3. IQ	.24	.03	.01	[.02, .04]	<.001	
4. Memory Score	.24	3.41	.73	[1.99, 4.84]	<.001	
5. Non-Aggression	.21	.43	.15	[.14, .72]	.004	.02
6. Compassion	.11	.22	.11	[.001, .45]	.049	.01
7. Honesty	−.25	−.51	.14	[−.78, −.24]	<.001	.03

Notes. N = 325.

Table 5

Exploratory Factor Analysis of Default Network Dorsal Medial and Core Subsystem Interconnectivity

Measure	STGS/TPJ–Core	dmPFC–Core
STGS/TPJ–PCC/Precuneus	**.81**	.24
STGS/TPJ–mPFC	**.70**	.32
STGS/TPJ–Precuneus/alPC	**.69**	.19
dmPFC– Precuneus/alPC	.33	**.72**
dmPFC–mPFC	.09	**.68**
dmPFC–PCC/Precuneus	.44	**.60**

Notes. N = 217. Numbers in bold indicate the highest loading in each row. Abbreviations: STGS = superior temporal gyrus and sulcus; TPJ = temporoparietal junction; alPC = anterolateral parietal cortex; PFC = prefrontal cortex; dmPFC = dorsomedial prefrontal cortex; mPFC = medial prefrontal cortex; PCC = posterior cingulate cortex.

Table 6

Multiple Regression Predicting Theory of Mind from DN Subsystem Interconnectivity

Measure	*β*	*B*	*SE*	*B 95% CI*	*p*	ΔR^2
1. Age	−.05	−.02	.03	[−.07, .03]	.47	
2. Sex (0 = F, 1 = M)	−.12	−.43	.23	[−.89, .03]	.07	
3. Full Scale IQ	.34	.04	.01	[.03, .06]	<.001	
4. Movement	−.07	−1.21	1.23	[−3.63, 1.21]	.33	
5. Memory Score	.25	3.49	.88	[1.76, 5.21]	<.001	
6. Core Interconnectivity	−.04	−.40	.81	[−1.99, 1.20]	.63	
7. Dorsal Medial Interconnectivity	−.08	−.66	.66	[−1.97, .64]	.32	
8. dmPFC–Core	−.10	−1.14	.92	[−2.96, .68]	.22	
9. STGS/TPJ–Core	.22	2.05	.73	[.60, 3.49]	.006	.03

Notes. $N = 217$. Abbreviations: STGS = superior temporal gyrus and sulcus; TPJ = temporoparietal junction; dmPFC = dorsomedial prefrontal cortex.

Table 7

Multiple Regression Predicting Lower-Order Agreeableness Traits from Default Network Subsystem Interconnectivity

	Compassion					Non-Aggression					Honesty				
Measure	***β***	***B***	***SE***	***95% CI***	***p***	***β***	***B***	***SE***	***95% CI***	***p***	***β***	***B***	***SE***	***95% CI***	***p***
1. Age	.07	.02	.01	[−.01, .04]	.27	−.05	−.01	.01	[−.03, .01]	.31	.06	.01	.01	[−.01, .03]	.25
2. Sex (0=F, 1=M)	−.12	−.23	.13	[−.50, .03]	.08	−.19	−.35	.10	[−.54, −.15]	.001	−.03	−.06	.11	[−.27, .16]	.60
3. Movement	.01	.13	.63	[−1.12, 1.38]	.84	−.06	−.55	.48	[−1.50, .39]	.25	−.03	−.23	.51	[−1.25, .78]	.65
4. Core Intercon.	.03	.18	.42	[−.65, 1.01]	.66	−.06	−.32	.32	[−.94, .31]	.32	.02	.11	.34	[−.57, .78]	.76
5. Dorsal Medial Intercon.	−.04	−.17	.35	[−.85, .51]	.63	−.08	−.31	.26	[−.83, .20]	.23	.06	.23	.28	[−.32, .78]	.42
6. STGS/TPJ–Core	.08	.41	.39	[−.35, 1.17]	.28	.05	.24	.29	[−.34, .81]	.41	−.14	−.64	.31	[−1.25, −.03]	.04
7. dmPFC–Core	.02	.15	.48	[−.80, 1.09]	.76	.06	.33	.36	[−.38, 1.05]	.36	.001	.01	.39	[−.76, .77]	.99
8. Compassion	–	–	–	–	–	.22	.20	.05	[.10, .30]	<.001	.11	.10	.06	[−.01, .21]	.07
9. Non-Aggression	.33	.35	.09	[.18, .53]	<.001	–	–	–	–	–	.58	.58	.06	[.46, .70]	<.001
10. Honesty	.14	.15	.09	[−.01, .32]	.07	.51	.51	.06	[.40, .61]	<.001	–	–	–	–	–

Notes. N = 217. Abbreviations: STGS = superior temporal gyrus and sulcus; TPJ = temporoparietal junction; dmPFC = dorsomedial prefrontal cortex.