AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Psychopathology and Clinical Science

ISSN: 2769-7541

2023, Vol. 132, No. 6, 733–748
https://doi.org/10.1037/abn0000846

Neurobehavioral Indices of Gaze Perception Are Associated With Social Cognition Across Schizophrenia Patients and Healthy Controls

Scott D. Blain[1, 2], Stephan F. Taylor[1, 3], Saige E. Rutherford[1, 4], Carly A. Lasagna[3], Beier Yao[5, 6], Mike Angstadt[7], Michael F. Green[8, 9], Timothy D. Johnson[10], Scott Peltier[7], Vaibhav A. Diwadkar[11], and Ivy F. Tso[1, 2]

[1] Department of Psychiatry, University of Michigan
[2] Department of Psychiatry & Behavioral Health, The Ohio State University
[3] Department of Psychology, University of Michigan
[4] Donders Center for Medical Neuroscience, Nijmegen, Netherlands
[5] Schizophrenia and Bipolar Disorder Program, McLean Hospital, Belmont, Massachusetts, United States
[6] Department of Psychiatry, Harvard Medical School
[7] Functional MRI Lab, University of Michigan
[8] Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles
[9] Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California, United States
[10] Department of Biostatistics, University of Michigan
[11] Department of Psychiatry & Behavioral Neurosciences, Wayne State University

Background: Gaze perception is a basic building block of social cognition, which is impaired in schizophrenia (SZ) and contributes to functional outcomes. Few studies, however, have investigated neural underpinnings of gaze perception and their relation to social cognition. We address this gap. ***Method:*** We recruited 77 SZ patients and 71 healthy controls, who completed various social-cognition tasks. During functional magnetic resonance imaging, participants (62 SZ, 54 controls) completed a gaze-perception task, where they judged whether faces with varying gaze angles were self-directed or averted; as a control condition, participants identified stimulus gender. Activation estimates were extracted based on (a) task versus baseline, (b) gaze-perception versus gender-identification, (c) parametric modulation by perception of stimuli as self-directed versus averted, and (d) parametric modulation by stimulus gaze angle. We used latent variable analysis to test associations among diagnostic group, brain activation, gaze perception, and social cognition. ***Results:*** Preferential activation to gaze perception was observed throughout dorsomedial prefrontal cortex,

This article was published Online First June 29, 2023.

Scott D. Blain https://orcid.org/0000-0002-2751-3629

Preliminary analyses were published as abstracts in *Biological Psychiatry* and *Schizophrenia Bulletin*. The article was also published as a preprint: https://psyarxiv.com/m7u65/.

The authors thank Pan Gu, Merranda McLaughlin, Savanna Mueller, Tyler Grove, Erica Shulte, Zarina Kraal, and Laura Locarno for their assistance.

This research was supported by the National Institutes of Health (5KL2TR000434, 5K23MH108823, R01MH122491 to Ivy F. Tso). This material is also based upon work supported by the National Science Foundation Graduate Research Fellowship (DGE-1841052 to Carly A. Lasagna).

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The study was conducted in accordance with the protocol approved by the Institutional Review Board of the University of Michigan Medical School (HUM00080457, "Neural Mechanisms of Gaze Perception in Psychosis"), and written informed consent was obtained from each participant. Legally, the data cannot be made publicly available online because of the wording of our IRB protocol and informed consent document. However, data and analytic code will be made available to readers upon reasonable request by contacting Scott D. Blain or Ivy F. Tso. This study was not preregistered.

Scott D. Blain served as lead for conceptualization, formal analysis, visualization, writing–original draft, and writing–review and editing. Stephan F. Taylor served in a supporting role for funding acquisition, investigation, project administration, resources, supervision, and writing–review and editing. Saige E. Rutherford served in a supporting role for formal analysis, methodology, and writing–review and editing. Carly A. Lasagna served in a supporting role for conceptualization, data curation, methodology, visualization, writing–original draft, and writing–review and editing. Beier Yao served in a supporting role for data curation and writing–review and editing. Mike Angstadt served in a supporting role for data curation, formal analysis, methodology, project administration, resources, and software. Michael F. Green served in a supporting role for conceptualization, investigation, methodology, project administration, and writing–review and editing. Timothy D. Johnson served in a supporting role for conceptualization, formal analysis, investigation, methodology, and project administration. Scott Peltier contributed equally to resources and served in a supporting role for conceptualization, data curation, funding acquisition, investigation, methodology, project administration, software, and supervision. Vaibhav A. Diwadkar served in a supporting role for conceptualization, investigation, methodology, project administration, supervision, and writing–review and editing. Ivy F. Tso served as lead for funding acquisition, project administration, resources, software, and supervision, contributed equally to writing–review and editing, and served in a supporting role for data curation, formal analysis, visualization, and writing–original draft. Scott D. Blain and Saige E. Rutherford contributed equally to data curation. Scott D. Blain and Ivy F. Tso contributed equally to methodology and investigation. Stephan F. Taylor and Ivy F. Tso contributed equally to conceptualization.

Correspondence concerning this article should be addressed to Scott D. Blain, Department of Psychiatry and Behavioral Health, The Ohio State University, 1960 Kenny Rd., Columbus, OH 43210, United States. Email: scott.blain@osumc.edu

superior temporal sulcus, and insula. Activation was modulated by stimulus gaze angle and perception of stimuli as self-directed versus averted. More precise gaze perception and higher task-related activation were associated with better social cognition. Patients with SZ showed hyperactivation within left pre-/post-central gyrus, which was associated with more precise gaze perception and fewer symptoms and thus may be a compensatory mechanism. ***Conclusions:*** Neural and behavioral indices of gaze perception were related to social cognition, across patients and controls. This suggests gaze perception is an important perceptual building block for more complex social cognition. Results are discussed in the context of dimensional psychopathology and clinical heterogeneity.

General Scientific Summary
Our study of 77 schizophrenia (SZ) patients and 71 healthy controls combined behavioral tasks, functional neuroimaging, and statistical modeling to examine how eye gaze perception—a low-level visual process—functions as a basic building block of social cognition. We found that both brain and behavioral indices of gaze perception were associated with participants' performance across a range of social-cognition tasks (e.g., emotion recognition and emotion regulation). This suggests that gaze processing may be an important perceptual building block for more complex social cognition and related real-world functioning, which could serve as a potential target for future interventions in those with SZ.

Keywords: social cognition, face processing, gaze, fMRI, latent variables

Supplemental materials: https://doi.org/10.1037/abn0000846.supp

Impaired social cognition is a core feature of schizophrenia (SZ) and a critical determinant of functional outcomes (Fett et al., 2011). Limited response to pharmacological treatments (Harvey et al., 2006) has inspired considerable effort to better understand the mechanisms of social cognitive deficits in SZ, which could eventually inform the development of better biobehavioral interventions. One way to better understand social cognition (and its disruption in mental illness) is to elucidate the basic building blocks that make social cognition possible. For instance, basic visual processing deficits are common in SZ and contribute to problems with social cognition and real-world functioning (Silverstein & Keane, 2011; Tso et al., 2014). One possible bridge between low-level visual processing and social cognition is eye gaze perception, the focus of our current study.

Gaze is a ubiquitous social cue that conveys important information about one's attention, emotion, and mental state (Emery, 2000). Gaze perception is developmentally foundational (Farroni et al., 2002) and is included as a core social process in the National Institute of Mental Health (NIMH) Research Domain Criteria (RDoC) Matrix. Understanding aberrant gaze processing in SZ—which is characterized by reduced perceptual precision, increased self-referential bias, and impaired spatial coding—may shed light on mechanisms underlying symptoms such as paranoia and delusions of reference (Abbott et al., 2018; C. Hooker & Park, 2005; Röder et al., 2015; Tso et al., 2012, 2014; Tso, Taylor, et al., 2021). The present study aimed to elucidate neural underpinnings of gaze perception, their disruption in SZ, and their associations with social cognition. We address these questions using neuroimaging, psychophysics, and latent variable analysis.

The neural substrates of gaze perception are distributed throughout regions associated with social cognition and broader sensory processing; these include the anterior insula, anterior cingulate cortex (ACC), dorsomedial prefrontal cortex (dmPFC), temporoparietal junction (TPJ), inferior parietal lobule (IPL), superior temporal sulcus/gyrus (STS/STG), angular gyrus, posterior cingulate cortex (PCC), and fusiform gyrus (Bristow et al., 2007; Carlin et al., 2011; Cavallo et al., 2015; Grosbras et al., 2005; C. I. Hooker et al., 2003; Itier & Batty, 2009; Sato et al., 2016). Many of these regions show aberrant function in SZ (Green et al., 2015), with specific evidence of frontal, occipital, and limbic hypoactivation as well as abnormal connectivity during gaze processing (Kohler et al., 2008; Pinkham et al., 2011; Tso, Angstadt, et al., 2021). Despite the prevalence of case–control studies in SZ research, there are also substantial *within-group* individual differences in social cognition. Given that many psychiatric symptoms and mechanisms cut across disorders and also relate to variation in normal-range traits (Insel et al., 2010; Kotov et al., 2017), it is important to understand not only how social cognition and its associated mechanisms differ *between* SZ and healthy controls (HC), but also to understand these associations *within and across* SZ and HC groups and whether any such associations differ by group.

Several studies have documented robust correlations among gaze processing, social cognition, and associated neural systems. For instance, general population studies have linked gaze perception to social cognitive ability and social functioning (Lasagna et al., 2020; McCrackin & Itier, 2021; Wastler & Lenzenweger, 2018). Gaze perception abilities and related brain activity have also been linked to broader social cognition and functional outcomes in SZ (Pinkham et al., 2011; Tso et al., 2012; Tso, Taylor, et al., 2021). Finally, research spanning HC and SZ has linked individual differences in social functioning to default network connectivity (Meda et al., 2014) and effective connectivity during gaze processing (Tso, Angstadt, et al., 2021).

Many studies only test *associations among observed variables* (e.g., accuracy on a single social-cognition task or brain activity within specific regions of interest), but single-task, performance-based indicators often show limited reliability (Apperly, 2012; Enkavi et al., 2019). Moreover, performance on any given task is influenced by task-specific factors that may or may not relate to the underlying construct of interest. The ability to reliably measure constructs and estimate their interrelations is improved by using latent variable analyses (Schumacker & Lomax, 2004). For example, a social cognition latent variable can be modeled as the shared variance of scores across measures of emotion recognition, mentalizing, and emotion regulation. Latent variables allow us to eliminate unsystematic error variance, facilitating more accurate estimates of variance and covariance for our constructs of interest (Blain, Grazioplene, et al., 2020; Blain, Longenecker, et al., 2020; Eisenberg et al., 2019; Enkavi et al., 2019). Latent variable analysis can also incorporate functional magnetic resonance imaging (fMRI) data, which can be used to examine how behavioral constructs like social cognition are associated with neural factors representing shared variance of brain activity across multiple regions recruited in various processes (Cooper et al., 2019; Kim et al., 2007; Lahey et al., 2012).

The current study investigated: (a) neural correlates of gaze perception in SZ and HC, (b) associations between gaze perception and social cognition, (c) associations between patterns of brain activity during gaze perception and social cognition, and (d) whether associations among gaze perception, social cognition, and associated brain activity differ between diagnostic groups. Participants completed a comprehensive battery of social-cognition tasks, spanning emotion recognition, perceptual theory of mind, and emotion regulation. Additionally, they completed a gaze perception task during fMRI. In this task, the gaze angles of face stimuli were sampled from a continuum ranging from self-directed (looking at the participant) to averted (looking away from the participant). Building on evidence of differential brain responses to direct versus averted gaze (Berchio et al., 2016; Boyarskaya et al., 2015; Kesner et al., 2018; Marquardt et al., 2017), we were motivated to examine how brain activity might be modulated by the degree to which each stimulus is directed toward the observer. Thus, we examined patterns of neural activation associated with gaze perception using four fMRI contrasts (i.e., activation during the task compared to baseline, during gaze perception compared to gender identification, and as a function of whether gaze is objectively/subjectively self-directed or averted). We then used latent variable analysis to examine how gaze perception performance and associated patterns of brain activity (i.e., exploratory latent neural factors representing shared variance in brain activation clusters across our four fMRI contrasts of interest) might be associated with diagnosis and social cognition. Follow-up analyses examined whether associations among gaze perception, social cognition, and neural factors differed between groups.

We hypothesized gaze perception would be associated with activity distributed throughout regions associated with visual processing and social cognition. Specifically, we anticipated preferential activation for gaze (vs. gender) in the insula, IPL/angular gyrus, TPJ, STS/STG, and dmPFC. We also hypothesized these regions would show attenuated activation in SZ. We hypothesized that better social cognition (i.e., shared variance in performance across relevant tasks) would be associated with gaze-perception performance (i.e., higher perceptual precision and lower self-referential bias). We also hypothesized latent neural factors—reflecting important aspects of neural response to gaze—would be associated with gaze-perception performance, social cognition, and diagnostic group. Finally, we explored whether associations among gaze perception, social cognition, and neural response to gaze differed between groups and whether these variables were associated with symptom severity within the SZ group.

Method

Participants

Seventy-seven individuals with schizophrenia or schizoaffective disorder (SZ) and 71 HC completed the study. Diagnoses were based on the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, text revision (DSM-IV-TR) and established using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-IV; First et al., 1995). Demographic characteristics of SZ and HC were well matched and are summarized in Table 1. In terms of race, the sample included 88 White (59.4%), 39 Black (26.3%), 14 Asian (9.5%), one Native American (0.7%), six participants identifying as multiracial or other (4.1%); and nine participants identified as Hispanic (6.1%). Of the 148 participants with behavioral data, 116 had valid fMRI data (62 SZ, 54 HC).

Participants were recruited through community advertisements and referrals from local clinics. Inclusion/exclusion criteria are documented in the online supplemental material. Clinical characteristics for the SZ group, including symptoms and chlorpromazine (CPZ) dose equivalents are reported in Table 1. The study was conducted in accordance with the protocol approved by the Institutional Review Board (IRB) of the University of Michigan Medical School (HUM00080457, "Neural Mechanisms of Gaze Perception in Psychosis"), and written informed consent was obtained from each participant.

A subset of participants (27 SZ, 22 HC) overlaps with those from previously published studies (Tso, Angstadt, et al., 2021; Tso, Burton, et al., 2021). These studies differed substantially from the current work, with one focusing on a region-of-interest analysis of patients with bipolar disorder (who are not included in the current study) compared to HC (Tso, Burton, et al., 2021) and the other focusing on effective connectivity during gaze perception (Tso, Angstadt, et al., 2021).

Assessments

Patients were assessed (Table 1) using either the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) or the Scale for the Assessment of Positive Symptoms and Scale for the Assessment of Negative Symptoms (SAPS/SANS; Andreasen & Grove, 1986). Social cognition was assessed using the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT; Mayer et al., 1999), Reading the Mind in the Eyes test (RME; Baron-Cohen et al., 2001), and Penn Emotion Recognition Task (ER-40; Kohler et al., 2003). The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) battery was used to assess various other aspects of cognitive ability (Marder & Fenton, 2004; Nuechterlein et al., 2008).

fMRI Data Acquisition and Processing

MRI scanning occurred on a 3.0T GE MR 750 Discovery scanner. fMRI data were processed using Statistical Parametric Mapping (Ashburner et al., 2021). See the online supplemental material for additional scanning procedures and preprocessing details. The online supplemental material also describes details of two subsamples ($n_1 = 49$, including 27 SZ and 22 HC; $n_2 = 67$, including 35 SZ and 32 HC) that made up the data set, as well as data harmonization procedure for the two studies from where the subsamples came.

Table 1
Demographic and Clinical Characteristics

Full sample	SZ ($n = 77$) $M \pm SD$	HC ($n = 71$) $M \pm SD$	t or χ^2	p
Age	33.2 ± 10.0	33.2 ± 11.1	0.0	.993
Sex (male/female)	41/36	35/36	0.2	.631
Education, years	14.4 ± 2.1	16.1 ± 1.9	−5.1	<.001
Parental education, years	15.5 ± 2.5	15.6 ± 3.2	−0.1	.903
MATRICS working memory	43.0	53.1	−6.1	<.001
MATRICS reasoning	45.3	51.2	−3.4	<.001
MATRICS composite	38.2	55.3	−8.9	<.001
PANSS positive	17.0 ± 5.7			
PANSS negative	14.6 ± 4.9			
SAPS	20.8 ± 16.9			
SANS	32.8 ± 17.5			
CPZeq	349.5 ± 351.8			
Diagnosis (schizophrenia/schizoaffective)	31/46			
fMRI sample	**SZ ($n = 62$) $M \pm SD$**	**HC ($n = 54$) $M \pm SD$**	**t or χ^2**	**p**
Age	32.9 ± 10.0	33.0 ± 11.1	0.0	.968
Sex (male/female)	30/32	27/27	0.0	.862
Education, years	14.4 ± 2.2	16.1 ± 2.0	−4.2	<.001
MATRICS working memory	42.2	51.2	−4.3	<.001
MATRICS reasoning	43.9	50.4	−3.4	<.001
MATRICS composite	38.8	52.4	−5.5	<.001
Parental education, years	15.5 ± 2.4	15.7 ± 3.4	−0.8	.788
PANSS positive	17.0 ± 5.6			
PANSS negative	14.6 ± 4.8			
SAPS	19.8 ± 17.3			
SANS	33.1 ± 17.3			
CPZeq	330.4 ± 344.9			
Diagnosis (schizophrenia/schizoaffective)	22/40			

Note. Parental education represents an average of maternal and paternal education. SZ = schizophrenia; HC = healthy controls; fMRI = functional magnetic resonance imaging; MATRICS = Measurement and Treatment Research to Improve Cognition in Schizophrenia; PANSS = Positive and Negative Syndrome Scale; SAPS = Scale for the Assessment of Positive Symptoms; SANS = Scale for the Assessment of Negative Symptoms; CPZeq = chlorpromazine equivalents.

fMRI Paradigm: Gaze Perception Task

Stimuli were naturalistic face images with nine different gaze angles presented in pseudorandomized order. These gaze angles encompass the psychophysical aspect of the task (Figure 1a), representing levels of eye-contact signal strength, ranging from 0.2 (averted), 0.3, … , to 1.0 (self-directed). For each face, participants indicated perceived eye contact or gender by pressing a button (Figure 1b). For the gaze condition, participants indicated whether or not each face was looking at them. For the gender condition, participants indicated whether each face was male or female. Participants in the first data set subsample completed 108 trials (6 trials × 6 blocks × 3 runs) for each task (eyes and gender). Participants in the second data set subsample completed 216 trials (6 trials × 6 blocks × 6 runs) for each task (eyes and gender). Stimuli were adopted from Gur et al. (2002).

Behavioral Data Analyses

Psychophysical Indices of Gaze Perception

For each participant, eye-contact endorsement rates across different gaze angles were modeled as a logistic function of gaze angle (i.e., objective eye-contact signal strength); two psychophysical properties (threshold and width) were derived from this curve (Figure 1c). Threshold represents how strong the eye contact signal needs to be for the individual to perceive eye contact 50% of the time. Lower thresholds indicate that weaker signals are needed to perceive gaze as directed toward oneself, thus representing stronger self-referential bias. Width represents the difference in signal strengths for which participants are endorsing 5% self-directed responses versus 95% self-directed responses. Narrower width (lower values) indicates higher perceptual precision, suggesting the participant is more sensitive to small changes in gaze angle.

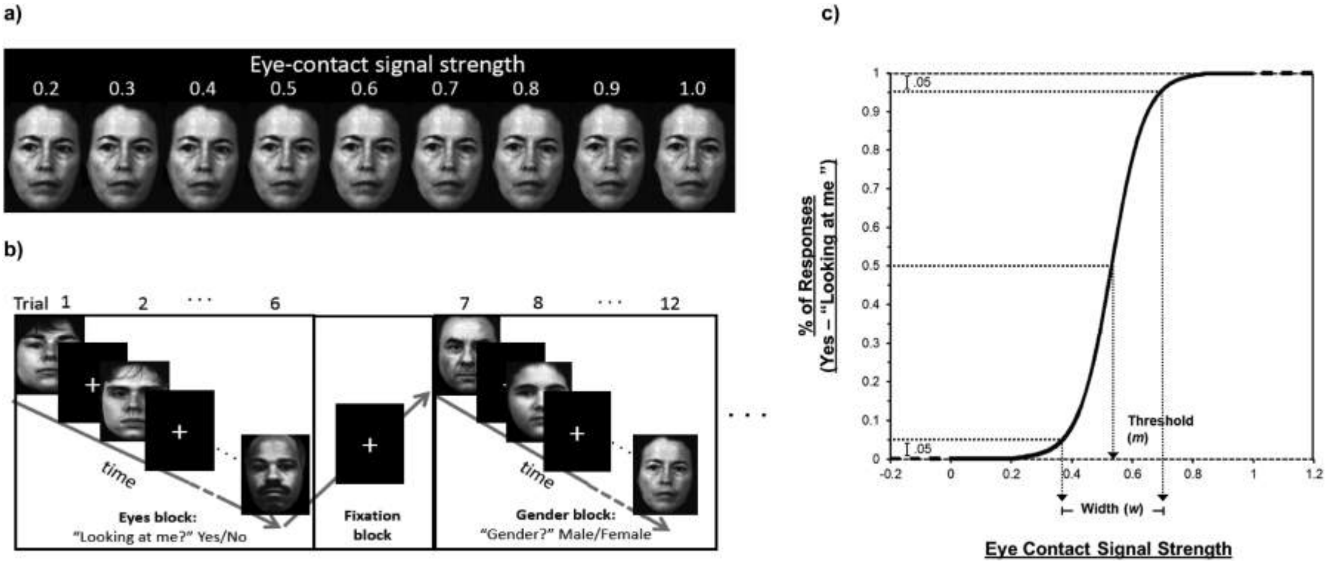
Group Differences in Social Cognition

Descriptive statistics were computed for social cognition measures and performance on the gaze task; to test whether SZ and HC differed, independent samples t tests were performed, and Cohen's d statistics were computed. Variable distributions, sorted by group, are displayed in Figure S1 in the online supplemental materials.

fMRI Data Analyses

fMRI data were modeled to identify patterns of brain activity associated with aspects of gaze perception. First, the anatomically

Figure 1
Face Stimuli and Design of the Gaze Perception Task



Note. (a) Face stimuli with nine gaze angles. (b) Eyes and gender trials were presented in alternating blocks with a fixation block between each task block. (c) After curve-fitting, two metrics were obtained from each participant's function: threshold (m) and width (w). Facial images were adapted with permission from Gur et al. (2002).

normalized time series was regressed on two regressors of interest (gaze trials and gender trials) and nuisance regressors (corresponding to motion parameters and runs), convolved with a canonical hemodynamic response function. An additional regressor was used to identify activation modulated by participants' individual perception (as indicated by corresponding behavioral endorsement) of each stimulus as self-directed (value of 1) or averted (value of 0), with averaged values within a given block used for trials with no response. In an alternative model, we used a parametric modulation regressor to identify activation that varies as a function of stimulus gaze angle (regressor values for each trial ranged from −4 to 4 in increments of 1, with larger values corresponding to more self-directed gaze).

For our second-level models, we examined which voxel clusters were (a) commonly activated across all participants and (b) differentially activated between SZ and HC. Models were computed for all-trials versus baseline, gaze versus gender, modulation by participants' perception of stimuli as self-directed, and modulation by stimulus gaze angle. Additional covariates were included corresponding to head motion (mean framewise displacement) and data set subsample. Second-level models were computed using a cluster-defining threshold of $p < .001$ and family-wise error rate correction of $p < .05$ (Eklund et al., 2016). After examining which clusters were significant at the whole-sample level and for the group-comparison contrasts (for task vs. baseline, gaze vs. gender, and two modulation contrasts), we extracted beta estimates (the first eigenvariate) within an 8 mm radius sphere from the peaks of significant clusters, which were used for subsequent analyses. Example variable distributions, sorted by group, are displayed in Figures S2 and S3 in the online supplemental materials.

Latent Variable Analysis

To test associations among diagnosis, gaze perception (i.e., width and threshold), and social cognition, we used Bayesian latent variable modeling, implemented with Mplus (L. K. Muthén & Muthén, 2017); Bayesian estimation of latent variable models provides optimal fit and is ideal for relatively small samples (B. Muthén & Asparouhov, 2012). A social cognition latent variable was indicated by MSCEIT branch scores, RME accuracy, and ER-40 accuracy; we also included an orthogonal methods factor for the MSCEIT. Our decision to use a single latent variable for social cognition was consistent with the results of a Velicer's minimum average partial (MAP) test, which achieved a minimum value with one factor ($MAP_1 = 0.07$). We examined correlations among social cognition, diagnosis ($HC = 0$, $SZ = 1$), and gaze-perception metrics.

To test whether patterns of brain activity engaged during gaze perception were associated with gaze-perception performance, social cognition, and diagnostic group, we used latent variable analysis with full-information robust weighted least squares means and variance adjusted (WLSMV) estimation. We fit exploratory neural factors (using an oblimin rotation) from our extracted beta estimates for significant clusters from whole-brain analyses, which were then used as statistical predictors of the social cognition latent variable, width and threshold metrics from the gaze task, and diagnostic group ($HC = 0$, $SZ = 1$). The number of neural factors was selected based on the results of a Velicer's MAP test (O'Connor, 2000).

Following our primary latent variable analyses in Mplus, factor scores for latent variables were computed and exported for follow-up moderation analyses and visualizations. To examine whether associations among gaze perception, social cognition, and neural factors differed between groups, we conducted follow-up regression models that included diagnosis and neural-factor-by-diagnosis interaction terms. We visualized key findings using scatter plots.

Finally, sensitivity analyses were conducted. First, we examined whether results changed when sex, age, cognitive ability (i.e., participants' average of working memory and reasoning MATRICS battery *t* scores), and data-collection subsample were included as covariates and when task-performance outliers were removed from the data set. Outlier datapoints were identified using Rosner's generalized extreme Studentized deviate (ESD) test (Rosner, 1983) and subsequently removed to form a new data set for these sensitivity analyses. At the suggestion of reviewers, we estimated additional models including only ER-40, RME, and the perceiving emotions branch score from the MSCEIT as indicators of social cognition.

Associations With Clinical Symptoms

Lastly, we examined associations of symptom severity with our gaze perception metrics, social cognition, and neural factors. Given that a portion of our participants completed the SAPS/SANS while others completed the PANSS, we converted scores on the SAPS and SANS to equivalent scores on the PANSS using the procedure outlined by Grot et al. (2021). Then, correlations were examined for positive, negative, and total symptoms. We examined zero-order correlations of symptoms with gaze perception, social cognition, and brain activation. Finally, we conducted additional analyses controlling for multicollinearity among brain-activation metrics.

Transparency and Openness

Legally, our data cannot be made publicly available online, due to the wording of our IRB protocol and informed consent document. Study materials, data, and analytic code will be made available to readers upon reasonable request, by contacting Scott D. Blain or Ivy F. Tso. Data were analyzed using R (R Core Team, 2020) and Mplus (L. K. Muthén & Muthén, 2017). The study was not preregistered.

Results

Behavioral Data

In the gaze task, SZ showed higher width and lower threshold, suggesting lower perceptual precision and higher self-referential bias. SZ also showed worse performance on *all six of six* observed social cognition variables. Means, standard deviations, and group differences are reported in Table S1 in the online supplemental materials, and variable distributions are presented in Figure S1 in the online supplemental materials. No behavioral or neural variables of interest were significantly correlated with SZ participants' CPZ dose equivalents and no variables significantly differed between those with versus without valid fMRI data.

fMRI Data

Second-level analyses across all participants revealed several significant clusters for each contrast (all-trials vs. baseline, gaze vs. gender, modulation by perception, and modulation by gaze angle). Clusters were distributed throughout regions canonically associated with social cognition, salience attribution, visual processing, and sensorimotor function (Figures S4–S9, Table S2 in the online supplemental materials). For the gaze versus gender and modulation contrasts, no clusters significantly differed between HC and SZ; for the task-versus-baseline contrast, SZ showed greater activation in a cluster centered on left pre-/postcentral gyrus (Figure 2, Figures S10–S15, Table S3 in the online supplemental materials). Moreover, in SZ, task-versus-baseline activation within this cluster was *correlated with better perceptual precision* (indicated by a negative correlation with width) on the gaze task ($r = -.269$, $p = .040$), but not self-referential bias ($r = .069$, $p = .604$) or latent social cognition ($r = -.051$, $p = .696$). Activation in the cluster was not associated with these variables in HC ($ps > .05$). These results are visualized in Figure 2 and were consistent with follow-up moderation analysis that showed a significant interaction between group and pre-/postcentral gyrus activation in predicting perceptual precision ($\beta = -.202$, $p = .008$).

Latent Variable Analysis

Model for Gaze Perception, Social Cognition, and Diagnosis

Our behavioral latent variable model showed excellent fit—root-mean-square error of approximation (RMSEA) = 0.015, 95% CI: [0.000, 0.081], comparative fit index (CFI) = 0.997, [0.936, 1.000], and Tucker–Lewis index (TLI) = 0.995, [0.888, 1.000]. Results are visualized in Figure 3. In our measurement model of social cognition, all observed variables were significantly loaded onto the latent factor. Better social cognition was associated with higher perceptual precision (i.e., a lower width parameter) but not with self-referential bias (i.e., threshold) on the gaze task; the strength of associations did not significantly differ between groups. Associations are visualized in Figure 4, using scatterplots of gaze-perception metrics and estimated social cognition factor scores. As in our *t* tests reported above, SZ showed lower perceptual precision, higher self-referential bias, and worse social cognition.

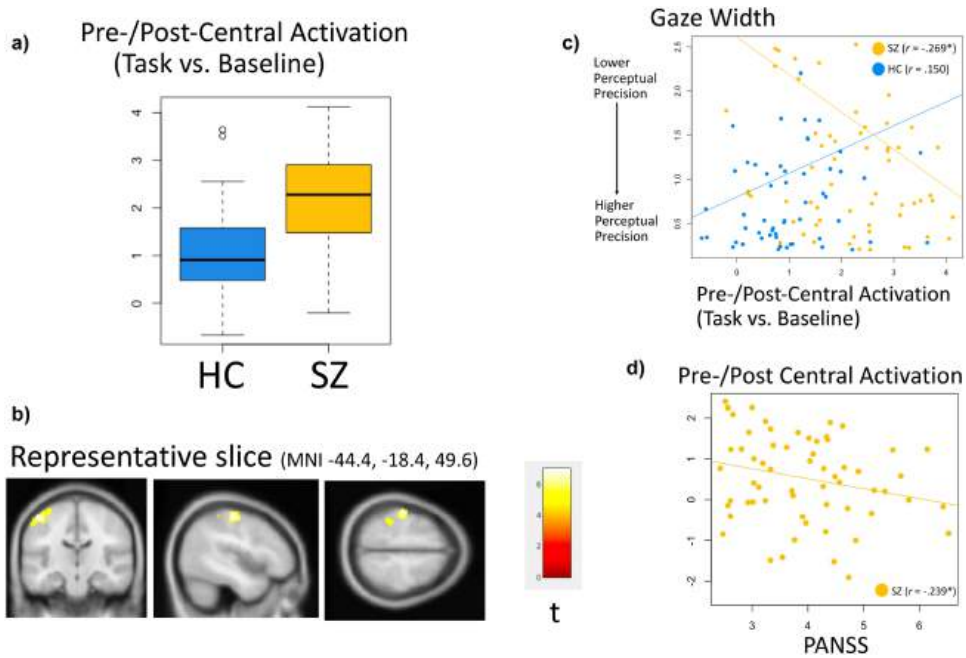
In general, associations among gaze perception metrics, social cognition, and diagnosis were similar if we controlled for sex, age, cognitive ability, and data set subsample or if task-performance outliers were removed from the data set (for details, see Figures S16, S17 and Table S6 in the online supplemental materials). A final analysis modeled social cognition using only the ER-40, RME, and the perceiving emotions branch from the MSCEIT (Figure S18 in the online supplemental materials); notably, associations between the gaze perception metrics and this reduced social cognition latent variable were stronger than those in our primary model ($r = .421$ for threshold and $r = -.657$ for width). This implies that the association between social cognition and gaze perception performance was driven largely by perceptual-based tasks.

Measurement Models for Neural Latent Variables

Our neurobehavioral latent variable model showed good fit—RMSEA = 0.022, 95% CI: [0.012, 0.030], CFI = 0.930, and TLI = 0.918. In terms of neural variables, six exploratory factors were extracted, based on the results of Velicer's MAP test ($MAP_6 = 0.017$). See Table S4 in the online supplemental materials for factors loadings. Neural factors corresponded to (a) global task activation versus baseline, (b) preferential activation for gaze (vs. gender), (c) preferential activation for gender (vs. gaze), (d) modulation (increase) of activation by perception of stimuli as self-directed and by viewing stimuli with objectively more self-directed gaze angles, (e) modulation (increase) of activation by viewing objectively more averted gaze angles, and (f) modulation (increase) of activation by perception of stimuli as averted. Note that voxel clusters that were more active in response to self-

Figure 2

Sensorimotor Hyperactivation and Perceptual Precision in SZ



Note. SZ participants showed hyperactivation (a) within left pre-/postcentral gyrus (b), which was associated with greater perceptual precision on the gaze task (c) and fewer symptoms (d). Findings are consistent with a moderation analysis that showed a significant interaction between group and pre-/postcentral activation predicting perceptual precision ($\beta = -.202$, $p = .008$). SZ = schizophrenia; HC = healthy controls; PANSS = Positive and Negative Syndrome Scale. See the online article for the color version of this figure.
* $p < .05$.

directed gaze—including those from the fMRI contrast for participant perception/behavioral endorsement and the contrast for objective gaze angle—loaded onto Factor 4, whereas voxel clusters that were more active in response to averted gaze were split across Factors 5 (for the objective gaze angle contrast) and 6 (for the participant perception/behavioral endorsement contrast). To facilitate interpretation, we choose to reverse signs when reporting associations of Factors 3, 5, and 6 with diagnosis, social cognition, and gaze perception, given that these factors were indicated by variables with negative mean values for the whole-sample, second-level fMRI contrasts. (Thus, more-negative values, before signs were reversed, would represent stronger activation in the group-mean direction.)

Associations of Neural Factors, Social Cognition, Gaze Perception, and Diagnostic Group

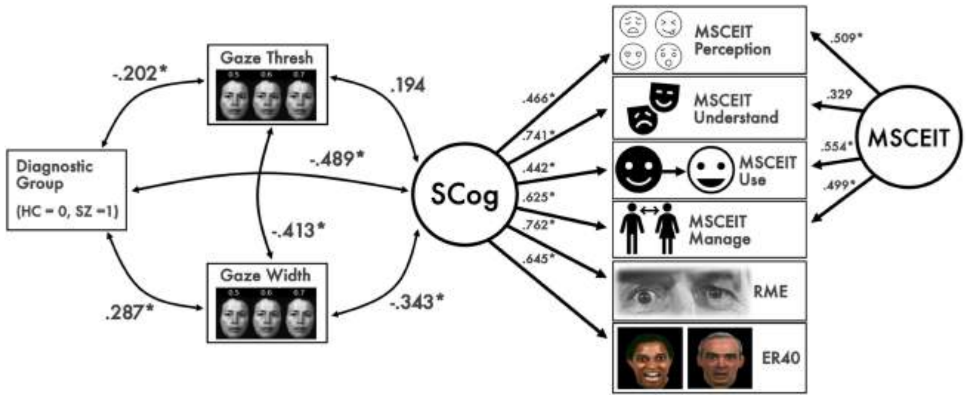
Path coefficients from neural variables to gaze-perception metrics, social cognition, and diagnosis are shown in Figure 5. Three out of six neural factors were significantly associated with social cognition. Specifically, better social cognition was associated with stronger activation across neural factors marking global task activation versus baseline (Factor 1), preferential activation for gaze versus gender (Factor 2), and preferential activation for gender versus gaze (Factor 3). Brain regions most strongly implicated in these neural factors included the following: fusiform gyrus, broad areas throughout PFC, hippocampus, and pre-/postcentral gyri (Factor 1), insula, dmPFC, and IPL (Factor 2), and dlPFC, PCC/precuneus, angular gyrus, and STS/STG (Factor 3).

In terms of neural factors and diagnostic group, only global task activation versus baseline (Factor 1) showed a significant association; importantly, this association was negative (the opposite direction compared to its association with social cognition), suggesting patients showed broadly attenuated activation during both task conditions (i.e., gaze and gender) versus baseline.

Reduced perceptual precision (i.e., a higher width parameter) was more specifically associated with reduced modulation of brain activity by viewing more direct versus more averted gaze; this included factors spanning modulation of activation by both objective and subjective gaze self-directedness (Factors 4–6). Brain regions most strongly implicated in these neural factors included the following: insula, dmPFC, intraparietal sulcus, and orbitofrontal cortex (Factor 4); PCC/precuneus, pre-/postcentral gyrus, STS/STG, and secondary visual cortex (Factor 5); fusiform gyrus, and precuneus (Factor 6). These findings suggest the psychophysical width parameter—an index of how strongly participants' perceptual decision-making relates to changes in stimulus properties (i.e., gaze angle)—may track particularly well with modulation of neural activation by these same stimulus properties.

Patterns of association between neural factors and self-referential bias on the gaze task mirrored several of those seen for social cognition and perceptual precision. Specifically, stronger self-referential

Figure 3
Behavioral Model of Diagnosis, Gaze Perception, and Latent Social Cognition



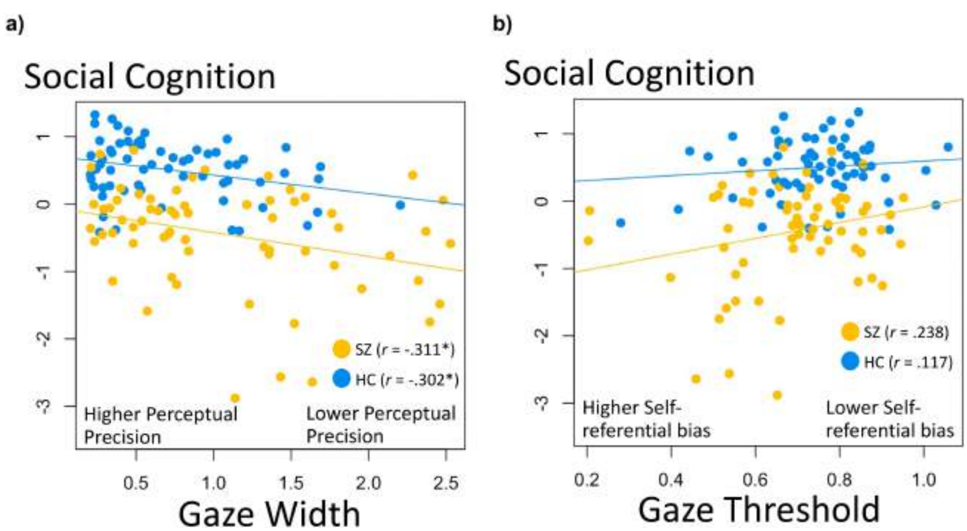
Note. The model depicts associations among diagnostic group, gaze perception metrics, and social cognition. Factor loadings for a social cognition latent factor (SCog) are shown on the right, as well as a method factor for the MSCEIT branch scores. Better social cognition was associated with greater perceptual precision. SZ showed worse social cognition, lower perceptual precision, and greater self-referential bias. MSCEIT = Mayer–Salovey–Caruso Emotional Intelligence Test; RME = Reading the Mind in the Eyes Test; ER-40 = Penn Emotion Recognition Test; SZ = schizophrenia; HC = healthy controls. Facial images were adapted with permission from Gur et al. (2002). In addition, the image of eyes featured in the RME box was adapted with permission from the Autism Research Center (Baron-Cohen et al., 2001). See the online article for the color version of this figure.
* $p < .05$.

bias (i.e., a lower threshold parameter) was associated with less preferential activation for gender versus gaze (Factor 3), as well as reduced modulation by viewing more direct versus more averted gaze stimuli (Factors 4 and 5).

Follow-Up Analyses

Associations were similar across the two diagnostic groups, as indicated by mostly nonsignificant interaction terms in follow-up

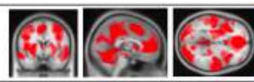
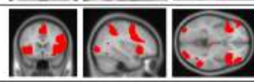
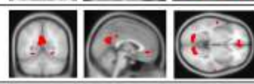
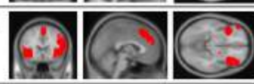
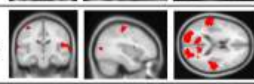
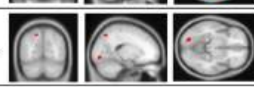
Figure 4
Scatter Plots of Gaze Perception Metrics and Social Cognition



Note. Across both diagnostic groups, a lower width parameter on the gaze task (i.e., greater perceptual precision) was associated with better social cognition ($r = -.343$, $p = .004$) (a). The threshold parameter (i.e., lower self-referential bias), however, was not significantly associated with social cognition ($r = .194$, $p = .084$) (b). Diagnosis did not significantly moderate the associations of social cognition with width ($\beta = -.030$, $p = .678$) or threshold ($\beta = .076$, $p = .278$). SZ = schizophrenia; HC = healthy controls. See the online article for the color version of this figure.
* $p < .05$.

Figure 5

Whole-Sample Regression Coefficients Among Key Study Variables

		Group (SZ = 1)	Social Cognition	Gaze Threshold	Gaze Width
Task Activation		-.211*	.201*	.123	-.036
Gaze > Gender		-.035	.219*	.010	-.011
Gender > Gaze		-.058	.231*	.196*	-.049
Self-directed Gaze		-.044	.081	.320**	-.369**
Objectively Averted Gaze		-.186	-.038	.224**	-.357**
Perceived Averted Gaze		.107	-.107	.063	-.204*

Note. The figure displays standardized path coefficients from each of the neural factors to criterion variables for diagnostic group, latent social cognition, gaze threshold (i.e., self-referential bias), and gaze width (i.e., perceptual precision). Neural factors correspond to (1) global task activation versus baseline, (2) preferential activation for gaze (vs. gender), (3) preferential activation for gender (vs. gaze), (4) modulation (increase) of activation by perception of stimuli as self-directed and by viewing stimuli with objectively more self-directed gaze angles, (5) modulation (increase) of activation by viewing objectively more averted gaze angles, and (6) modulation (increase) of activation by perception of stimuli as averted. Voxel clusters that were more active in response to self-directed gaze—including those from the fMRI contrast for participant perception/behavioral endorsement and the contrast for objective gaze angle—loaded onto Factor 4, whereas voxel clusters that were more active in response to averted gaze were split across Factors 5 (for the objective gaze angle contrast) and 6 (for the participant perception/behavioral endorsement contrast). Significant positive associations are colored in red and negative correlations in blue. SZ = schizophrenia; fMRI = functional magnetic resonance imaging. See the online article for the color version of this figure.
* $p < .05$. ** $p < .01$.

regression analyses (Table S5 in the online supplemental materials). Significant interactions were, however, found between group and the neural factor marking task activation versus baseline, when statistically predicting social cognition ($\beta = .321$, $p < .001$) and self-referential bias on the gaze task ($\beta = .237$, $p = .034$); interactions were positive, suggesting that the association between greater task activation and better social cognition (as well as lower self-referential bias) was particularly strong in SZ. Inspection of scatterplots and within-group correlations (Figure 6) suggested the association between social cognition and global task activation was driven by the SZ group and was not significant when examined in the HC group alone.

In general, associations among gaze perception metrics, social cognition, neural factors, and diagnosis were similar if we controlled for sex, age, cognitive ability, and data set subsample or if task-performance outliers were removed from the data set. A final analysis modeled social cognition using only the ER-40, RME, and the perceiving emotions branch from the MSCEIT; again, results were mostly similar to those from the primary analyses. For details, see Figures S19–S21 and Tables S6–S8 in the online supplemental materials.

Associations With Clinical Symptoms

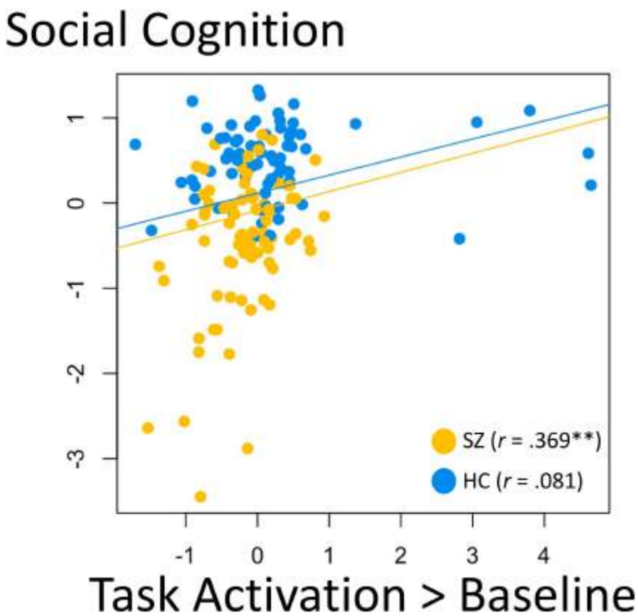
Several significant associations were found between PANSS scores and our variables of interest (Table 2). More severe symptoms were associated with worse social cognition and perceptual precision. More severe symptoms were also associated with decreased preferential activation for gender versus gaze and less modulation of activation by averted gaze. Finally, greater activation in the pre-/postcentral gyrus cluster that differed between HC and SZ was associated with *lower* symptom severity, mirroring the association of better perceptual precision with activation in this region. Across task performance and neural variables, stronger associations were seen for negative and total symptoms compared to associations with positive symptoms.

Discussion

In this study, we combined fMRI, psychophysics, social-cognition tasks, and latent variable analysis to elucidate whether gaze perception and associated brain activity are correlates of social cognition. We found that both behavioral and neural indicators of gaze perception were associated with individual differences in social cognition, across SZ and HC, and that these neurobehavioral indicators were particularly related to negative symptom severity. We also found broadly attenuated task-related activation—but enhanced sensorimotor activation—in SZ. Strength of this sensorimotor hyperactivation was associated with better perceptual precision on the gaze task and decreased symptom severity.

Figure 6

Plot of the Task Activation Neural Factor, Diagnostic Group, and Social Cognition



Note. Diagnostic group was a significant moderator of the association between social cognition and the global task activation neural factor ($\beta = .254$, $p < .001$); the positive association observed in the whole sample was largely driven by an association within the SZ group. SZ = schizophrenia. See the online article for the color version of this figure.
$^{**}p < .01$.

Table 2

Associations of Symptoms With Gaze Perception, Social Cognition, and Neural Activation

Behavioral or neural variable	PANSS negative	PANSS positive	PANSS total
Task performance			
Gaze width	0.433**	0.269*	0.506**
Gaze threshold	−0.207†	−0.124	−0.239†
Social cognition	−0.236*	−0.190†	−0.306**
Neural activation (zero-order)			
Pre-/postcentral activation	−0.173	−0.134	−0.224†
Task activation	−0.123	−0.071	−0.140
Gaze > gender	0.029	−0.063	−0.021
Gender > gaze	−0.218†	−0.117	−0.242*
Direct gaze	−0.148	−0.062	−0.153
Objectively averted gaze	−0.213†	−0.141	−0.255*
Perceived averted gaze	−0.330**	−0.040	−0.272*
Neural activation (partial)			
Pre-/postcentral activation	−0.191	−0.136	−0.239*
Task activation	−0.018	−0.069	−0.062
Gaze > gender	−0.096	−0.184	−0.199
Gender > gaze	−0.244†	−0.201	−0.323*
Direct gaze	−0.166	−0.036	−0.151
Objectively averted gaze	−0.044	−0.105	−0.106
Perceived averted gaze	−0.335**	−0.021	−0.267*

Note. PANSS = Positive and Negative Syndrome Scale.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

Reinforcing and Extending Previous Social Cognition Research

Behavioral results from the current study reinforce previous research on social cognitive deficits in SZ. We replicate findings of reduced perceptual precision and increased self-referential bias during gaze perception in SZ (Tso et al., 2012; Tso, Taylor, et al., 2021; Yao et al., 2018), as well as poorer social cognition as measured with the MSCEIT, RME, and ER-40 (Pinkham et al., 2018; Tso et al., 2014). The significant association between latent social cognition and diagnostic group suggests that fairly broad—rather than task-specific—social cognitive deficits are characteristic of SZ.

The fact that social cognition and perceptual precision were associated not only with diagnostic group but also with symptom severity is relevant to explaining heterogeneity within SZ. Social cognition and gaze perception were most strongly related to negative symptoms, suggesting features such as anhedonia, avolition, and withdrawal may contribute to social cognitive dysfunction and vice versa. This is consistent with several other studies that suggest negative symptoms are most strongly associated with social cognition and related functional outcomes (Abplanalp et al., 2022; Burton et al., 2019; Strassnig et al., 2015; Ventura et al., 2009).

Our fMRI analyses suggest that gaze perception—as well as its modulation by stimulus properties and perception—involves multiple, widely distributed brain regions (e.g., insula, dmPFC, IPL/angular gyrus, PCC, TPJ/STS, precentral/postcentral gyri, and visual cortex); many of these regions have been previously discussed as having roles in the visual, salience, and default mode networks (Uddin et al., 2019). Parametric modulation of brain activation during the gaze task (throughout clusters in insula, dmPFC, STS/TPJ, and IPL) was correlated with a psychophysical index of perceptual precision for task performance, suggesting a close mapping between behavioral and neural response to salient features of gaze. This is consistent with the known roles of these brain regions in mechanisms critical to processing self-related visual information: for instance, the anterior insula for salience attribution, IPL for visuospatial processing, STS/TPJ for perception of social cues, and dmPFC for self-referential processing (Boyarskaya et al., 2015; Itier & Batty, 2009; Schilbach et al., 2006; Schobert et al., 2018; Uddin, 2015).

Because the processes involved in gaze perception are also fundamental to broader social cognition, it is fitting that we also found behavioral and neural indices of gaze perception were associated with participants' performance across social-cognition tasks. These results add to a growing body of work linking social cognitive abilities to individual differences in structure and function of the dmPFC, insula, IPL, and TPJ/STS (Allen et al., 2017; Eres et al., 2015; Hou et al., 2017; Meda et al., 2014; Tso, Angstadt, et al., 2021; Udochi et al., 2022), extending past research through the use of multiple behavioral tasks, latent variable modeling, and a relatively large sample spanning patients and controls. The current findings are particularly novel and of note, as we show that gaze perception—a basic perceptual building block for more complex social cognition—is related to various social cognitive abilities, spanning emotion perception, perceptual theory of mind, and emotion regulation.

Diminished Global but Enhanced Sensorimotor Activation in SZ

In our latent variable analysis, a global task activation neural factor was lower in SZ and associated with worse social cognition within the SZ group. This could be interpreted to suggest that broad patterns of task-related hypoactivation—rather than specific differences related to the more nuanced aspects of gaze perception indexed by our other contrasts—might underly many of the social cognitive deficits seen in SZ. This interpretation is consistent with work showing hypoactivation across the precuneus, insula, and prefrontal cortex in SZ for a variety of general and social cognitive tasks (Green et al., 2015; Kohler et al., 2008; Soldevila-Matías et al., 2022). In contrast, more specific patterns of neural activation that did not show group differences or interactions in the current study —such as those underlying the processing of gaze versus gender and direct versus averted gaze—may underly individual differences in social cognition that are not specific to SZ.

Contrary to the diminished global activation in SZ, across both task conditions versus baseline, the SZ group showed enhanced activation within a cluster centered on left precentral/postcentral gyri—spanning motor, premotor, and somatosensory cortices. Counterintuitively, this "abnormal" finding was also correlated with better gaze perception—*higher perceptual precision*—suggesting that it may be a compensatory mechanism. This is consistent with our finding that activation in this cluster was also associated with *lower levels of symptom severity*. In other words, some individuals with SZ may successfully engage in sensorimotor-driven processes to make up for broader patterns of diminished neural activation (such as those captured in our global task activation factor) that may negatively impact gaze perception.

Sensorimotor hyperactivation could be helpful to gaze processing in several ways. For instance, disambiguating gaze direction may be challenging to SZ participants due to dysfunction of regions such as the pMFC, insula, and TPJ/STS, and increased sensorimotor engagement may increase processing of stimuli and control of motor output, helping prevent patients from misinterpreting averted gaze as self-directed. Another possibility is that viewing various levels of averted gaze automatically triggers attentional shifts, with or without eye movements, toward the direction of gaze (Frischen et al., 2007); this is consistent with a meta-analytic finding that gaze perception shows more similarities in functional neuroanatomy with reflexive than with voluntary shifts of attention (Grosbras et al., 2005). Since shifting eyes or attention to the gazed-at location of the viewed face is inappropriate in the context of the gaze task, it requires inhibition or prompt disengagement from the gazed-at location to do well on the task. Prior studies have shown intact automatic attentional orienting to gaze (Langdon et al., 2017; Schwartz et al., 2010) but difficulties disengaging from a gazed-at location once shared attention is established in SZ (Langdon et al., 2017). Therefore, over-recruitment of the sensorimotor network may help overcome this difficulty and redirect attention to the task. Future studies could test whether sensorimotor over-recruitment is truly a compensatory neural mechanism in SZ by experimentally manipulating activation of this region in SZ (e.g., using brain stimulation techniques) and assessing the impact on social cognition.

Interpretations and Explanations of Null Effects

Contrary to our hypotheses, there were no significant diagnostic group differences for brain activation during the gaze versus gender conditions or for modulation of activation by viewing direct versus averted gaze. These null findings at least partially contradict previous work showing group differences in neural activity during gaze perception (Kohler et al., 2008; Pinkham et al., 2011) and a larger body of work on the neural bases of social cognitive deficits in SZ (Green et al., 2015). Nonetheless, the current study utilized a comparatively large sample and is consistent with several other studies that did not detect significant group differences in regional brain activations, despite significant associations between brain function and dimensional measures of social functioning (Abram et al., 2017; Fox et al., 2017; Horan et al., 2014; Oliver et al., 2021).

It is possible our task design contributed to the lack of group differences. Although our Gaze–Gender contrast reveals brain activation specific to explicit gaze perception, it is by no means a complete account of all processes involved in gaze perception. This is because gender identification still entails face processing and implicit gaze perception. Thus, some signals that are inherently part of gaze perception were likely subtracted away in the gaze–gender results. For example, basic visual processing deficits are well documented in SZ (Silverstein & Keane, 2011; Tso et al., 2014) and may affect face processing during both explicit (i.e., distinguishing self-directed from averted gaze) and implicit (i.e., gender identification) gaze perception; likewise, both conditions may involve processes broadly related to social cognition and face perception. Unfortunately, these possibilities cannot be revealed with the current task design and will need to be explored in future investigations, which could use a control condition that does not allow normal face processing (e.g., scrambled faces or faces with the eyes region covered).

Within-Group Heterogeneity and Related Future Directions

Whereas task design may be one factor contributing to our lack of group differences, effects may also be masked by within-group variability in social cognition. Although a majority of SZ patients show social deficits, the magnitude of such deficits varies widely among patients and may be related to symptom severity (Hajdúk et al., 2018). Indeed, in the current study, we identified neural factors that were significantly related to symptom severity within the SZ group, despite not showing significant SZ versus HC differences. Future studies could more carefully probe the role of various symptom dimensions in social cognitive deficits and associated neural pathways.

In addition to there being considerable within-group variability in social cognition itself, even similar overt deficits may emerge from heterogenous neurobiological mechanisms and symptom profiles (i.e., equifinality; Cicchetti & Rogosch, 1996). For instance, social cognitive deficits in SZ could result from any combination of affective flattening, hypermentalizing, or cognitive disorganization, each of which may stem from unique symptom profiles and neurobiology (Bliksted et al., 2016; Frith, 2004; Green et al., 2015; Madeira et al., 2016). Better understanding of these alternative pathways to overt social deficits likely requires more advanced computational models of social cognition to parse the constituent cognitive and neural components. Future research should also incorporate a broader range of social cognition tasks, given that our current indices of social cognition were limited to the MSCEIT, RME, and ER-40, all of which focus on emotional processing. Studies that assess higher-level social cognitive abilities such as empathy, mental state attribution, mentalizing, or sarcasm perception (Abell et al., 2000; Buck et al., 2017; Corcoran et al., 1995; Johannesen et al., 2018; Stiller & Dunbar, 2007), examine the factor structure of social cognition, and examine associations of gaze processing with multiple social cognition factors would allow for even stronger conclusions. For instance, given that we found stronger associations with gaze perception when our social cognition latent variable included only lower-level/perceptual indicators (see Figure S18 in the online supplemental materials vs. Figure 3), it is plausible that associations with gaze perception would further differ in magnitude when using tasks that measure higher-level theory of mind or empathy.

Future work should also strive to elucidate finer-grained links among SZ symptomatology, neurocognitive mechanisms, and social functioning. These questions could be further probed by including participants at elevated risk for SZ, such as first-degree relatives and those with high schizotypy; these groups also tend to show abnormal social cognition (Stuke et al., 2021; Wastler & Lenzenweger, 2018), and research in nonpatient samples can help overcome limitations related to medication use in SZ. Meanwhile, it could also be useful to incorporate additional diagnoses and symptom dimensions in future work, as social cognitive deficits are also prominent in autism (Pantelis & Kennedy, 2017), social anxiety (Jun et al., 2013; Schulze et al., 2013), and personality disorders (Roepke et al., 2013; Winter et al., 2017). Such an approach would be consistent with frameworks such as the NIMH RDoC (Insel et al., 2010) and the Hierarchical Taxonomy of Psychopathology (Kotov et al., 2017), which seek to understand psychopathology in terms of underlying mechanisms and dimensions rather than diagnosis. Finally, although our sample size was comparable to or larger than those used in several previous studies using latent variable modeling/multivariate analyses to investigate brain function in SZ (Lincoln et al., 2020; Liu et al., 2021; Plis et al., 2014), future research with even larger samples should be undertaken to confirm results from the current study.

Conclusion

Findings add to a growing body of work characterizing the neural substrates of gaze perception and their relation to social cognition. Given that both neural and behavioral indices of gaze perception were associated with performance across a range of social-cognition tasks, it appears that gaze perception may represent a key perceptual building block for social cognitive processes spanning emotion recognition, perceptual theory of mind, and emotion regulation.

In terms of clinical implications, it is becoming increasingly apparent that social cognitive deficits are a key dimensional feature related to psychopathology and functional outcome. Individual differences in social cognition across clinical and normal-range functioning appear to be supported by a complex set of perceptual processes and associated brain regions acting in tandem, which influence social abilities and outcomes via mechanisms that do not always clearly map onto diagnostic categories such as SZ.

References

Abbott, J., Middlemiss, M., Bruce, V., Smailes, D., & Dudley, R. (2018). The effect of arousal and eye gaze direction on trust evaluations of stranger's faces: A potential pathway to paranoid thinking. *Journal of Behavior Therapy and Experimental Psychiatry*, *60*, 29–36. https://doi.org/10.1016/j.jbtep.2018.02.007

Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, *15*(1), 1–16. https://doi.org/10.1016/S0885-2014(00)00014-9

Abplanalp, S. J., Braff, D. L., Light, G. A., Nuechterlein, K. H., Green, M. F., & Consortium on the Genetics of Schizophrenia-2. (2022). Understanding connections and boundaries between positive symptoms, negative symptoms, and role functioning among individuals with schizophrenia: A network psychometric approach. *JAMA Psychiatry*, *79*(10), 1014–1022. https://doi.org/10.1001/jamapsychiatry.2022.2386

Abram, S. V., Wisner, K. M., Fox, J. M., Barch, D. M., Wang, L., Csernansky, J. G., MacDonald, A. W., & Smith, M. J. (2017). Fronto-temporal connectivity predicts cognitive empathy deficits and experiential negative symptoms in schizophrenia. *Human Brain Mapping*, *38*(3), 1111–1124. https://doi.org/10.1002/hbm.23439

Allen, T. A., Rueter, A. R., Abram, S. V., Brown, J. S., & Deyoung, C. G. (2017). Personality and neural correlates of mentalizing ability. *European Journal of Personality*, *31*(6), 599–613. https://doi.org/10.1002/per.2133

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. https://doi.org/10.1038/nrn1884

Andreasen, N. C., & Grove, W. M. (1986). Evaluation of positive and negative symptoms in schizophrenia. *Psychiatry and Psychobiology*, *1*(2), 108–122. https://doi.org/10.1017/S0767399X00003199

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*(1), 29–52. https://doi.org/10.1111/nyas.12360

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(5), 825–839. https://doi.org/10.1080/17470218.2012.676055

Ashburner, J., Barnes, G., Chen, C.-C., Daunizeau, J., Flandin, G., Friston, K. J., Gitelman, D., Glauche, V., Henson, R., Hutton, C., Jafarian, A., Kiebel, S., Kilner, J., Litvak, V., Mattout, J., Moran, R., Penny, W., Phillips, C., Razi, A., & Zeidman, P. (2021). *SPM12 Manual*. Wellcome Trust Centre for Neuroimaging.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251. https://doi.org/10.1017/S0021963001006643

Beck, A. T., & Steer, R. A. (1993). *Manual for the Beck Depression Inventory*. Psychological Corporation.

Berchio, C., Rihs, T. A., Piguet, C., Dayer, A. G., Aubry, J. M., & Michel, C. M. (2016). Early averted gaze processing in the right fusiform gyrus: An EEG source imaging study. *Biological Psychology*, *119*, 156–170. https://doi.org/10.1016/j.biopsycho.2016.06.008

Bernhardt, B. C., Klimecki, O. M., Leiberg, S., & Singer, T. (2014). Structural covariance networks of the dorsal anterior insula predict females' individual differences in empathic responding. *Cerebral Cortex*, *24*(8), 2189–2198. https://doi.org/10.1093/cercor/bht072

Blain, S. D., Grazioplene, R. G., Ma, Y., & DeYoung, C. G. (2020). Toward a neural model of the openness-psychoticism dimension: Functional connectivity in the default and frontoparietal control networks. *Schizophrenia Bulletin*, *46*(3), 540–551. https://doi.org/10.1093/schbul/sbz103

Blain, S. D., Longenecker, J. M., Grazioplene, R. G., Klimes-Dougan, B., & DeYoung, C. G. (2020). Apophenia as the disposition to false positives: A unifying framework for openness and psychoticism. *Journal of Abnormal Psychology*, *129*(3), 279–292. https://doi.org/10.1037/abn0000504

Bliksted, V., Videbech, P., Fagerlund, B., & Frith, C. (2016). The effect of positive symptoms on social cognition in first-episode schizophrenia is

Journal of Abnormal Psychology

ISSN: 0021-843X

2020, Vol. 129, No. 3, 279–292
http://dx.doi.org/10.1037/abn0000504

Apophenia as the Disposition to False Positives: A Unifying Framework for Openness and Psychoticism

Scott D. Blain
University of Minnesota Twin Cities

Julia M. Longenecker
University of Pittsburgh and the Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, Pennsylvania

Rachael G. Grazioplene
Yale University

Bonnie Klimes-Dougan and Colin G. DeYoung
University of Minnesota Twin Cities

Positive symptoms of schizophrenia and its extended phenotype—often termed psychoticism or positive schizotypy—are characterized by the inclusion of novel, erroneous mental contents. One promising framework for explaining positive symptoms involves apophenia, conceptualized here as a disposition toward false-positive errors. Apophenia and positive symptoms have shown relations to openness to experience (more specifically, to the openness aspect of the broader openness/intellect domain), and all of these constructs involve tendencies toward pattern seeking. Nonetheless, few studies have investigated the relations between psychoticism and non–self-report indicators of apophenia, let alone the role of normal personality variation. The current research used structural equation models to test associations between psychoticism, openness, intelligence, and non–self-report indicators of apophenia comprising false-positive error rates on a variety of computerized tasks. In Sample 1, 1,193 participants completed digit identification, theory of mind, and emotion recognition tasks. In Sample 2, 195 participants completed auditory signal detection and semantic word association tasks. Psychoticism and the openness aspect were positively correlated. Self-reported psychoticism, openness, and their shared variance were positively associated with apophenia, as indexed by false-positive error rates, whether or not intelligence was controlled for. Apophenia was not associated with other personality traits, and openness and psychoticism were not associated with false-negative errors. Findings provide insights into the measurement of apophenia and its relation to personality and psychopathology. Apophenia and pattern seeking may be promising constructs for unifying the openness aspect of personality with the psychosis spectrum and for providing an explanation of positive symptoms. Results are discussed in the context of possible adaptive characteristics of apophenia as well as potential risk factors for the development of psychotic disorders.

General Scientific Summary
Our research using two samples taken from the general population suggests that the personality aspect of openness (individual differences in creativity and imagination, representing half of a broader openness/intellect trait) and symptoms of the psychosis spectrum are related to each other and may share underlying mechanisms. In the current work, we provide novel evidence that apophenia—the tendency to observe meaningful patterns where there are none present—may be an important cognitive mechanism at the core of what is shared between openness and risk for psychosis.

Keywords: schizotypy, openness to experience, false positives, schizophrenia, intelligence

Supplemental materials: http://dx.doi.org/10.1037/abn0000504.supp

Scott D. Blain, Department of Psychology, University of Minnesota Twin Cities; Julia M. Longenecker, Mental Illness Research Education Clinical Center (MIRECC), University of Pittsburgh, and the Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, Pennsylvania; Rachael G. Grazioplene, Department of Psychology, Yale University; Bonnie Klimes-Dougan and Colin G. DeYoung, Department of Psychology, University of Minnesota Twin Cities.

Data were provided in part by the HCP, WU-Minn Consortium, funded by the National Institutes of Health (1U54MH091657) and by the McDonnell Center at Washington University. Scott D. Blain was supported by the National Science Foundation Graduate Research Fellowship (1348264). Study protocols conducted in Sample 1 were approved by the Institutional Review Board of Washington University (201204036). Protocols involving Sample 2 were approved by the University of Minnesota Institutional Review Board (1410S54484). This article was published as a digital preprint, in efforts to facilitate open science (https://psyarxiv.com/d9wkc/). Preliminary results were presented by Scott D. Blain at the International Consortium for Schizotypy Research and International Society for the Study of Individual Differences meetings and as part of a colloquium at Ruhr University, Bochum.

This work was funded, in part, by the National Institutes of Health (Van Essen and Ugurbil; 1U54MH091657).

Correspondence concerning this article should be addressed to Scott D. Blain, Department of Psychology, University of Minnesota Twin Cities, N616 Elliot Hall, 75 East River Parkway, Minneapolis, MN 55455. E-mail: blain058@umn.edu

Psychosis refers to a loss of contact with reality that is present in various forms of psychopathology, most prominently schizophrenia. Symptoms of schizophrenia involve maladaptive alterations in affect, cognition, and behavior and are typically separated into negative, disorganized, and positive symptoms. Negative symptoms include constricted affect, anhedonia, and social withdrawal. Disorganized symptoms consist of disrupted thought processes that can produce incoherent or bizarre speech and behavior. Positive symptoms—hallucinations and delusions—are perhaps the most recognizable and defining feature of psychotic disorders, described as positive because they involve the addition of novel and erroneous mental content. One construct that may be useful for understanding positive symptoms across psychotic disorders is apophenia.

The term apophenia was coined by Klaus Conrad to describe a core feature of psychosis: the perception of meaningful patterns where none, in fact, exist (Mishara, 2010). In our current conceptualization, apophenia can be seen as synonymous with a disposition toward type I errors (false positives) in both perception (as in hallucinations) and belief (as in delusions). This conceptualization of apophenia is in line with the concept of aberrant salience—the inappropriate assignment of attention or meaning to external objects and internal representations—a proposed mechanism of psychosis that fits well into broader theoretical frameworks involving the function of dopamine in salience and pattern detection (Kapur, 2003). Apophenia may involve the assignment of unwarranted salience to information in one's environment, thus leading to an increase in false-positive errors. These tendencies may remain relatively widespread throughout the population because in healthy functioning, things perceived as salient are likely to involve real and meaningful patterns.

Given that hallucinations and delusions can both be seen as severe instances of type I errors, it is not surprising that individuals with schizophrenia show high rates of such errors in behavioral tasks (Blakemore, Sarfati, Bazin, & Decety, 2003). Apophenia is not limited to those with severe psychopathology, though, and it also occurs regularly throughout the general population. Apophenia, as understood today, can include any instance in which a pattern is falsely detected or labeled as meaningful when it is actually absent or attributable to chance. Everyday examples could include believing your name was called when hearing random sounds, seeing animals in the clouds, and believing in concepts such as astrology or good-luck charms. Importantly, apophenia may be the result of heightened pattern seeking, which leads to increased rates of false-positive errors, as a tradeoff for decreased false-negative (type II) errors. However, high levels of pattern seeking are not identical to apophenia. The result of a sensitive pattern detection system can be either adaptive (e.g., innovative) or maladaptive (e.g., departing from reality), but if it is to be adaptive, it must be accompanied by the capacity to distinguish between true and false patterns among those detected (DeYoung, Grazioplene, & Peterson, 2012).

The majority of existing research on apophenia has operationalized the construct using self-report measures assessing positive symptoms of schizophrenia and its extended phenotype—schizotypy. Individual differences in schizotypy are conceptualized as variation in traits that are continuously distributed and correspond to the positive, negative, and disorganized symptoms of schizophrenia (Kwapil & Barrantes-Vidal, 2014). Positive schizotypy, also known as psychoticism, is associated with ideas of reference, magical thinking, and unusual perceptual experiences—milder versions of delusions and hallucinations (Krueger, Derringer, Markon, Watson, & Skodol, 2012). This construct should not be confused with Eysenck's psychoticism, a misleadingly named trait reflecting impulsive nonconformity, low agreeableness, and low conscientiousness but not psychosis proneness (Knezevic et al., 2019). Advantages of studying psychoticism in the general population include (a) fewer confounds stemming from use of psychoactive medications, comorbidity, and functional impairment; (b) greater capacity to participate in research; and (c) a broader range of the variables of interest. Although previous attempts to measure and characterize apophenia have largely focused on self-report measures designed to assess psychoticism and related traits (e.g., DeYoung et al., 2012), a handful of studies have also attempted to measure apophenia with behavioral tasks (Brugger & Graves, 1997; Chen, Hsiao, Hsiao, & Hwu, 1998; Fyfe, Williams, Mason, & Pickup, 2008; Galdos et al., 2011; Grant, Balser, Munk, Linder, & Hennig, 2014; Mohr, Graves, Gianotti, Pizzagalli, & Brugger, 2001).

Although existing research on the non–self-report assessment of apophenia is somewhat sparse, we now provide an overview. Brugger and Graves (1997) assessed apophenia using a task in which participants are asked to complete a short puzzle game with reinforcement based on participants' response time; after the puzzle task, participants are asked whether other aspects of their behavior (which did not, in reality, influence reinforcement frequency) were related to reinforcement frequency. Another task used to assess apophenia is a theory of mind test in which simple geometric shapes move in ways that are either social or random in nature (in the social condition, the shapes move like interacting agents, and most people can clearly recognize behaviors such as help seeking, encouragement, fear, etc.); participants are asked to label the motion in each trial as either social or random (Fyfe et al., 2008). Galdos et al. (2011) used an auditory signal detection task in which participants listen to white noise and speech masked with white noise and are asked to indicate whether or not speech was present. Grant et al. (2014) used a word detection task in which participants are asked to state whether words or nonwords were present inside 15 letter strings. Finally, word association tasks are used to assess apophenia by asking participants whether or not a series of word pairs or groups are semantically related (Mohr et al., 2001). Across the aforementioned tasks and studies, false-positive errors and erroneous associations were positively associated with psychoticism.

A major weakness for many of these studies has been use of only a single task at a time to assess apophenia. False positives on any given task may be influenced by aspects of task-specific performance that are irrelevant to a general tendency toward apophenia. In the present study, therefore, we use multiple tasks and a latent variable approach to assess the general tendency toward false-positive errors. Furthermore, the current research uses two large community samples with a total of 1,388 participants, an obvious improvement over the majority of previous studies on this topic, which have typically used single samples with fewer than 100 participants (e.g., Brugger & Graves, 1997; Fyfe et al., 2008; Mohr et al., 2001).

Although apophenia has been primarily studied in relation to psychoticism, elements of normal personality may also play a role.

Efforts are increasingly being made to integrate models of normal personality and psychopathology because of the recognition that the major dimensions of risk for psychopathology correspond fairly closely to dimensions of normal personality (DeYoung & Krueger, 2018; Kotov et al., 2017). The Five-Factor Model or Big Five is well established as a description of the major dimensions of normal personality variation, and four of its dimensions (neuroticism, extraversion, agreeableness, and conscientiousness) clearly correspond to major dimensions of psychopathology (negative affect or internalizing, detachment, antagonism, and disinhibition, respectively). Psychoticism has been the dimension of psychopathology most difficult to integrate with normal personality models because it often does not seem to correspond well with the openness/intellect dimension of the Big Five, despite theoretical expectations that psychoticism resembles maladaptive openness and occasional empirical associations consistent with those expectations (De Fruyt et al., 2013; Lo et al., 2017; Piedmont, Sherman, Sherman, Dy-Liacco, & Williams, 2009; Suzuki, Samuel, Pahlen, & Krueger, 2015).

Recent empirical developments have clarified the source of this difficulty. Despite its frequent use as the label for one of the Big Five dimensions, openness to experience is actually just one of two major subfactors or aspects of the broader openness/intellect dimension, with the other being intellect (DeYoung, Quilty, & Peterson, 2007; DeYoung, Quilty, Peterson, & Gray, 2014). The openness aspect encompasses fantasy-proneness and aesthetic interests, whereas intellect reflects intellectual confidence and intellectual engagement. (When we use the term openness by itself, we are referring to the subfactor, never the broader openness/intellect dimension.) Although apophenia and psychoticism are typically at least weakly negatively associated with IQ and intellect, they are positively associated with openness and can be conceived as maladaptive variants of this personality trait, forms of openness to implausible patterns (Chmielewski, Bagby, Markon, Ring, & Ryder, 2014; DeYoung et al., 2012, 2016; Wiggins & Pincus, 1989). A number of studies show that psychoticism and openness share variance, although often less than the other Big Five traits and their psychopathological variants, at least in part because their association is suppressed by their opposite associations with intellect and IQ (DeYoung et al., 2012, 2016; Chmielewski et al., 2014). Furthermore, although psychoticism, like almost all forms of psychopathology, has a sizable association with neuroticism, it loads primarily on a factor with openness, not neuroticism, when intellect and openness are measured separately (De Fruyt et al., 2013; DeYoung et al., 2012, 2016). Additionally, evidence suggests at least partially overlapping biological correlates of openness and psychoticism (Blain, Grazioplene, Ma, & DeYoung, 2019; Grazioplene, Chavez, Rustichini, & DeYoung, 2016; Jung, Grazioplene, Caprihan, Chavez, & Haier, 2010; Smeland et al., 2017; Wright, Pahlen, & Krueger, 2017). Finally, another possible confound in clearly establishing a relation between psychoticism and openness is failure to separate positive and negative schizotypy dimensions because negative schizotypy is negatively related to openness (Chmielewski et al., 2014; Ross, Lutz, & Bailey, 2002). The present research seeks to clarify the relation between openness and psychoticism by examining their association with non-self-report measures of apophenia.

When studying the role of apophenia in psychoticism and openness, it is important to consider more general cognitive deficits as well. Low intelligence is a common risk factor for most forms of psychopathology and may confer particular vulnerability to schizophrenia (Khandaker, Barnett, White, & Jones, 2011; Zammit et al., 2004). Moreover, individuals with psychosis show specific deficits in social cognition and working memory (Mier & Kirsch, 2015; Park & Holzman, 1992), and psychoticism is also negatively correlated with performance in these domains (Blain, Peterman, & Park, 2017; Park & McTigue, 1997). Pertinent to the current study, intelligence is positively associated with stimulus discrimination (Melnick, Harrison, Park, Bennetto, & Tadin, 2013) and inversely associated with tendencies toward type I errors (Tomporowski & Simpson, 1990). Research and theory suggest that intelligence may moderate the association between apophenia and dysfunction (DeYoung et al., 2012; Grant et al., 2014). Therefore, when examining associations between psychoticism and apophenia, it is important to control for covariance explained by general cognitive ability. Few studies have examined the relation between psychoticism and non–self-report indices of apophenia in large community samples, and even fewer have considered the possible explanatory roles of intelligence and openness.

The current research sought to overcome these limitations and also incorporate a multiindicator design that assessed apophenia using a variety of computerized tasks, representing a range of perceptual and cognitive domains. As noted above, behavior in any given task is influenced by a number of task-specific factors, such that measuring a construct reliably using tasks and avoiding underestimation of true effect sizes typically requires multiple tasks. Indeed, research suggests that the psychometric quality of individual behavioral tasks is often deficient and that using multiindicator designs in a latent variable modeling framework can correct many such deficiencies (Campbell & Fiske, 1959; Nosek & Smyth, 2007). This is likely to be especially important when assessing the tendency toward broad classes of error, across diverse tasks, because errors on any single task should be influenced by specific abilities and reactions to that particular task as well as by the general tendency toward a given type of error (in this case type I errors).

Because psychoticism, openness, and apophenia all involve tendencies toward pattern seeking and perceptual sensitivity, they can be expected to share at least a moderate amount of variance. In the current research, we attempted to clarify relations among these constructs. Specifically, we investigated the relations among self-report measures of psychoticism, openness, and apophenia as indexed by false-positive errors across a range of behavioral tasks that require participants to draw connections or detect patterns. We hypothesized that latent factors corresponding to psychoticism and openness would be positively correlated and that psychoticism, openness, and a latent factor accounting for their shared variance would be associated positively with participants' disposition toward false-positive responses. Finally, we anticipated these patterns of association would be specific to false-positive errors, rather than also generalizing to variance shared with or specific to false negative errors.

Method

Participants

Sample 1 included 1,193 participants (656 females) from the Human Connectome Project. The sample included individuals between the ages of 22 and 37 years (M = 28.8, SD = 3.7). Exclusion criteria included a history of severe psychiatric, neurological, or medical disorders. Participants were not excluded for mild to moderate psychopathology (i.e., psychopathology that did not result in hospitalization or treatment over a period of greater than 1 year). Thus, given population estimates, approximately 15–20% of the sample would likely warrant a Diagnostic and Statistical Manual of Mental Disorders, fifth edition (*DSM–5*) diagnosis (National Institute of Mental Health, 2017).

Sample 2 included 195 undergraduate participants (119 females) between the ages of 18 and 44 years ($M = 21.1$, $SD = 4.8$) recruited on the University of Minnesota Twin Cities campus. They were not screened or excluded for psychiatric illness.

Self-Report Measures

Sample 1.

NEO Five-Factor Inventory (NEO-FFI; Costa & McCrea, 1992). The NEO-FFI is a measure of the Big Five personality traits. It consists of 60 items taken from the longer NEO Personality Inventory, Revised (NEO PI-R; Costa et al., 1992) and uses a 5-point Likert scale. The NEO-FFI does not include subscales for openness and intellect within the broader openness/intellect factor; thus, to create an openness aspect scale, correlations of items from the full NEO-FFI openness scale were examined in relation to openness and intellect scales from the Big Five Aspect Scales (BFAS; DeYoung et al., 2012). Previous work has been done to extract a similar openness aspect scale designed using the full NEO PI-R, based on item-associations with the BFAS in three independent samples (Ross & DeYoung, 2018); this latter scale has been used in previous research examining relations of openness and intellect with psychoticism (Suzuki, Griffin, & Samuel, 2017). Items from this NEO PI-R openness scale that are also included in the NEO-FFI were selected to create an FFI openness scale for use in the current research. These included "I don't like to waste my time daydreaming (reverse coded)," "I am intrigued by the patterns I find in art and nature," "Poetry has little or no effect on me (reverse coded)," and "Sometimes when I am reading poetry or looking at a work of art, I feel a chill or wave of excitement."

For the current study, additional validation was conducted using the Eugene Springfield Community Sample (described in our online supplementary materials). Based on item correlations in that sample, all items included in our FFI openness scale had a correlation with BFAS openness of .30 or greater and this correlation with openness was at least .15 greater than the correlation with intellect. Furthermore, there was a very strong positive correlation (approaching unity) between latent variables indicated by the four items of the FFI openness aspect scale and BFAS openness items, $r = 1.0$, $p < .001$, whereas the latent correlation between the FFI openness aspect scale and BFAS Intellect, $r = .41$, $p = < .001$ was significantly smaller ($z = 47.1$, $p < .001$). Nonetheless, to assuage concerns regarding use of an ad hoc shortened scale, we also repeated all analyses with the full NEO-FFI openness scale, which yielded results substantively identical to our main findings, likely because of the higher degree of representation of openness compared with intellect in the FFI. As per the suggestion of a

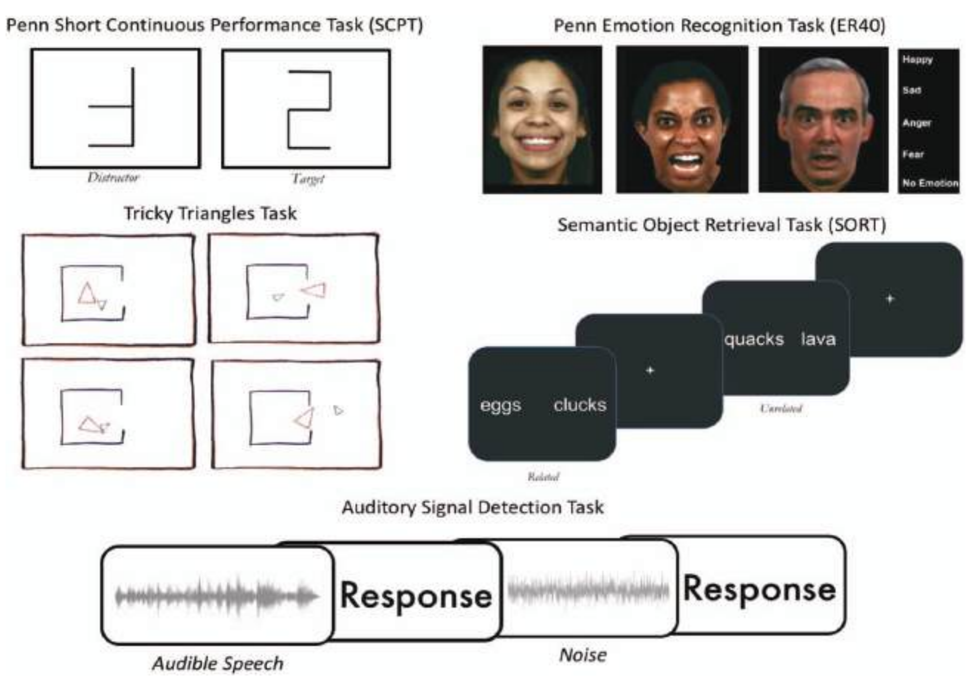


Figure 1. Behavioral apophenia tasks. Images of faces were taken from the ER40 task from the research database https://www.med.upenn.edu/bbl/2dfaces.html, which includes publicly available research images described in "A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies," by R.C. Gur, R. Sara, M. Hagendoorn, O. Marom, P. Hughett, L. Macy, T. Turner, R. Bajcsy, A. Posner, and R.E. Gur, 2002, *Journal of Neuroscience Methods*, 115, pp. 137–143. See the online article for the color version of this figure.

reviewer, neuroticism and conscientiousness items from the FFI were used in discriminant validity analyses.

Achenbach Self-Report (ASR; Achenbach, 2009). Participants were administered the ASR, an instrument used to assess psychopathology. Each item uses a three-point Likert scale. In the current study, only psychoticism items were used, as the broader thought problems scale from which they were taken includes items related to obsessiveness and self-injury. Items included "I hear sounds or voices that other people think aren't there," "I see things that other people think aren't there," "I do things that other people think are strange," and "I have thoughts that other people would think are strange."

Sample 2.

Big Five Aspect Scales (DeYoung et al., 2007). The BFAS consists of 100 items that use a 5-point Likert scale. The questionnaire is based on the Five-Factor Model and breaks down each of the factors into two aspects. In the current study, the openness aspect scale was used. For latent variable modeling, items were broken into four parcels, using the first four items to create two parcels of two items each and the remaining six items to create two parcels of three items each. The neuroticism and conscientiousness items from the BFAS were used in discriminant validity analyses.

Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008). Participants completed 34 MPQ items that form the absorption subscale. The full name of the construct measured is openness to absorbing and self-altering experiences (Tellegen & Atkinson, 1974), which is a good representation of the shared variance of openness and psychoticism (DeYoung, Carey, Krueger, & Ross, 2016). Each item used a 5-point Likert scale.

Personality Inventory for the DSM–5 (PID-5; Krueger et al., 2012). The PID-5 is a questionnaire that includes 220 items rated on a 4-point Likert scale. This inventory is used to assess symptoms of personality disorder in the alternative model from Section III of *DSM–5*. The PID-5 includes 25 facet scales that reflect five higher-order dimensions: negative affect, detachment, psychoticism, antagonism, and disinhibition. For the current study, the eccentricity, perceptual dysregulation, and unusual beliefs and experiences facet scales (from the broader psychoticism domain) were used.

Short Oxford-Liverpool Inventory of Feelings and Experiences (Mason, Linney, & Claridge, 2005). This inventory is a multidimensional measure that assesses positive, negative, and disorganized schizotypy as well as nonconformity. Each item uses a true or false response. In the current study, we used only a sum score for the unusual experiences subscale.

Intelligence Measures

Sample 1. Intelligence was assessed using the National Institutes of Health Toolbox Picture Vocabulary and List Sorting tests as well as the Penn Matrix Test. In the Picture Vocabulary test, participants selected which picture from a selection set most closely matched the meaning of a presented word. For the List Sorting test, participants were required to remember and sort a list

Table 1
Descriptive Statistics

Variables	*M* (*SD*)	Skewness	[Minimum, maximum]
Sample 1			
Openness	2.2 (0.8)	0.1	[0, 4]
Psychoticism	0.4 (0.9)	2.5	[0, 6]
ToM false positives (%)	8.5 (10.0)	4.6	[0, 100]
SPCPT false positives	28 (5.5)	1.5	[0, 28]
Emotion false positives	0.9 (1.3)	2.0	[0, 8]
ToM false negatives (%)	2.7 (8.3)	3.5	[0, 60]
SPCPT false negatives	3.6 (2.3)	0.9	[0, 15]
Emotion false negatives	3.6 (2.3)	0.9	[0, 15]
Matrix test accuracy	16.7 (4.9)	−0.6	[4, 24]
Picture vocabulary	116.6 (9.9)	0.1	[90.7, 153.1]
List sorting	110.9 (11.3)	0.2	[80.8, 144.5]
Sample 2			
PID-5 eccentricity	1.1 (0.7)	0.2	[0, 3]
PID-5 unusual beliefs	0.7 (0.6)	0.7	[0, 2.3]
PID-5 perceptual	0.7 (0.5)	0.6	[0, 2.3]
O-LIFE positive	9.2 (2.4)	−0.9	[0, 12]
MPQ absorption	37.7 (18.8)	0.3	[4, 92]
BFAS openness	3.5 (0.5)	0.2	[2.1, 5]
Semantic false positives	5.7 (4.6)	1.9	[0, 29]
Auditory false positives	0.9 (1.9)	3.6	[0, 12]
Semantic false negatives	9.0 (5.0)	1.1	[0, 30]
Auditory false negatives	11.9 (7.3)	1.9	[0, 57]
WAIS vocabulary scaled	12.3 (2.4)	−0.3	[2, 19]
WAIS similarities scaled	11.9 (2.4)	−0.1	[6, 18]
WAIS matrix scaled	10.7 (2.6)	−0.2	[3, 18]
WAIS block design scaled	11.3 (3.1)	0.0	[4, 19]

Note. ToM = theory of mind task; SPCPT = short Penn continuous performance task; PID-5 = Personality Inventory for the *DSM–5*; O-LIFE = Oxford-Liverpool Inventory of Feelings and Experiences; WAIS-IV = Weschler Adult Intelligence Scale, fourth edition; MPQ = Multidimensional Personality Questionnaire; BFAS = Big Five Aspect Scales.

Table 2
Pearson Correlations Among Sample 1 Measures

Variables	O-1	O-2	O-3	O-4	P-1	P-2	P-3	P-4	PMAT	Vocab	List	ToM FP	SPCPT FP	Emo FP	ToM FN	SPCPT FN	Emo FN
O-1	1.0																
O-2	.18	1.0															
O-3	.14	.40	1.0														
O-4	.12	.49	.67	1.0													
P-1	.02	.05	.06	.07	1.0												
P-2	.02	.02	.06	.07	.48	1.0											
P-3	.08	.11	.12	.16	.18	.15	1.0										
P-4	.12	.15	.12	.16	.16	.18	.61	1.0									
PMAT	.14	.13	−.02	.01	−.04	−.07	−.01	.05	1.0								
Vocab	.23	.19	.02	.04	−.08	−.10	−.05	.02	.47	1.0							
List	.11	.10	−.04	−.03	.00	−.05	−.04	.02	.37	.34	1.0						
ToM FP	−.01	.01	.07	.06	.14	.08	.07	.08	−.09	−.10	−.06	1.0					
SPCPT FP	.00	.01	.01	.03	.07	.08	.08	.09	−.21	−.17	−.15	.06	1.0				
Emo FP	.04	.04	.09	.09	.01	.00	.08	.05	−.12	−.02	−.10	.06	.07	1.0			
ToM FN	−.10	−.09	.01	−.01	−.03	.03	−.01	−.02	−.23	−.30	−.14	.07	.10	.02	1.0		
SPCPT FN	−.07	.01	.03	.05	.03	.03	.07	.00	−.19	−.20	−.14	.00	.19	.09	.09	1.0	
Emo FN	−.07	−.12	−.02	−.03	.02	−.02	.03	.00	−.17	−.18	−.10	.04	.08	−.04	.10	.05	1.0

Note. O 1–4 = openness NEO-FFI items –4; P1-4 = Achenbach psychoticism items 1–4; PMAT = Penn Matrix Test; Vocab = Picture Vocabulary Test; List = List Sorting Test; ToM FP = theory of mind task false positives; SPCPT FP = short Penn continuous performance task false positives; Emo FP = emotion recognition task false positives; ToM FN = theory of mind task false negatives; SPCPT FN = short Penn continuous performance task false negatives; Emo FN = emotion recognition task false negatives. Correlations with absolute value >.05 are significant at a two-tailed alpha value of .05.

Table 3
Pearson Correlations Among Sample 2 Measures

Variables	Eccentric	Unusual	Perceptual	O-LIFE Pos	MPQ Abs	BFAS-O 1	BFAS-O 2	BFAS-O 3	BFAS-O 4	Semantic FP	Auditory FP	Semantic FN	Auditory FN	Vocab	Similarity	Matrix	Block Design
Eccentric	1.0																
Unusual	.60	1.0															
Perceptual	.66	.74	1.0														
O-LIFE Pos	.48	.56	.67	1.0													
MPQ-Abs	.65	.74	.79	.61	1.0												
BFAS-O 1	.07	.04	.08	.10	.27	1.0											
BFAS-O 2	.13	.10	.17	.12	.36	.43	1.0										
BFAS-O 3	.16	.21	.18	.23	.32	.57	.41	1.0									
BFAS-O 4	.13	.20	.10	.19	.29	.30	.34	.35	1.0								
Semantic FP	.07	.15	.14	.20	.13	.11	.03	.04	.02	1.0							
Auditory FP	.04	.05	.01	.14	.08	.05	.08	.00	−.09	.10	1.0						
Semantic FN	−.19	−.04	−.13	−.10	−.10	−.17	−.05	−.12	−.18	.04	.13	1.0					
Auditory FN	.08	.18	.18	.13	.06	−.07	−.04	−.02	−.11	.13	−.03	.15	1.0				
Vocabulary	.18	.03	.12	.08	.18	.20	.12	.23	.19	−.07	.01	−.24	−.25	1.0			
Similarities	.19	.10	.18	.07	.24	.22	.22	.14	.27	−.08	−.04	−.13	−.12	.45	1.0		
Matrix	.04	.03	−.06	.07	.04	.00	.10	.06	.13	−.02	.05	−.09	−.22	.17	.11	1.0	
Block design	.08	.11	.05	−.05	.10	.19	.10	.05	.16	−.07	.01	−.29	−.17	.25	.18	.20	1.0

Note. Eccentric, Unusual, Perceptual = PID-5 psychoticism sub-scales; O-LIFE = Oxford-Liverpool Inventory of Feelings and Experiences; MPQ-Abs = Absorption scale from the Multidimensional Personality Questionnaire; BFAS-O 1–4 = item parcels 1–4 for Big Five Aspect Scales Openness; FP = raw false positive rates; FN = raw false negative rates. Correlations with absolute value >.13 are significant at a two-tailed alpha value of .05.

of items. In the Penn Matrix Test, participants completed a set of visual patterns. The National Institutes of Health Toolbox measures of intelligence have shown a high degree of reliability and validity, as indexed by test-retest reliability and associations with gold-standard measures of intelligence (Heaton et al., 2014).

Sample 2. Intelligence was assessed using the Weschler Adult Intelligence Scale, fourth edition (WAIS-IV; Wechsler, 2008). Four subtests of the WAIS-IV were used: block design, matrix reasoning, vocabulary, and similarities. These are the four subtests recommended by Wechsler as a brief but accurate estimate of full-scale IQ.

Behavioral Apophenia Measures (see Figure 1)

Sample 1.

Tricky triangles task (Abell, Happe, & Frith, 2000). Participants were presented a series of computerized animations of shapes interacting either randomly or socially. Participants indicated whether each was random or social. The social condition included a variety of interaction types, such as seducing, coaxing, and mocking. The random condition included shapes that were drifting, bouncing, or moving in other ways that did not imply social interaction. Each participant was presented five random and five social animations, with intermixed presentation order. False positives were operationalized as random trials labeled social. False negatives were operationalized as social trials labeled random.

Penn short continuous performance task (Gur et al., 2010). Participants were presented a series of sets of vertical and horizontal lines that either formed a target letter/digit or a distractor. Participants indicated whether each stimulus was a target or distractor. Items were presented at a rate of one per second, with 300 ms viewing time per stimulus. Thirty-six targets were presented randomly among a total of 360 items. False positives were operationalized as distractors labeled targets. False negatives were operationalized as targets labeled distractors.

Penn emotion recognition task (Gur et al., 2001). Participants were presented a series of 40 faces and asked to identify what emotion each was expressing. Emotion options included happy, sad, anger, fear, and no emotion. For each emotion, eight faces were presented. False positives were operationalized as no emotion trials labeled happy, sad, anger, or fear. False negatives were operationalized as total number of incorrect responses across the happy, sad, anger, and fear trials.

Sample 2.

Semantic object retrieval task (Assaf et al., 2009). Participants were presented a series of 92 word pairs that were either related or unrelated. Participants indicated whether the two words were related. Examples of related items included popcorn and theater. Examples of unrelated items included books and kernel. False positives were operationalized as unrelated items labeled related. False negatives were operationalized as related items labeled unrelated.

Auditory signal detection task (Galdos et al., 2011). Participants were asked to indicate whether speech was present, in a series of auditory stimuli, presented in four conditions: (a) white noise only, (b) barely audible speech masked with white noise, (c) audible speech masked with white noise, and (d) speech only. Volume of trials was consistent among participants. False positives were operationalized as white-noise-only trials labeled as having speech present. False negatives were operationalized as speech trials labeled noise.

Analysis

Descriptive statistics were calculated for self-report, intelligence, and false-negative and false-positive rates (Table 1). Variables with a skew >2.0 were logarithmically transformed to approximate normality. Structural equation modeling was then used to assess relations among latent factors. Using raw false-positive and false-negative variables for each task, apophenia variables were created for each task by residualizing false-positive scores by regressing them on false-negative scores, resulting in variables representing unique variance of false-positive tendencies for each task. This eliminates the potential confound that poor performance on a task is likely to involve both false positives and false negatives (indeed, the two types of error tended to be positively correlated within each task; Tables 2 and 3). Similar variables were created to represent unique variance in false negatives.

For Sample 1, the corresponding items or tasks were allowed to load onto latent variables for psychoticism, openness, apophenia, and intelligence. Because two psychoticism items from the ASR were strongly semantically related, both entailing that others find the target strange, an a priori decision was made to allow their residuals to correlate. If they were not allowed to correlate, model fit decreased significantly ($\Delta\chi^2 = 265.1$, $p < .001$). For Sample 2,

Table 4
Model Fit Statistics

Models	χ^2	p	RMSEA	95% CI	TLI
Sample 1					
Model 1 (psychoticism and O)	71.2	<.001	.050	[.038, .062]	.960
Model 2 (psychoticism, IQ, and apophenia)	65.8	<.001	.033	[.022, .044]	.958
Model 3 (O, IQ, and apophenia)	144.9	<.001	.058	[.049, .068]	.897
Model 4 (O-psychoticism, IQ, and apophenia)	228.5	<.001	.047	[.040, .053]	.911
Sample 2					
Model 1 (psychoticism and O)	43.7	.012	.060	[.028, .089]	.971
Model 2 (psychoticism, IQ, and apophenia)	46.4	.048	.046	[.004, .074]	.961
Model 3 (O, IQ, and apophenia)	39.0	.183	.032	[.000, .064]	.961
Model 4 (O-psychoticism, IQ, and apophenia)	125.1	.002	.048	[.029, .066]	.951

Note. CI = confidence interval; RMSEA = root mean square error of approximation; TLI = Tucker-Lewis index; O = openness.

scale scores were used as indicators for psychoticism and openness, and, in models including both psychoticism and openness, MPQ absorption was added as a manifest indicator with cross-loadings from both psychoticism and openness. Intelligence in Sample 2 was indicated by four subtests of the WAIS, with residuals from the verbal intelligence subscales (vocabulary and similarities) allowed to correlate. For Sample 2, given only two tasks, factor loadings for the apophenia factor were constrained to be equal.

Maximum likelihood estimation was used to fit all models. First, confirmatory factor analytic models were fit to examine the relations between openness and psychoticism. Structural equation models were then fit to examine the prediction of apophenia by (a) psychoticism and intelligence, (b) openness and intelligence, and (c) shared variance of openness and psychoticism (henceforth O-P) and intelligence. At the request of reviewers, a number of supplemental models were also fit. First, all of the aforementioned associations were tested without controlling for intelligence and with intelligence-by-O-P interactions. Second, openness and psychoticism were examined as simultaneous predictors of apophenia. Third, we conducted a set of discriminant validity analyses, using neuroticism and conscientiousness to predict our criterion variables. Finally, we computed models using residualized false negatives as criterion latent variables, as a further test of discriminant validity.

Although a strength of the present study was using multiple tasks simultaneously to assess apophenia, we nonetheless fit a set of auxiliary models to assess the relation of our questionnaire variables with apophenia scores on individual tasks. In addition to testing our hypothesis regarding the advantages of using multiple tasks to assess apophenia, this analysis also provided closer replications of several previous studies (Chen et al., 1998; Fyfe et al., 2008; Galdos et al., 2011). The same latent predictors were used as in the main models, with the observed apophenia scores (residualized false positives) from each task used, in turn, as the criterion variables.

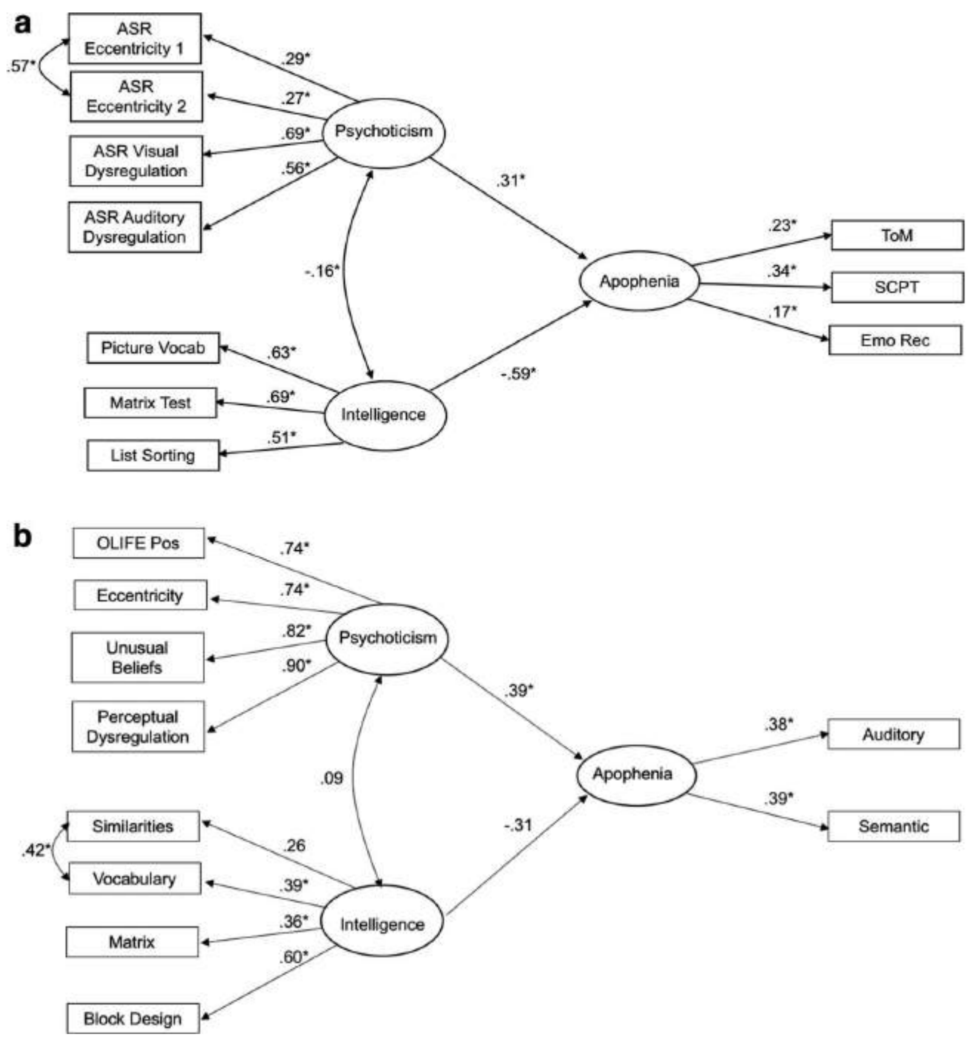


Figure 2. Model of psychoticism, intelligence, and apophenia, Samples 1(a) and 2(b). ASR = Achenbach Self-Report; ToM = theory of mind task; OLIFE = Oxford-Liverpool Inventory of Feelings and Experiences; SCPT = Penn short continuous performance task. $^{*}p < .05$.

Results

Descriptive statistics are presented in Table 1. See Tables 2 and 3 for Pearson correlations among study measures. Structural equation modeling fit statistics for our primary models are presented in Table 4. All models showed acceptable fit using the root mean square error of approximation criterion, with values ranging from .032 to .060. Factor loadings of assigned manifest variables onto their corresponding latent variables were significant, across models, albeit small in magnitude for behavioral indicators of apophenia.

Latent psychoticism was positively correlated with latent openness, in Sample 1, $r = .14$, $p = .005$, and Sample 2, $r = .31$, $p = .001$. Latent psychoticism was positively associated with apophenia, in Sample 1 (Figure 2a) and Sample 2 (Figure 2b). Like psychoticism, openness was significantly positively associated with apophenia in Sample 1 (Figure 3a); the path from openness to apophenia did not reach statistical significance in Sample 2 (Figure 3b), but magnitude of the standardized path coefficient in this model was similar to that found in Sample 1 (which had a larger *N* and therefore higher statistical power). If openness and psychoticism were used as simultaneous predictors (while controlling for Intelligence), their individual regression weights were somewhat diminished, and only psychoticism was a significant predictor of apophenia in either sample (Supplemental Figures 1 and 2). This pattern of suppression is consistent with the hypothesis that the shared variance of openness and psychoticism is associated with apophenia. Indeed, in our final set of models, apophenia was positively associated with O-P shared variance in Sample 1 (see Figure 4) and Sample 2 (see Figure 5). The relations of latent scores for O-P and apophenia are visualized using scatterplots in Figure 6.

Associations between O-P and apophenia remained largely consistent across samples and variables, whether or not Intelligence was controlled for (Supplemental Table 1), and these associations were not present for false negatives (Supplemental Table 2). In Sample 1, intelligence was a significant negative predictor of apophenia, in two of three models. Intelligence strongly negatively predicted false negatives across samples and models. Across mod-

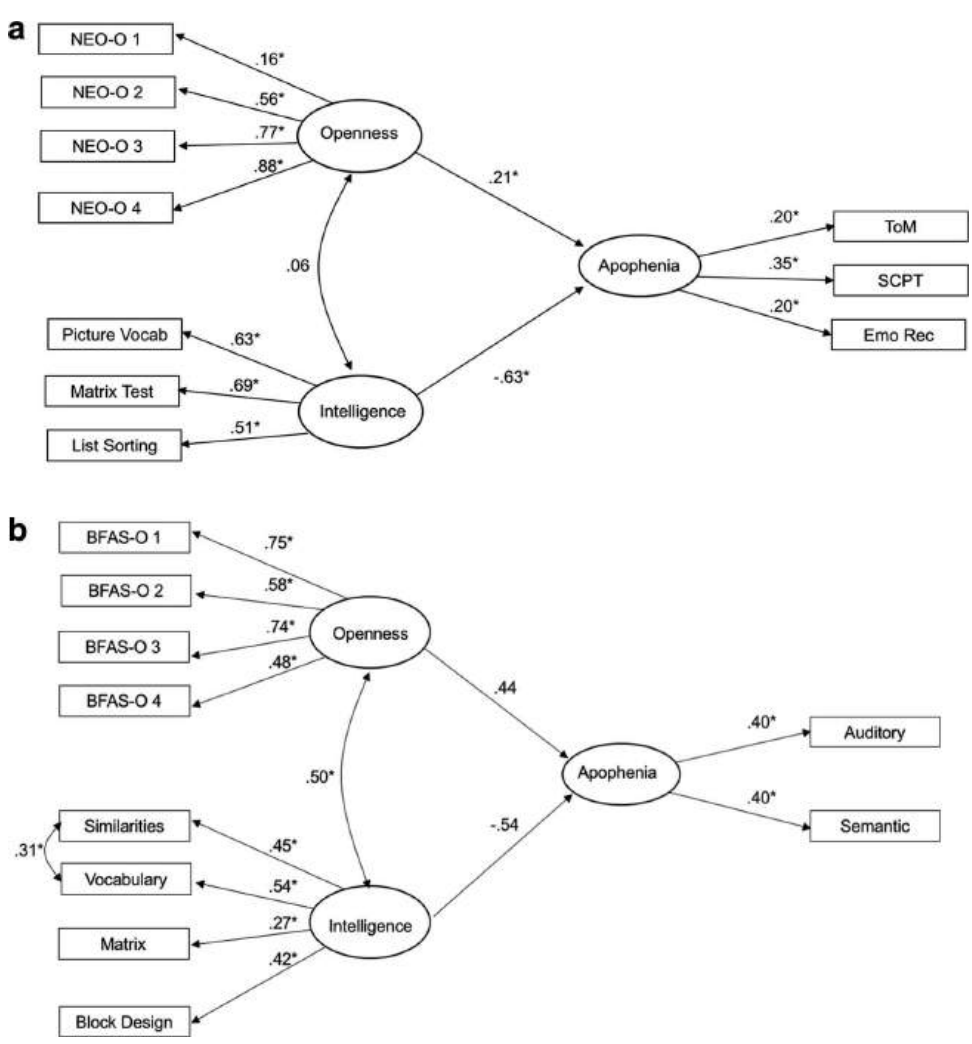


Figure 3. Model of openness, intelligence, and apophenia, Samples 1(a) and 2(b). ToM = theory of mind task; BFAS = Big Five Aspect Scales; SCPT = Penn short continuous performance task. $^{*} p < .05$.

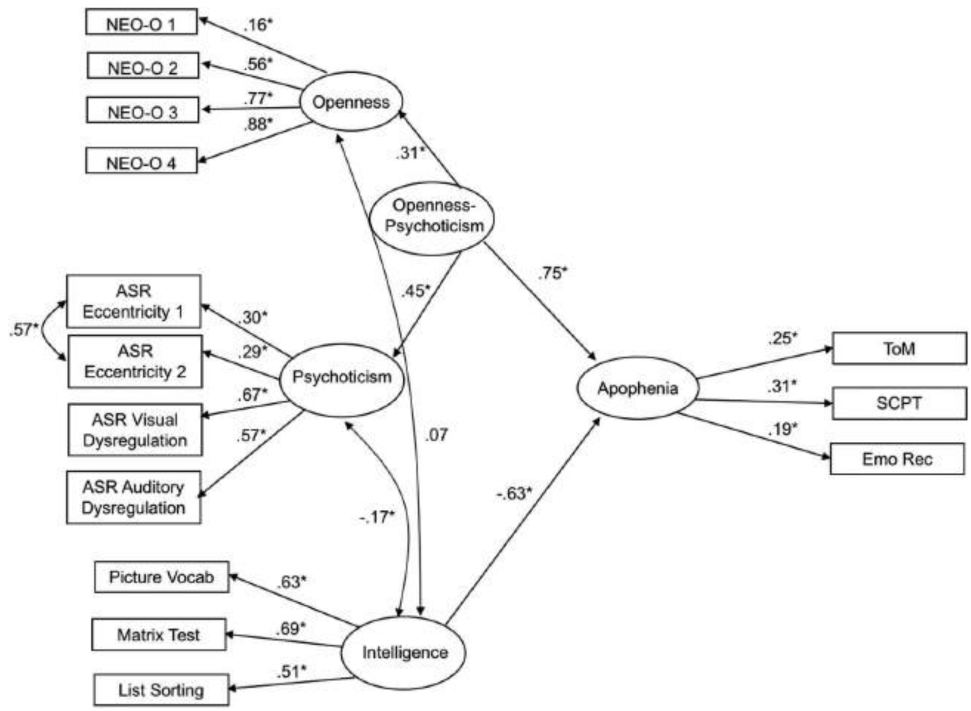


Figure 4. Model of openness-psychoticism, intelligence, and apophenia, Sample 1. ASR = Achenbach Self-Report; ToM = theory of mind task; SCPT = Penn short continuous performance task. * $p < .05$.

els, findings remained substantively identical if we used raw, rather than residualized, false-positive and -negative variables. At the suggestion of a reviewer, we also tested for interactions between intelligence and O-P, in predicting apophenia (Supplemental Table 1). These were mostly nonsignificant across variables and samples, although there was a significant interaction of intelligence and O-P in Sample 1 ($\beta = -.31$, $p < .005$). Tests of discriminant validity showed that apophenia was not predicted by conscientiousness or neuroticism (Supplemental Table 3). Positive associations of psychoticism, openness, and O-P with apophenia were found for the majority of specific tasks, but relations were least robust for the auditory signal detection task (Supplemental Table 4). Path coefficients from O-P variables to observed apophenia variables in the specific task models were smaller in magnitude, compared with our main models containing latent apophenia factors.

Discussion

Findings supported our hypothesis that apophenia, as indexed by false-positive rates, is positively associated with self-report measures of psychoticism and the openness aspect of openness/intellect. Furthermore, associations were specific to false positives because psychoticism and openness were unrelated to false negatives, so our findings do not simply represent a generalized decrement in task performance. Relations were robust and replicable in two samples, across different tasks and, importantly, did not hold for other potentially relevant personality traits, including neuroticism and conscientiousness. Standardized regression coefficients for predictions of apophenia by shared O-P variance were high, ranging from .61 to .75, suggesting apophenia may be a good representation of what openness and psychoticism have in common. Apophenia of the relatively mild sort represented by false positives on behavioral tasks likely reflects core tendencies toward pattern seeking and perceptual sensitivity, which may, in turn, result in openness to implausible patterns. Our findings suggest that apophenia may be a key cognitive process contributing to risk for psychosis. Findings also contribute to the growing body of research that identifies transdiagnostic risk factors linked to normal personality variation (DeYoung et al., 2012, 2016; Kotov et al., 2017).

The current findings replicate and unite previous empirical work. False alarms in digit identification, social triangles, white noise, and semantic association tasks have each independently been found to correlate positively with psychosis-spectrum measures (Chen et al., 1998; Fyfe et al., 2008; Galdos et al., 2011; Mohr et al., 2001). Our results extend these individual studies, suggesting that single-task indicators of apophenia may underestimate the relation between self-reported psychoticism and behavioral apophenia. Indeed, in our present results, effect sizes for associations with individual-task false positives were much smaller in magnitude than associations with latent behavioral apophenia. Future studies would benefit from assessing a range of behavioral apophenia tasks to overcome the limitations associated with task-specific variance.

Although the increased false positives associated with apophenia may seem deleterious, a tendency toward pattern seeking and perceptual sensitivity (which seems likely to underlie apophenia, openness, and psychoticism) could have provided evolutionary

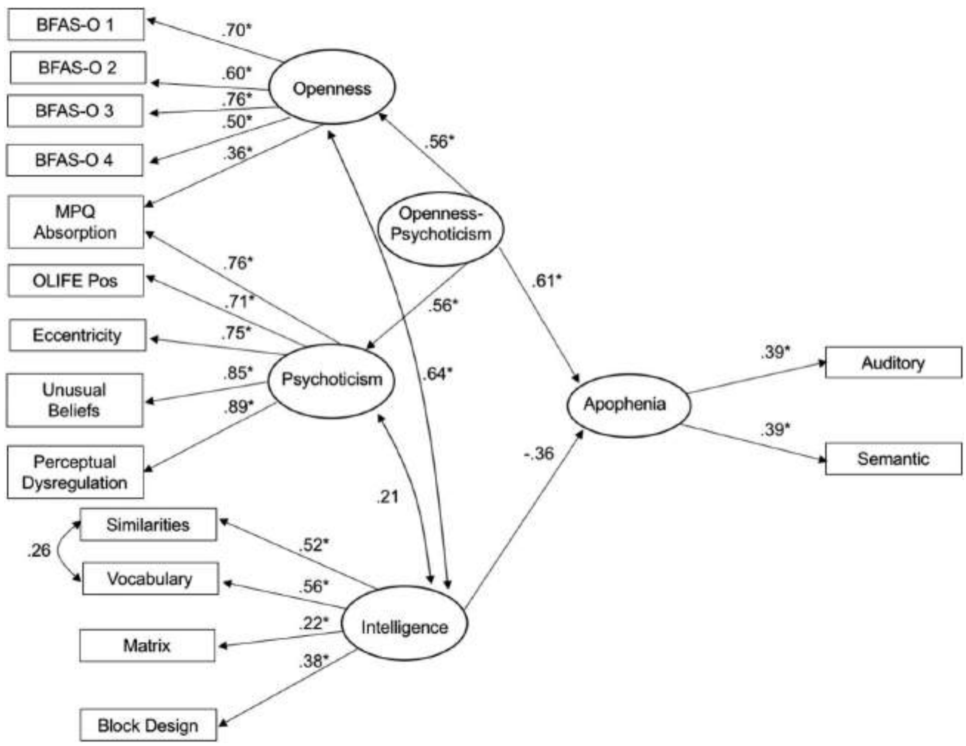


Figure 5. Model of openness-psychoticism, intelligence, and apophenia, Sample 2. BFAS = Big Five Aspect Scales; MPQ = Multidimensional Personality Questionnaire; OLIFE = Oxford-Liverpool Inventory of Feelings and Experiences. * $p < .05$.

advantages. Like any animal, a human being must balance false positives against false negatives in modeling regularities in its experience. Favoring false positives can aid in avoiding threat, even if it sometimes leads to an excess of caution. In contrast, a bias toward false negatives can result in failure to detect actual threats and thus constitutes a threat to survival. Furthermore, favoring false positives may confer advantages in domains in which creative pattern detection is advantageous, for example, among artists, inventors, and even psychics—precisely those populations that are high in openness (Beck & Forstmeier, 2007; Powers, Kelley, & Corlett, 2017). Fitness advantages of high openness may explain the persistence of apophenia throughout the population, despite the detrimental effects of full-blown psychosis. Apophenia might therefore have had a U-shaped relation to fitness, with a moderate level being an acceptable cost for high creativity. Indeed, research suggests both openness and psychoticism predict creativity and creative achievement, regardless of whether or not shared variance with intelligence or intellect is partialed out (Kaufman et al., 2016). Although apophenia and related traits or mechanisms (e.g., openness and pattern seeking) might be associated with certain adaptive characteristics at moderately high levels, when extreme or coupled with other risk factors (such as negative or disorganized symptoms), they may result in severe psychopathology.

Importantly, high intelligence may play a protective role when present in individuals with high openness, allowing better detection of which perceived patterns are likely to be useful rather than spurious (DeYoung et al., 2012). After a false-positive error has initially occurred, intellectual deficits may contribute to an inability to separate real patterns from spurious ones, possibly leading to suspiciousness, paranoia, and delusional thinking associated with psychosis. Perhaps an optimal balance between type I and type II error rates occurs when an individual has high levels of intelligence as well as openness (which is typical, given that they are positively correlated), allowing high levels of creativity and cognitive exploration, without functional impairment (DeYoung, 2015). This personality profile would likely be associated with a high degree of initial type I errors (because of high openness), paired with effective screening of these errors (because of high intelligence). In other words, some tendency toward apophenia could be advantageous when coupled with intact reality testing. Thus, individuals high in openness who have intact cognitive functioning may exhibit above average functional outcomes, despite a tendency toward apophenia, given that the combination of high levels of pattern seeking and intelligence likely facilitates the generation and application of creativity and innovative thinking.

Results from Sample 1 are consistent with this account because apophenia was negatively correlated with intelligence. In Sample 2, however, there was no relation between intelligence and apophenia, although intelligence was strongly associated with false-negative error rates. This difference might reflect either the fact that different tasks were used or that the second sample was exclusively college students, a group with higher-than-average intelligence. Further research is needed to clarify the possible protective role of intelligence in relation to apophenia and the psychosis spectrum.

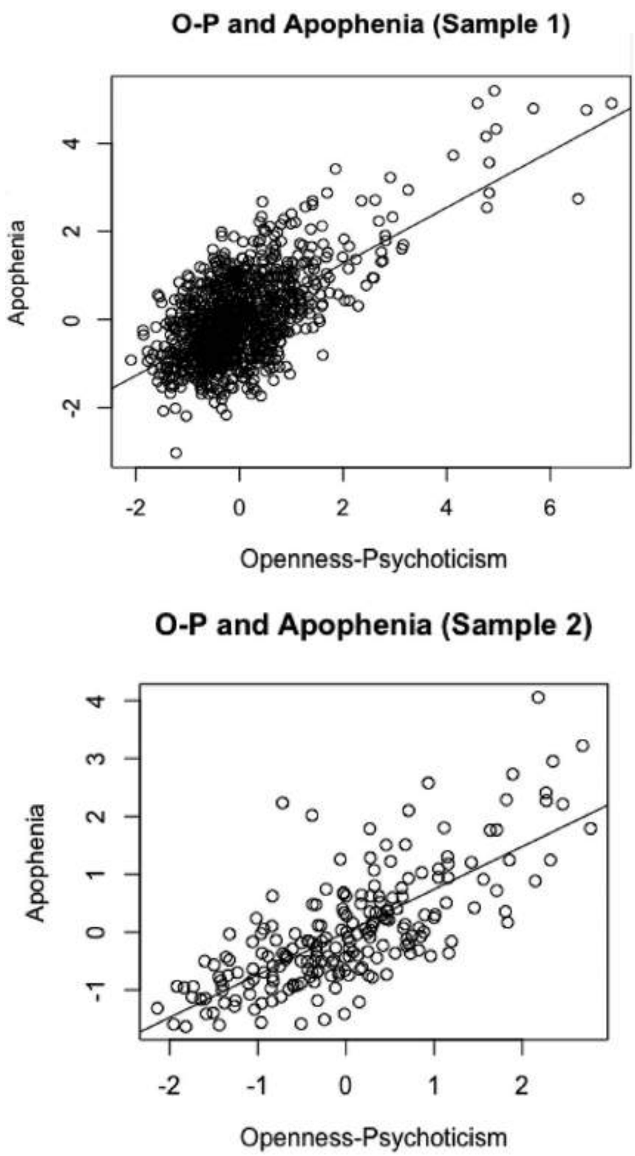


Figure 6. Scatterplots of latent variable relations.

The current study provides insights into the measurement of apophenia using behavioral indicators, an important step in better characterizing cognitive correlates of openness and psychoticism. Apophenia is a scientifically informative transdiagnostic feature seen across a number of disorders characterized by psychosis, including schizotypal and paranoid personality, schizophrenia, and depression and bipolar disorder with psychotic features (Chmielewski et al., 2014; Hanssen et al., 2003; Mishara, 2010; Narayan, Allen, Cullen, & Klimes-Dougan, 2013). Further elucidating the role of apophenia and pattern detection as a cognitive mechanism may prove useful for explaining variation in the tendency to make false-positive errors, in populations ranging from healthy individuals to those with functionally impairing psychiatric concerns. Such an approach is in line with the Hierarchical Taxonomy of Psychopathology conceptualization of psychiatric illness (Kotov et al., 2017) and with the National Institute of Mental Health's Research Domain Criteria initiative, particularly its cognitive systems dimension (Cuthbert & Insel, 2013). Furthering research into the non–self-report assessment of apophenia using tasks such as those in the current study could allow for research to circumvent problems inherent to questionnaire-based research, an important step for better characterizing correlates and mechanisms underlying the psychosis spectrum, as well as psychopathology more broadly.

Although the current study had multiple strengths, some limitations are worth noting. Despite the fact that our current measures were selected to directly mirror and conceptually replicate past research and have shown association with openness-psychoticism in previous studies and the current sample, it is worth noting that the factor loadings of apophenia variables onto their latent factor were small in magnitude. This is likely to be an unavoidable feature of assessing false positives across a diverse range of tasks, but future work could incorporate additional or even more direct measures of apophenia, such as the Salience Attribution Test (Roiser et al., 2009).

Despite the potential of latent variables comprising multiple task or questionnaire indicators to capture truer estimates of effect size, compared with models using only individual manifest variables, this method also runs the risk of inflating effect sizes beyond their population values, when indicators share only a small portion of variance. Thus, it is important to interpret results of the current study in terms of the general trends seen across models, indicators, and traits (i.e., the consistently demonstrated, yet mild-in-magnitude relations among openness, psychoticism, and apophenia manifest and latent variables), rather than merely focusing on effect sizes of our most comprehensive structural models.

Our measures of openness and psychoticism were self-reported and could be usefully supplemented in future research by peer reports or clinician ratings. Furthermore, we cannot tell how well the current results would generalize to clinical populations, and future research should examine the roles of apophenia in those with active psychosis or at clinical high risk. Finally, although our samples included a variety of psychoticism and personality measures and tested apophenia across a wide variety of tasks, Sample 1 had brief questionnaire measures with relatively few items, and Sample 2, although it had better questionnaire measurement, had a considerably smaller number of participants. These limitations in our current work could be addressed in the future by recruiting additional large samples with extensive, high-quality measures of psychoticism, openness, and apophenia.

Conclusion

The current study advances research on openness, psychoticism, and their possible underlying mechanisms. Apophenia (i.e., false-positive error rate) was shown to be associated with both the openness aspect of openness/intellect and with psychoticism, which is consistent with the role pattern seeking and perceptual sensitivity play in all of these traits. Apophenia, as indexed by behavioral measures of the disposition toward false positives, may provide a useful transdiagnostic construct for studying the cognitive correlates and possible mechanisms of psychosis and related phenomena, across a range of mental disorders and normal personality variation. Understanding how psychoticism relates to normal personality variation and specific cognitive mechanisms is crucial for advancing our understanding of psychosis risk in the general population. Our current work furthers this aim and also adds to a growing body of literature demonstrating the promise of characterizing the underlying correlates and mechanisms of psychiatric features through the use of large, nonclinical samples.

References

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15,* 1–16. http://dx.doi.org/10.1016/S0885-2014(00)00014-9

Achenbach, T. M. (2009). *Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications*. Burlington: University of Vermont Research Center of Children, Youth & Families.

Assaf, M., Jagannathan, K., Calhoun, V., Kraut, M., Hart, J., Jr., & Pearlson, G. (2009). Temporal sequence of hemispheric network activation during semantic processing: A functional network connectivity analysis. *Brain and Cognition, 70,* 238–246. http://dx.doi.org/10.1016/j.bandc.2009.02.007

Beck, J., & Forstmeier, W. (2007). Superstition and belief as inevitable by-products of an adaptive learning strategy. *Human Nature, 18,* 35–46. http://dx.doi.org/10.1007/BF02820845

Blain, S. D., Grazioplene, R. G., Ma, Y., & DeYoung, C. G. (2019). Toward a neural model of the Openness-Psychoticism dimension: Functional connectivity of the default and frontoparietal control networks. *Schizophrenia Bulletin*. Advance online publication. http://dx.doi.org/10.1093/schbul/sbz103

Blain, S. D., Peterman, J. S., & Park, S. (2017). Subtle cues missed: Impaired perception of emotion from gait in relation to schizotypy and autism spectrum traits. *Schizophrenia Research, 183,* 157–160. http://dx.doi.org/10.1016/j.schres.2016.11.003

Blakemore, S. J., Sarfati, Y., Bazin, N., & Decety, J. (2003). The detection of intentional contingencies in simple animations in patients with delusions of persecution. *Psychological Medicine, 33,* 1433–1441. http://dx.doi.org/10.1017/S0033291703008341

Brugger, P., & Graves, R. E. (1997). Testing vs. believing hypotheses: Magical ideation in the judgement of contingencies. *Cognitive Neuropsychiatry, 2,* 251–272.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105. http://dx.doi.org/10.1037/h0046016

Chen, W. J., Hsiao, C. K., Hsiao, L. L., & Hwu, H. G. (1998). Performance of the Continuous Performance Test among community samples. *Schizophrenia Bulletin, 24,* 163–174. http://dx.doi.org/10.1093/oxfordjournals.schbul.a033308

Chmielewski, M., Bagby, R. M., Markon, K., Ring, A. J., & Ryder, A. G. (2014). Openness to experience, intellect, schizotypal personality disorder, and psychoticism: Resolving the controversy. *Journal of Personality Disorders, 28,* 483–499. http://dx.doi.org/10.1521/pedi_2014_28_128

Costa, P. T., & McCrea, R. R. (1992). *Revised neo personality inventory (neo pi-r) and neo five-factor inventory (neo-ffi)*. Odesa, FL: Psychological Assessment Resources.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11,* 126. http://dx.doi.org/10.1186/1741-7015-11-126

De Fruyt, F., De Clercq, B., De Bolle, M., Wille, B., Markon, K., & Krueger, R. F. (2013). General and maladaptive traits in a five-factor framework for *DSM–5* in a university student sample. *Assessment, 20,* 295–307. http://dx.doi.org/10.1177/1073191113475808

DeYoung, C. G. (2015). Openness/Intellect: A dimension of personality reflecting cognitive exploration. *APA Handbook of Personality and Social Psychology: Personality Processes and Individual Differences, 4,* 369–399.

DeYoung, C. G., Carey, B. E., Krueger, R. F., & Ross, S. R. (2016). Ten aspects of the Big Five in the Personality Inventory for. *DSM–5*. *Personality Disorders, 7,* 113–123. http://dx.doi.org/10.1037/per0000170

DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality, 46,* 63–78. http://dx.doi.org/10.1016/j.jrp.2011.12.003

DeYoung, C. G., & Krueger, R. F. (2018). A cybernetic theory of psychopathology. *Psychological Inquiry, 29,* 117–138.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93,* 880–896. http://dx.doi.org/10.1037/0022-3514.93.5.880

DeYoung, C. G., Quilty, L. C., Peterson, J. B., & Gray, J. R. (2014). Openness to experience, intellect, and cognitive ability. *Journal of Personality Assessment, 96,* 46–52. http://dx.doi.org/10.1080/00223891.2013.806327

Fyfe, S., Williams, C., Mason, O. J., & Pickup, G. J. (2008). Apophenia, theory of mind and schizotypy: Perceiving meaning and intentionality in randomness. *Cortex, 44,* 1316–1325. http://dx.doi.org/10.1016/j.cortex.2007.07.009

Galdos, M., Simons, C., Fernandez-Rivas, A., Wichers, M., Peralta, C., Lataster, T., . . . van Os, J. (2011). Affectively salient meaning in random noise: A task sensitive to psychosis liability. *Schizophrenia Bulletin, 37,* 1179–1186. http://dx.doi.org/10.1093/schbul/sbq029

Grant, P., Balser, M., Munk, A. J. L., Linder, J., & Hennig, J. (2014). A false-positive detection bias as a function of state and trait schizotypy in interaction with intelligence. *Frontiers in Psychiatry, 5,* 135. http://dx.doi.org/10.3389/fpsyt.2014.00135

Grazioplene, R. G., Chavez, R. S., Rustichini, A., & DeYoung, C. G. (2016). White matter correlates of psychosis-linked traits support continuity between personality and psychopathology. *Journal of Abnormal Psychology, 125,* 1135–1145. http://dx.doi.org/10.1037/abn0000176

Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., . . . Gur, R. E. (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology, 25,* 766–776. http://dx.doi.org/10.1016/S0893-133X(01)00278-0

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., . . . Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods, 187,* 254–262. http://dx.doi.org/10.1016/j.jneumeth.2009.11.017

Hanssen, M., Peeters, F., Krabbendam, L., Radstake, S., Verdoux, H., & van Os, J. (2003). How psychotic are individuals with non-psychotic disorders? *Social Psychiatry and Psychiatric Epidemiology, 38,* 149–154. http://dx.doi.org/10.1007/s00127-003-0622-7

Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., . . . Gershon, R. (2014). Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *Journal of the International Neuropsychological Society, 20,* 588–598. http://dx.doi.org/10.1017/S1355617714000241

Jung, R. E., Grazioplene, R., Caprihan, A., Chavez, R. S., & Haier, R. J. (2010). White matter integrity, creativity, and psychopathology: Disentangling constructs with diffusion tensor imaging. *PLoS ONE, 5,* e9818. http://dx.doi.org/10.1371/journal.pone.0009818

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia.

Theory of Mind and the Agreeableness-Antagonism Dimension:

Differential Associations with Callousness, Aggression, and Manipulativeness

Scott D. Blain[1]*, Aisha L. Udochi[2], Timothy A. Allen[3], Muchen Xi[2], & Colin G. DeYoung[2]

University of Michigan Department of Psychiatry[1], University of Minnesota – Twin Cities Department of Psychology[2], University of Pittsburgh Department of Psychiatry[3]

*Corresponding Author:

Scott D. Blain

University of Michigan Department of Psychiatry

Rachel Upjohn Building 2742

4250 Plymouth Road

Ann Arbor, MI 48109

(270) 287-8688

blscott@med.umich.edu

Author Note:

SDB and ALU were supported by National Science Foundation Graduate Research Fellowships, during the completion of this work. Additional support was provided by Grant No. K01-MH123915 to TAA.

Preliminary results were presented by SDB at various conferences and campus talks, including a biopsychology colloquium at Ruhr University Bochum (2019), the Association for Research in Personality (2019), International Society for the Study of Individual Differences (2019), and Society for Personality and Social Psychology (2021). Finally, this work was used in part to fulfill Ph.D. degree requirements for SDB and was included in the dissertation "Individual Differences in Social Cognition and Behavior: A Personality Psychology Framework."

Analytical scripts, study data, a list of all procedures and measures included in this study, and a list of references also utilizing the same broader dataset (currently none but other manuscripts are in preparation) will be made publicly available at the Open Science Framework (https://osf.io/3qwjr/?view_only=8ff037d28a5247c585c780faedb534bb). This article was also published as a digital preprint, in efforts to facilitate open science.

Abstract

Theory of Mind (ToM) refers to how we identify and understand the mental states of others. ToM abilities vary with dimensions of normal-range personality and can be seriously impaired in a number of mental disorders, particularly those related to the Antagonism domain. The current study used a multi-task design to examine how ToM relates to Agreeableness-Antagonism subfactors, replicating and extending previous work. Participants (*N* = 335) completed self-report measures of the Big Five, empathy, and personality pathology, as well as tasks spanning mental state attribution, affect recognition, and mentalizing. Exploratory structural equation modeling was used to assess the impact of Agreeableness-Antagonism subfactors on ToM. A three-factor structure was derived for Agreeableness-Antagonism, with factors corresponding to Compassion-Callousness, Pacifism-Aggression, and Honesty-Manipulativeness. While higher Aggression and lower Compassion predicted worse ToM ability, higher Manipulativeness predicted better ToM ability. Findings replicate and extend work suggesting differential relations of specific Agreeableness-Antagonism subfactors with social cognition. We discuss our results with a focus on the importance of dimensional psychopathology models and facet-level research.

Key Words: Social Cognition, Personality, Theory of Mind, Dark Triad, Emotional Intelligence

Theory of Mind and the Agreeableness-Antagonism Dimension: Differential Associations with Callousness, Aggression, and Manipulativeness

When navigating social interactions, humans rely upon a variety of social cognitive processes—including the ability to perceive emotions and empathize with others (Barrett et al, 2011; Singer & Klimecki, 2014). One important social cognitive process is known as *theory of mind (ToM)* or *mentalizing*, which refers to a person's ability to recognize, understand, and utilize the thoughts, feelings, and beliefs of other people (Premack & Woodruff, 1978). People vary in their ToM capacity, and research in psychopathology has consistently reported that social cognitive abilities covary with a broad range of diagnoses and dimensions—including Autism Spectrum Quotient, Psychoticism, and Antagonism (Baron-Cohen et al., 1986; Bora & Pantelis, 2013; Dolan & Fullam, 2004; Krueger & Markon, 2014; Preißler et al., 2010). Better ToM is also associated with functional outcomes such as increased social competence (Liddle & Nettle, 2006; Jensen-Campbell et al., 2002) and reduced antisocial behavior (Mohr et al., 2007; Meier et al., 2006). Given that the field of personality psychology seeks to provide comprehensive taxonomies and explanatory models for understanding individual differences in human cognition and behavior, looking to existing models of normal-range personality variation provides one promising avenue for better understanding ToM and its potential correlates.

Social Cognition and Personality

The *Big Five* personality traits capture five broad dimensions of personality that can comprehensively organize most personality traits and descriptors (John et al., 2008). Each of the Big Five has been linked to individual differences in specific psychological processes and associated self-regulatory functions (DeYoung, 2015; DeYoung & Blain, 2020). For instance, Extraversion is related to reward responsivity and Openness/Intellect is related to information processing and pattern sensitivity (Blain, Grazioplene et al., 2020; Blain, Longenecker et al., 2020; Blain, Sassenberg et al., 2020; DeYoung et al., 2012; 2014; Lucas et al., 2000; Smillie et al., 2012; 2019). ToM and social cognition appear most related to the Big Five domain of Agreeableness. Agreeableness has been associated with many of the same functional outcomes as ToM, including social competence and social network size (Liddle & Nettle, 2006; Jensen-Campbell et al., 2002; Meier et al., 2006). Research directly examining the relation between Agreeableness and ToM has also begun to emerge, and there is some evidence for a positive correlation between the two constructs, such as findings presented by Nettle & Liddle (2008). Further insight into the individual difference correlates of ToM has been gained by simultaneously examining its association with normal-range and pathological personality traits.

The Big Five dimensions are very similar to the dimensions that emerge from patterns of covariation in symptoms of psychopathology—including not only personality disorders but other disorders too (DeYoung & Krueger, 2018; Kotov et al., 2010; Kotov et al., 2017). Many psychiatric symptoms can be described as risky or maladaptive variants of behaviors described by normal personality variation (DeYoung & Krueger, 2018). For instance, maladaptively low Agreeableness has been labeled “Antagonism” (Gore & Widiger, 2013; Suzuki et al., 2015). Individual differences in ToM ability have been linked to both Agreeableness and Antagonism, including factors representing the shared variance of Agreeableness and Antagonism scales (Allen et al., 2017; Nettle & Liddle, 2008). Importantly, Agreeableness and Antagonism, like all domain-level traits, can be broken down into a variety of lower-order traits, which are typically labeled aspect and facets (DeYoung et al., 2007). For instance, Agreeableness can be decomposed into the two aspects of *Compassion* and *Politeness*, which describe peoples’ ability to empathize with others and peoples’ ability to control socially unacceptable behaviors (e.g., expressions of aggression), respectively (DeYoung et al., 2007). Facets of Antagonism include callousness, aggression, and manipulativeness (Gore & Widiger, 2013; Krueger et al., 2013).

Allen et al. (2017) directly examined differential associations between ToM ability and lower-order subfactors of the Agreeableness-Antagonism domain. Allen et al. (2017) found that ToM was positively correlated with Compassion and negatively correlated with Politeness. Subsequent analyses of multiple Agreeableness and Antagonism facet-level scales suggested that Politeness could be subdivided into two subfactors that differentially predicted ToM ability. Like the Compassion aspect, a Non-aggression or Pacifism subfactor positively predicted ToM, but an Honesty subfactor negatively predicted ToM. Nonetheless, these findings regarding subfactors were discovered in *post hoc* analyses and warrant replication using a confirmatory framework and additional measures of ToM ability. In the current research, we attempted to replicate and extend the work of Allen et al. (2017), further examining whether ToM ability was related to these three Agreeableness-Antagonism subfactors of Compassion-Callousness, Honesty-Manipulativeness, and Pacifism-Aggression.

Reliability, Multi-task Designs, and Latent Variable Modeling

To justify claims regarding the underlying associations among constructs—for example, Agreeableness and social cognition—we must first be able to assess each of those constructs, individually, in a way that is reliable and valid. Concerns of reliability and validity are especially important when using behavioral tasks, as even tasks that can detect robust effects at the group level (e.g., tests of implicit bias or self-regulation) often fail to produce reliable measurement of individual differences (Hedge et al., 2018; Enkavi et al., 2019a; Schnabel et al., 2008). Fortunately, questionnaire measures of personality and tests of general or social cognitive ability tend to have better reliabilities than many of the measures commonly used in other areas of psychology (Hedge et al., 2018; Morrison et al., 2019; Pinkham et al., 2018; Vellante et al., 2013). Regardless of their reliability, however, single-task performance-based indicators are often limited in their scope and measure constructs more narrowly than those they purport to represent (Apperly, 2012; Blain, Longenecker et al., 2020). Performance on any given task is influenced by a number of task-specific factors but using multi-indicator designs can allow us to move toward measuring constructs more reliably as what is shared across multiple tasks, thereby avoiding underestimation of true effect sizes (Blain, Longenecker et al., 2020; Campbell & Fiske, 1959; Eisenberg et al., 2019; Enkavi et al., 2019a; 2019b; Nosek & Smythe, 2007).

We can further increase our ability to reliably measure constructs such as ToM ability and estimate their associations with other variables by using latent variable methods, such as structural equation modeling (SEM), which models the prediction of latent variables by other latent variables. Latent variables represent the shared variance of multiple measured (or *manifest*) variables (Schumacker & Lomax, 2004). For example, a latent social cognitive ability variable might be modeled as the shared variance of accuracy scores across different social cognition tasks. Assessing variables of interest at the latent level allows for more robust conclusions, as latent variables capture only the shared variance of their indicators, thereby eliminating error variance and more accurately capturing variability in the underlying constructs of interest (Keith, 2006). Modeling social cognitive ability as the shared variance in performance across tasks should give a better representation of true variance in social cognitive ability by factoring out unique task variance (which includes a combination of task-specific variance and error).

The Current Study

We hoped to replicate and extend the findings from Nettle & Liddle (2008) and Allen et al. (2017) by analyzing the relation of ToM ability and individuals' trait levels along the Agreeableness-Antagonism continuum. We further break down the personality hierarchy to explore how Agreeableness-Antagonism subfactors (i.e., Compassion-Callousness, Pacifism-Aggression, and Honesty-Manipulativeness) relate to ToM.

First, we hypothesized that accuracy scores from multiple tests of ToM ability would be positively correlated with one another and would map onto a single latent factor, producing a well-fitting measurement model. Mirroring the pattern of findings from Allen et al. (2017), we also hypothesized that when computing an oblique, three-factor solution while factor analyzing multiple self-report measures of Agreeableness and Antagonism, dimensions would emerge that corresponded to Compassion-Callousness, Honesty-Manipulativeness, and Pacifism-Aggression. We hypothesized that ToM accuracy (modeled as a latent variable and as scores from individual tasks) would positively correlate with subfactors for Compassion and Pacifism and negatively correlate with Honesty. In an effort to more clearly replicate previous research, we also included models predicting performance on the individual tasks, in addition to our model using a latent variable for ToM accuracy.

Although the work of Allen et al. (2017) showed support for three subfactors of Agreeableness-Antagonism that differentially predict ToM abilities, their analyses were *post hoc* and warrant replication. Further, our study utilizes a broad battery of Agreeableness-Antagonism facet scales, a multitask design, and an exploratory structural equation modeling (ESEM) analytical approach—all of which are advantages over previous work done on this topic.

Method

Participants

Participants were recruited via a combination of Qualtrics panels and from the campus of a large public research university in the Midwestern US as part of a study on social cognition, personality, and psychopathology. No explicit exclusion criteria for psychopathology were implemented in an attempt to sample a broad, representative range of pathological and normal personality variation from the general population. The original sample consisted of 389 individuals, but 54 individuals were excluded for having high amounts of random or invalid responding, leaving a total valid sample of 335. Participants ranged from 18 to 75 in age (M = 26.4, SD = 13.6). There were 267 females (79.7%), 67 males (20%), and 1 intersex individual (0.3%). In terms of race/ethnicity, 235 participants identified as White or Caucasian (70.1%), 59 as Asian or Pacific Islander (17.6%), 7 as Black or African American (2.1%), 4 as Latino or Hispanic (1.2%), and 30 as multiracial or other (9.0%). 283 participants were native English speakers (84.5%).

Participants reviewed an online informed consent document before beginning the study, then completed an online battery of questionnaires and behavioral tasks, including self-report measures of demographics, personality, psychopathology, and social functioning and several tests of social cognition. All protocols were approved by the Institutional Review Board.

Self-Report Measures

Big Five Aspect Scales

The Big Five Aspect Scales (BFAS; DeYoung et al., 2007) is a 100-item questionnaire that assesses the Big Five personality domains, including two component aspects for each of the Big Five. Each aspect is measured by a total of ten items, including a combination of standard and reverse-coded items. Participants answered each question using a five-point Likert scale ranging from 1 ("Strongly disagree") to 5 ("Strongly agree"). The current study used the two Agreeableness aspects scales, measuring Compassion and Politeness.

Computer Adaptive Test of Personality Disorders: Static Form

The Computer Adaptive Test of Personality Disorders: Static Form (CAT-PD SF; Simms et al., 2011; Wright & Simms, 2014), a selection of 212 items from the 1366-item CAT-PD, is a measure that assesses 33 maladaptive personality traits (e.g., Callousness and Manipulativeness) that can be grouped into five broad categories similar to the Big Five (i.e., Negative Emotionality, Detachment, Antagonism, Disconstraint, and Psychoticism). Participants rated items on a 5-point scale ranging from 1 ("very untrue of me") to 5 ("very true of me"). The current study used the Antagonism-related scales of Callousness, Manipulativeness, Hostile Aggression, and Domineering.

Externalizing Spectrum Inventory Brief Form

The Externalizing Spectrum Inventory Brief Form (ESI-BF; Patrick et al., 2013) is a shortened version of the 415-item ESI. This 160-item questionnaire assesses general Disinhibition, Substance Abuse, Callous Aggression, and 23 lower-order facets of the externalizing spectrum. Participants rated each item on a 4-point scale, with higher scores corresponding to greater agreement with the item. The current study used the lower-order facet scales most strongly correlated with Antagonism and Agreeableness, including Physical Aggression, Relational Aggression, Destructive Aggression, Fraud, Theft, Empathy, and Honesty.

Interpersonal Reactivity Index

The Interpersonal Reactivity Index (IRI; Davis et al., 1980) is a 28-item questionnaire that assesses individual differences in empathy across four dimensions—Empathic Concern, Fantasy, Personal Distress, and Perspective Taking. Each dimension is measured by a 7-item subscale, and participants rated items on a 5-point Likert scale ranging from 0 ("does not describe me well”) to 4 (“describes me very well”), with a combination of standard and reverse-scored items. In the current study, we only used the Empathic Concern scale (i.e., other-oriented emotions centered on helping people in need) because it is the IRI dimension most strongly associated with Agreeableness (Melchers et al., 2016).

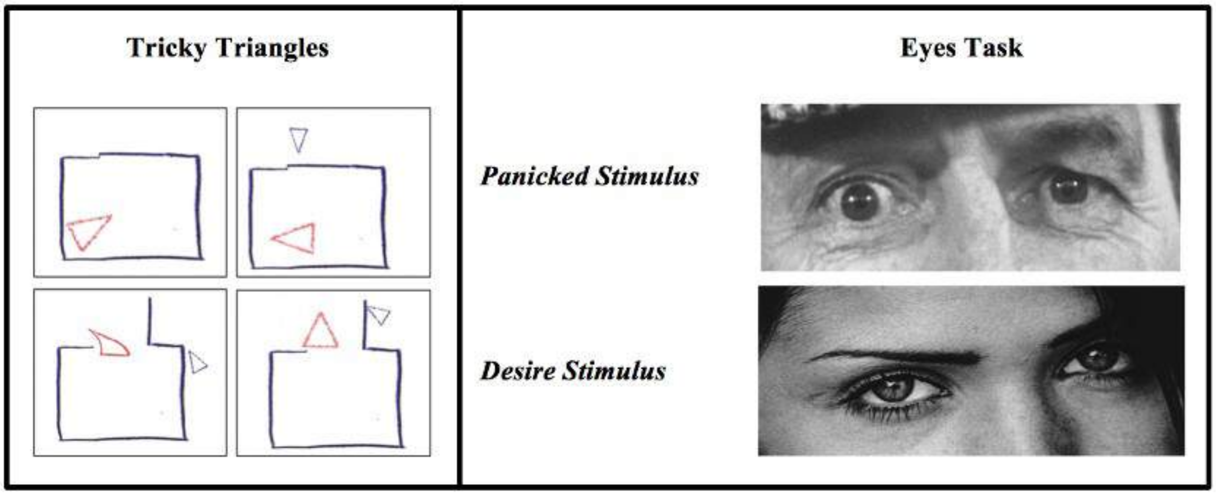
Personality Inventory for DSM-5

The Personality Inventory for DSM-5 (PID-5; Krueger et al., 2012) is a 220-item questionnaire that assesses 25 maladaptive personality trait facets that, like the CAT-PD, can be grouped into five categories (i.e., Antagonism, Detachment, Disinhibition, Negative Affect, and Psychoticism). Facet scales range from 4 to 14 items and are rated on a 4-point scale ranging from 1 (“very false or often false”) to 4 (“very true or often true”). The present study used three of the facets belonging to the Antagonism domain, including Callousness, Deceitfulness, and Manipulativeness.

Theory of Mind Tasks

Participants also completed three tests of social cognition, including a triangle animation task in which participants labeled animations as random, physical, or social (Abell et al., 2000), a mentalizing stories task in which participants had to answer questions about characters’ factual and social knowledge (Stiller & Dunbar, 2007), and a perceptual ToM task using eye stimuli (Baron-Cohen et al., 1997). Example stimuli from the triangles and eyes tasks are pictured in Figure 1.

Figure 1.



Note. Social cognition tasks

Theory of Mind Vignettes

The ToM vignette task (Stiller & Dunbar, 2007) comprises five short stories depicting social situations. Each story describes a social interaction involving multiple characters. Participants read each story, after which they answered five ToM questions and five 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 from two to six in each question. Performance on memory questions within the task can be used as a covariate to ensure that any associations with variables of interest are due to participants' ToM ability rather than their memory for the details of the story. ToM questions required 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. For example, in second-level questions, participants tracked their own mental state and the mental state of one character (e.g., "John wanted to go home after work"). In fourth-level questions, participants tracked the mental state of three characters (e.g., "John thought that Penny knew what Sheila wanted to do”). To assess performance on the task, we adopted the procedure used by Nettle and Liddle (2008) and Allen et al. (2017) and computed simple sums of correct responses to memory questions and ToM questions for each participant.

Tricky Triangles Task

In the triangles task (Abell et al., 2000), participants are presented with a series of computerized animations of shapes interacting in a way that was random, physical, or social. In the random condition, the shapes did not interact with each other, but rather moved around purposelessly (e.g., bouncing or drifting). In the physical condition, the shapes moved in a goal-directed manner without invoking ToM or mentalizing (e.g., fishing or swimming). In the social condition, shapes enacted a social sequence, such as coaxing, seducing, or mocking. Participants were tasked with indicating whether each animation was random, physical, or social in nature, then scored for their accuracy in correctly categorizing each animation in a series of 22 clips.

Reading the Mind in the Eyes Task

The eyes task (Baron-Cohen et al., 2001) consists of 36 grey-scale photos of people taken from magazines. These photos were cropped and rescaled so that only the area around the eyes could be seen. Each photo was accompanied by four mental state terms, from which the participant was instructed to choose the word that best described what the person in the photo was thinking or feeling. Only one of the four items was correct (as judged by consensus from an independent panel of judges in the initial psychometric study). Participants were scored for their accuracy across all 36 stimuli.

Analyses

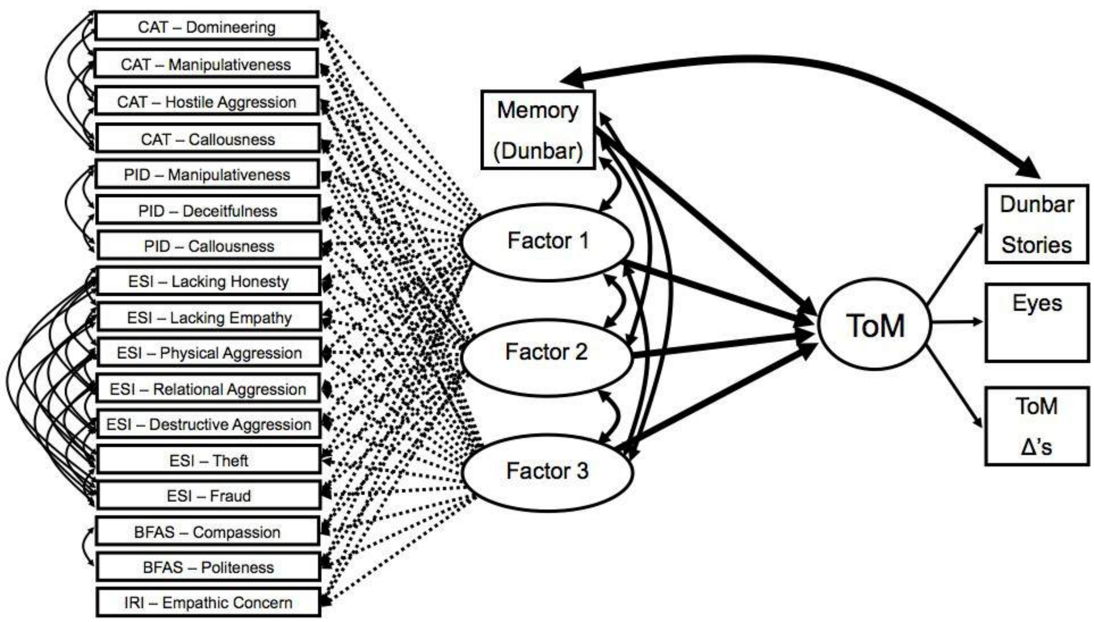
Descriptive statistics were calculated for all task performance and personality variables. Cronbach’s α (Cronbach, 1951) and ω_t (McDonald, 1999; Revelle & Condon, 2019) were computed to assess internal consistency reliability. *Mplus* was then used for latent variable modeling (Muthén & Muthén, 2017); all models were estimated using full-information, robust maximum likelihood estimation (MLR). First, we computed a single-factor confirmatory factor analytic (CFA) model using accuracy scores from the three ToM tasks to examine how well these tasks represent a single coherent latent variable. To further assess whether a single factor model was appropriate for explaining shared variance across tasks, ω_t was also computed.

Next, Exploratory Structural Equation Modeling (ESEM) was implemented. Exploratory Agreeableness-Antagonism subfactors were derived from relevant BFAS, IRI, ESI-BF, PID-5, and CAT-PD SF subscales. First, we conducted a Velicer's minimum average partial (MAP) test (O'Connor, 2000) to see how many factors were empirically suggested in this data. Then, a three-factor solution was computed, given our hypotheses and aims to conceptually replicate and extend the work of Allen et al. (2017). We used factoring with a constrained oblimin rotation ($\gamma = -.80$). Relative to traditional structural equation modeling (SEM) with confirmatory factors, computing exploratory factors better accounts for the nontrivial cross-loadings of indicators (Asparouhov & Muthén, 2009). ESEM also allows for more accurate model estimation vs. simple use of observed factor score estimates in SEM. Nonetheless, results for models in the current study were substantively equal if factor score estimates derived using exploratory factor analysis followed by the regression method were used (rather than full ESEM).

Residual covariances of subscales from the same questionnaire (e.g., BFAS Politeness and Compassion) were freely estimated, which resulted in better fit vs. constraining these covariances to zero (ΔS-B $\chi^2 = 218.3$, $p < .001$). Additionally, the residual covariance between memory accuracy and ToM accuracy from the vignette task was freely estimated, resulting in significantly improved fit (ΔS-B $\chi^2 = 23.4$, $p < .001$).

Subsequently, we predicted latent ToM accuracy (shared variance of accuracy across the three tasks) from the Agreeableness-Antagonism factors, allowing predictors to correlate and including performance on the memory questions from the vignette task as an additional, correlated predictor variable. The full ESEM model is presented in Figure 2.

Figure 2.



Note. Exploratory structural equation model of Agreeableness-Antagonism factors and ToM

Satorra-Bentler adjusted fit indices were computed and 95% confidence intervals (with standard errors derived using the Huber-White sandwich estimator) were estimated for the path coefficients from predictor variables to latent ToM accuracy (Huber, 1967; Muthén & Muthén, 2017; Satorra, & Bentler, 2001; White, 1980). Finally, for visualization purposes, factor score estimates were computed for Agreeableness-Antagonism and ToM latent variables, using the regression method. ToM factor scores were residualized for scores on the memory condition of the vignette task and standardized, then plotted against standardized factor scores for each of the Agreeableness-Antagonism factors (residualizing for variance in the other two Agreeableness-Antagonism factors).

In a final model, which sought to more directly replicate previous results, we examined the effects of Agreeableness-Antagonism factors on observed accuracy scores for each of the individual ToM tasks (controlling for performance on the memory conditioning from the story task). These models are presented in our supplement. Results from models that did not freely estimate residual covariances of variables from the same task/questionnaire are reported in our supplemental methods and results.

Results

Descriptive statistics are presented in Table A1. Performance was generally high for the behavioral tasks, but variables showed no prominent ceiling effects. Several of the personality and task performance variables showed moderate skewness; thus, analytical methods robust to non-normality were used (i.e., MLR estimation implemented with MPLUS). Values of ω_t and α indicated acceptable internal consistency for the majority of questionnaire and task variables.

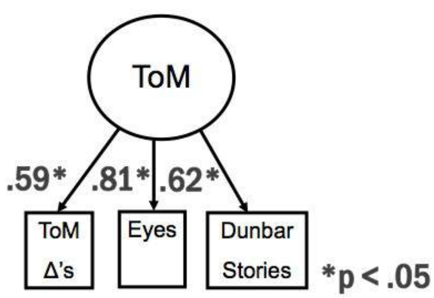
Table 1. Descriptive statistics

	Mean (SD)	Skewness	[Minimum, Maximum]	α	ω_t
Dunbar ToM	19.4 (3.2)	-1.1	[8, 25]	.66	.77
Dunbar Memory	22.5 (3.7)	-1.6	[10, 26]	.81	.85
Eyes Task	24.3 (5.7)	-0.8	[5, 34]	.77	.80
Triangles	12.6 (3.3)	-0.5	[3, 19]	.62	.68
BFAS Compassion	4.0 (0.6)	-0.4	[2.4, 5.0]	.86	.86
BFAS Politeness	3.8 (0.6)	-0.6	[2.1, 5.0]	.77	.77
IRI Empathic Concern	3.8 (0.7)	-0.2	[1.9, 5.0]	.77	.78
PID Callousness	1.4 (0.5)	2.0	[1.0, 3.8]	.91	.94
PID Deceitfulness	1.6 (0.6)	1.0	[1.0, 3.7]	.88	.89
PID Manipulativeness	1.8 (0.8)	0.6	[1.0, 4.0]	.84	.85
CAT Callousness	1.8 (0.8)	1.1	[1.0, 4.4]	.91	.91

CAT Domineering	2.1 (0.8)	0.7	[1.0, 5.0]	.85	.85
CAT Hostile Aggression	1.6 (0.8)	1.7	[1.0, 5.0]	.93	.93
CAT Manipulativeness	1.7 (0.7)	1.2	[1.0, 4.3]	.87	.87
ESI Theft	1.3 (0.7)	2.1	[1.0, 4.0]	.90	.91
ESI Fraud	1.3 (0.7)	2.2	[1.0, 4.0]	.89	.90
ESI Honesty	1.9 (0.6)	0.4	[1.0, 4.0]	.76	.77
ESI Physical Aggression	1.4 (0.6)	2.1	[1.0, 4.0]	.91	.92
ESI Destructive Aggression	1.3 (0.5)	2.6	[1.0, 4.0]	.93	.93
ESI Relational Aggression	1.4 (0.6)	1.6	[1.0, 4.0]	.89	.89
ESI Empathy	1.6 (0.5)	0.9	[1.0, 3.2]	.88	.89

Bivariate correlations are presented in Table A1. Across measures, Agreeableness variables positively predicted task performance while Antagonism variables negatively predicted task performance. The CFA model showed that accuracy scores on the three ToM tasks loaded onto a single latent variable (Figure 3). Factor loadings were moderately high, with performance on the eyes task being the strongest. Fit statistics for all models are presented in Table 2. The CFA model was just identified, so meaningful model fit evaluation based on standard fit indices was not possible. Nonetheless, there was evidence that a substantial portion of variance could be explained by a single underlying social cognitive ability factor ($\omega_t = .71$).

Figure 3.



Note. Model of ToM tasks

Table 2. Fit statistics for structural equation models

Models	RMSEA	95% C.I.	SRMR	S-B $\chi 2$	*p*	CFI	TLI
1. ToM	.000	[.000, .000]	.000	0.00	< .001	1.0	1.0
2. ToM, Mem, and Antagonism Factors	.060	[.050, .069]	.023	262.4	< .001	.98	.96
3. Individual Tasks, Mem, and Antagonism Factors	.061	[.051, .071]	.023	255.1	< .001	.98	.96
S1. ToM, Mem, and Antagonism Factors (constrained residual covariances)	.088	[.081, .097]	.042	550.7	< .001	.94	.91
S2. Individual Tasks, Mem, and Antagonism Factors (constrained residual covariances)	.087	[.078, .095]	.028	506.0	< .001	.94	.91

Next, we used ESEM to identify latent factors from Agreeableness and Antagonism scales. Consistent with the notion that Agreeableness can be separated into two correlated aspects and with previous work using a similar set of Agreeableness-Antagonism scales (Allen et al., 2017), conducting a Velicer's MAP test indicated the presence of two factors across the 17 scales. Correlations between these two extracted factors and relevant BFAS variables showed that the first factor approximated low Politeness ($r = -.66$, $p < .001$) and the second factor strongly resembled Compassion ($r = .79$, $p < .001$).

Since we were interested in parsing subfactors within Politeness and replicating the discriminant validity of these two factors in predicting ToM (Allen et al., 2017), we then chose to extract three factors. Factor loadings for the three-factor solution are presented in Table 3.

Table 3. Factor loadings of Agreeableness-Antagonism scales on three exploratory factors

Scale	Aggression	Manipulativeness	Compassion
CAT – Domineering	.21	**.57***	-.18*
CAT – Manipulativeness	**.47***	.46*	-.14*

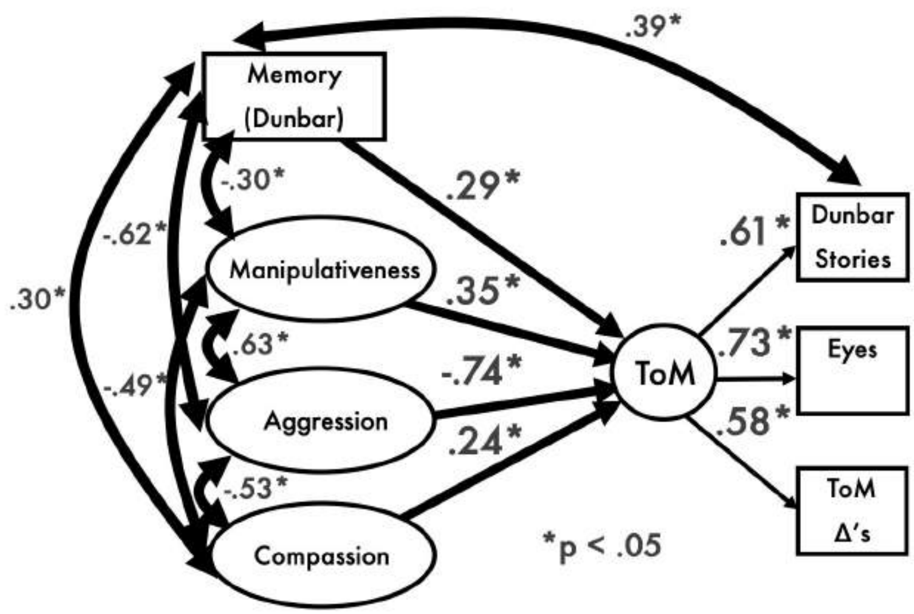
PID – Manipulativeness	.24	**.60***	-.02
PID – Deceitfulness	.28	**.61***	-.12*
ESI – Honesty	-.02	**-.39***	.16
BFAS – Politeness	-.07	**-.51***	.26*
ESI – Relational Aggression	**.53***	.39*	-.10*
ESI – Physical Aggression	**.82***	.03	-.03
ESI – Destructive Aggression	**.90***	-.03	-.03
ESI – Theft	**.78***	.13	.05
ESI – Fraud	**.70***	.20	-.03
CAT – Hostile Aggression	**.82***	.05	.14*
PID – Callousness	**.66***	.06	.31*
CAT – Callousness	.40*	.16*	**-.52***
ESI – Empathy	-.29*	-.04	**.70***
BFAS – Compassion	-.02	-.08	**.78***
IRI – Empathic Concern	-.05	-.10	**.71***

Note. *$p < .05$ (based on the z-distribution and standard errors computed using the Huber-White sandwich estimator); bolded values indicate which factor the scale had the largest loading for.

Based on their content and on previous research, we labeled these three factors Pacifism-Aggression, Honesty-Manipulativeness, and Compassion-Callousness. The first factor corresponded to Pacifism-Aggression, having strong positive loadings for the ESI aggression subscales and various relevant scales from the CAT-PD SF and PID-5. The second factor corresponded to Honesty-Manipulativeness, showing the strongest positive loadings for PID-5 Deceitfulness and Manipulativeness, CAT-PD Domineering and Manipulativeness, and Relational Aggression, as well as negative loadings for BFAS Politeness and ESI Honesty. A final factor corresponded to Compassion-Callousness, with strong positive loadings for the BFAS Compassion, IRI empathic concern, and ESI Empathy scales, as well as negative loadings for the CAT-PD SF and PID-5 Callousness scales. Significant cross-loadings were relatively common, across all factors.

Subsequently, we examined the effects of the Aggression, Callousness, and Manipulativeness factors on ToM accuracy across tasks. Results of the full structural model and ToM measurement model are displayed in Figure 4. (The full measurement model for Agreeableness-Antagonism factors is not displayed here, due to visual complexity). Residual correlations accounting for shared instrument variance are presented in Table A2.

Figure 4.

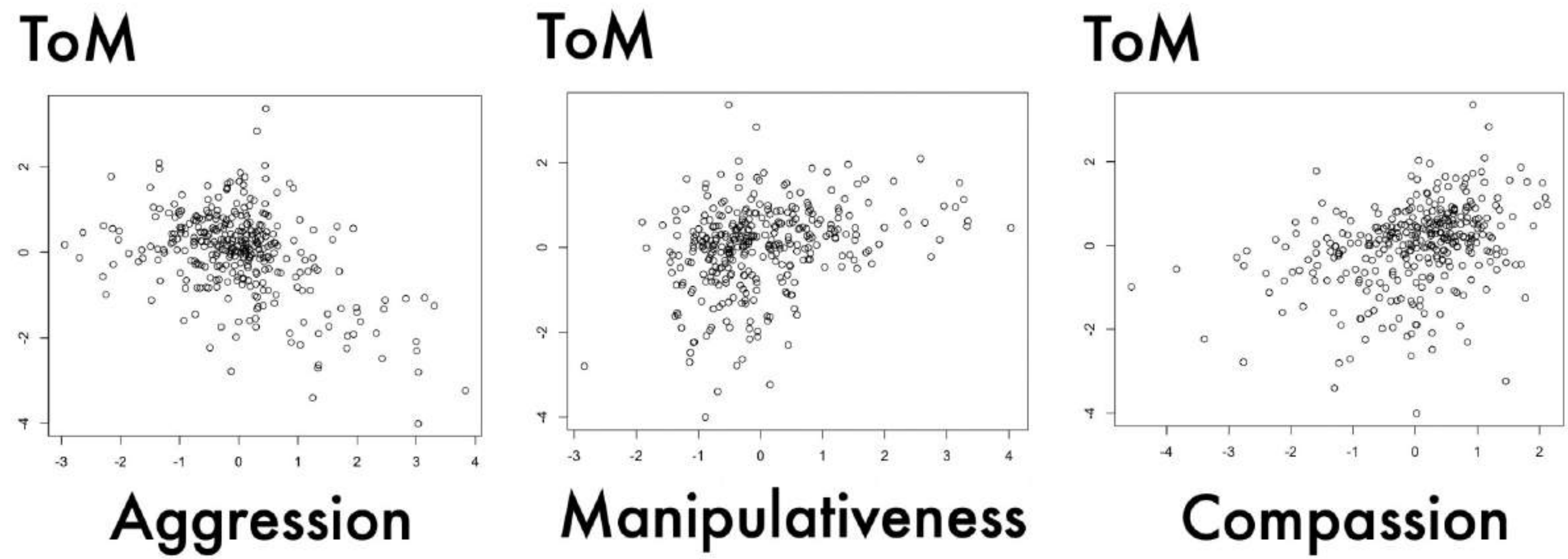


Note. Model of Agreeableness-Antagonism factors and ToM

We found that ToM was negatively predicted by Aggression (95% CI β: [-.96, -.52]) and positively by Compassion (95% CI β: [.08, .41]) as well as Manipulativeness (95% CI β: [.17., .52]). Memory was a positive predictor of ToM (95% CI β: [.12, .46]). Model fit was acceptable (when allowing correlated error terms for scales within each given questionnaire), as indicated by SRMR and RMSEA ≤ .08 and TLI/CFI ≥ .95 (Hooper et al., 2008; Hu & Bentler, 1999). Similar results were obtained whether or not memory was included as a covariate.

Associations among residualized factor score estimates are visualized in Figure 5.

Figure 5.



Note. Scatterplots of ToM and Agreeableness-Antagonism factor score estimates

A final model examined the effects of the Aggression, Manipulativeness, and Compassion factors on observed accuracy scores for the three ToM tasks, controlling for memory performance on the stories task. Factor loadings for Agreeableness-Antagonism scales onto their corresponding factors were nearly equivalent to those obtained in our latent ToM model and are fully reported in the supplement (Table S1). Path coefficients and 95% confidence intervals for Agreeableness-Antagonism factors predicting individual ToM task accuracy scores are presented in Table S2. Effects were all in the same direction of those from the model predicting latent ToM, but their magnitudes were smaller and inconsistently significant; it is possible these individual-task effects were not statistically significant at the $\alpha = .05$ threshold due to sampling variability or the lower reliability of using single tasks as criterion variables.

For both our latent ToM and individual-task models, model specifications that did not freely estimate residual covariances (of subscales taken from the same questionnaire and of memory and ToM accuracy scores from the stories task) yielded results that were substantively equivalent to those reported here. Fit of these models was marginal, as indicated by SRMR $\leq .08$, RMSEA between .08 and .10, and CFI/TLI between .90 and .95 (Hu & Bentler, 1999). Full results from these models are included in our supplement (Figure S1 and Tables S3–S5).

Discussion

The current study investigated how subfactors of the Agreeableness-Antagonism dimension are associated with individual differences in ToM ability. Specifically, we sought to replicate and extend the work of Allen et al. (2017) using multiple behavioral tasks, a more extensive assessment of normal-range and pathological personality traits, and an ESEM approach. Using ESEM, we derived a three-factor structure for a selection of Agreeableness-Antagonism scales. In line with theory and previous work, we labeled these factors Compassion-

Callousness, Pacifism-Aggression, and Honesty-Manipulativeness. Results showed that while higher Compassion and lower Aggression predicted higher ToM ability, higher Manipulativeness also predicted higher ToM ability. These findings replicate the post hoc analyses of Allen et al. (2017) and demonstrate that the effects generalize to multiple behavioral measures of social cognition.

The current findings also extend previous work by Nettle and Liddle (2009), who found that Agreeableness was positively associated with ToM as assessed by the mentalizing vignettes task, but *not* the eyes task. Based on this pattern of results, they concluded that Agreeableness is positively related to "social-cognitive" ToM (understanding mental states based on language and reasoning), but not "social-perceptual" ToM (understanding mental states based on visual or other sensory clues). In contrast, our results indicate that Agreeableness-Antagonism is positively associated with the shared variance *across* ToM tasks, including both the mentalizing vignettes and the eyes task. This suggests that the null finding for Agreeableness and the eyes task reported by Nettle and Liddle may have been due to the use of observed dependent variables, which confound shared variance related to ToM abilities with task-specific unique variance and measurement error. Thus, it seems utilization of latent variable modeling using multiple tasks or multilevel modeling of task data (e.g., Rouder & Haaf, 2019) may be particularly useful when evaluating associations between social cognitive ability and related traits or outcomes. Similar to Allen et al. (2017), we also found that associations between ToM and Agreeableness were specific to certain subfactors; this suggests Nettle and Liddle's finding of an association between ToM and global Agreeableness may have arisen from using a measure highly weighted toward the Compassion aspect. In addition to successfully replicating and extending previous work, the current study extends our understanding of Agreeableness-Antagonism and has broader implications for personality and individual differences research.

Understanding Agreeableness and Antagonism

It has now been shown in two independent samples that individual differences in social cognitive abilities are positively associated with Compassion and Pacifism but *negatively* related to Honesty, suggesting that better ToM abilities may enable individuals to be more compassionate and less aggressive, but also more deceitful and manipulative. As there is a robust history of research relating Agreeableness—and particularly its Compassion aspect—to questionnaire measures of trait empathy (del Barrio et al., 2004; Graziano et al., 2007; Mooradian et al., 2008; 2011; Penner et al., 1995), it is unsurprising that the compassion subfactor was positively related to ToM ability. The Compassion-Callousness subfactor in the current study showed strong positive loadings for IRI Empathic Concern, BFAS Compassion and ESI Empathy, as well as a strong negative loading for CAT-PD Callousness. Our finding of a positive association between the compassion factor and ToM ability, paired with the previous work of Allen et al. (2017) and Nettle and Liddle (2009), complements research showing that ToM, trait empathy, and Compassion are associated with positive real-world social outcomes (Cassidy et al., 2003; Devine et al., 2016; Stiller & Dunbar, 2007; Sun et al., 2017). Findings are also in line with a long tradition of research documenting social cognitive deficits in those with elevated psychopathy and children with callous-unemotional traits (Jones et al., 2010; Shamay-Tsoory et al., 2010; Pardini et al., 2003). Thus, it is possible that individual differences in the ability to accurately perceive, decipher, and react to the mental states of others is one core psychological mechanism underlying Compassion and prosocial behavior, as well as callousness and antisocial behavior. Future work should further explore the relations between social cognition and Compassion by incorporating a broader array of behavioral tasks that tap into additional components of social cognitive ability: for instance, tasks capturing additional components of empathy such as emotion contagion or affective resonance (Zaki et al., 2008). It is possible that performance on these tasks would be even more strongly related to Compassion than tasks like those in the current study, which primarily assess cognitive and perceptual aspects of social processing.

In addition to showing a positive association with Compassion, ToM ability was also positively associated with individual differences in Pacifism (or lower scores on the Aggression subfactor). The Pacifism-Aggression subfactor in the current study showed strong positive loadings for several ESI subscales related to aggression, as well as the CAT-PD Hostile Aggression scale. Counterintuitively, the factor loading of PID-5 Callousness was actually higher for the Pacifism-Aggression factor than the Compassion-Callousness factor; this is likely due to the large number of aggression items in that subscale and is consistent with the fact that the current PID-5 Callousness scale was originally designed to measure two separate facets (Callousness and Aggression) but was eventually collapsed into a single scale based on analyses using item response theory (DeYoung et al., 2016; Krueger et al., 2012). Somewhat unexpectedly, the Politeness scale from the BFAS did not significantly load onto the Aggression subfactor (but instead only loaded on the Manipulativeness subfactor). The finding of a negative association between Aggression and ToM result replicates a *post hoc* finding from Allen et al. (2017) and is consistent with other work documenting negative associations between Aggression and ToM ability (Meier et al., 2006; Mohr et al., 2007). These results suggest that facets of Agreeableness other than just Compassion (and trait empathy) are positively associated with individual differences in ToM ability. Future work could better determine to what extent the specific psychological mechanisms linking lower aggression to better ToM ability are similar to or different from those underlying the association between ToM and Compassion.

The only Agreeableness-Antagonism subfactor showing a diverging pattern of association with ToM ability was Honesty-Manipulativeness. The Manipulativeness subfactor was marked by positive loadings for PID-5 Manipulativeness and Deceitfulness and CAT-PD Manipulativeness and Domineering, as well as negative loadings for BFAS Politeness and ESI Honesty. Paired with the original finding in Allen et al. (2017), the current replication of a positive association between Manipulativeness and ToM provides further empirical support for the longstanding notion that being able to successfully persuade, manipulate, and deceive others is partially reliant on the ability to understand the thoughts and emotions of others (Byrne, 1996; Byrne & Whiten, 1994; Ding et al., 2015; Lonigro et al., 2014; Slaughter et al., 2013; Talwar et al., 2007). Findings are also directly in line with some previous work examining social cognition in the dark triad framework—focusing on the traits of Narcissism, Machiavellianism, and Psychopathy (Furnham et al., 2013; Paulhus & Williams, 2002). Research on these traits and their relation to social cognition is mixed, but tends to suggest that while psychopathy is negatively associated ToM abilities, Machiavellianism and Narcissism (characterized by entitlement, grandiosity, and attention seeking) may be unrelated to or even positively associated with individual differences in ToM (Jonason & Krause, 2013; Kajonius & Björkman, 2020; Paal & Bereczkei, 2007; Schimmenti et al., 2019; Stellwagen & Kerig, 2013; Vonk et al., 2015; Wai & Tiliopoulos, 2012). Taken with previous work and theory, the current findings seem to suggest that specific traits often conceptualized as pathological (i.e., dishonesty and manipulativeness) may actually be associated with enhanced abilities and outcomes in some contexts. Indeed, this is consistent with how many researchers have discussed potential adaptive qualities of

Machiavellianism, which can be particularly apparent in high-power occupations such as attorneys, executives, and salespeople (Byrne & Whiten, 1994; Furnham & Treglown, 2021; Grover & Furnham, 2021). Nonetheless, whether being able to successfully deceive others leads to positive functional outcomes at the individual and group level will typically depend on situational context and one's other trait levels. Indeed, those with the highest levels of ToM ability in the current dataset have both high Manipulativeness and Compassion, suggesting these individuals may use deceit and manipulation for prosocial means.

It is worth noting that the current results contrast with a few studies that found a negative association or no association between ToM ability and Machiavellianism (e.g., Ali & Chamorro-Premuzic, 2010; Lyons et al., 2010). To explain these associations, some have suggested those with elevated dark triad traits possess the capacity and ability to engage in ToM-related processes, but lack the disposition—and, in most situations, the motivation—to do so (Kajonius & Björkman, 2020). Nonetheless, scales used to assess Machiavellianism in many of these studies from the dark triad perspective tend to conflate dishonesty and manipulation with other facets of Antagonism, such as immorality or mistrust, indicating that the Honesty-Manipulativeness dimension as construed in the current work and by Allen at al. (2017) may, indeed, be associated with patterns of ToM that diverge from associations with the broader Agreeableness-Antagonism domain. Moreover, studies that do not control for subfactors predicting ToM abilities in opposite directions may have true effects in either direction masked due to statistical suppression (Martinez Gutierrez & Cribbie, 2021; Tzelgov & Henik, 1991).

Additional Implications for Individual Differences Research

Individual differences in social cognition are conceptually similar to several other constructs that exist in disparate lines of work—for instance emotional intelligence and

Gardner's interpersonal intelligence (Gardner, 2011; McEnrue & Groves, 2006). In recent years, emotional intelligence has gained widespread popularity as a construct in both popular culture and among researchers, with some arguing emotional intelligence is predictive of broad positive life outcomes and should be used to inform hiring decisions (Bar-On, 2001; Emanuel & Gudbranson, 2018; Fox & Spector, 2000; Goleman, 1996; Stein & Book, 2011; Watkin, 2000; van der Linden et al., 2010; 2012; 2017). In contrast, others have argued that measures of emotional intelligence (particularly those relying only on self-report) provide little incremental validity over general intelligence and Conscientiousness, when it comes to predicting important occupational and educational outcomes (Amelang & Steinmayr, 2006; Antonakis, 2004; Gottfredson, 1997; Landy, 2005; Ones et al., 2012; Schmidt et al., 2008; Thorndike & Stein, 1937; Van Rooy & Viswesvaran, 2004; Willoughby & Boutwell, 2018). Our current results add to a growing body of research suggesting that while performance-based measures of social cognition may be highly correlated with measures of general cognitive ability, they do indeed show incremental validity in predicting socially relevant personality traits and outcomes (Allen et al., 2017; Stiller & Dunbar, 2007).

The current study provides insights into the theory and measurement of social cognition and its association with Agreeableness-Antagonism subfactors. Because problems with ToM and related interpersonal outcomes are characteristic of multiple mental disorders and symptom domains, elucidating their association with normal-range personality traits and improving their measurement may eventually help facilitate more effective methods for assessment and treatment. Such an approach is in line with the Hierarchical Taxonomy of Psychopathology's conceptualization of psychiatric illness, the National Institute of Mental Health's (NIMH's) Research Domain Criteria (RDoC) initiative, and theories of psychopathology that emphasize continuity with normal personality variation and impairments in cybernetic functioning (DeYoung & Krueger, 2018; Insel et al., 2010; Kotov et al., 2017). Future work on this topic could incorporate additional tasks to span a range of social cognitive and interpersonal abilities, including more of those recommended by the NIMH's Workgroup on Tasks and Measures for RDoC (Barch et al., 2016). Furthermore, it remains to be seen whether the personality correlates of social cognitive abilities in populations with more extreme levels of Antagonism (i.e., criminal offenders or those diagnosed with antisocial or narcissistic personality disorder) would reflect the patterns observed in the general population. Finally, another topic worth exploring further is whether the personality correlates and mechanisms of social cognitive deficits are consistent or divergent across different psychopathology dimensions; in particular, research should explore whether ToM deficits associated with psychoticism and positive schizotypy or autistic traits and symptoms are similar in etiology and mechanisms to those associated with Antagonism and related personality disorders.

Limitations

Though the current study had numerous strengths, several limitations are worth noting. First, the current sample had an overrepresentation of females and people of European and Asian ancestry; future work should attempt to collect more demographically representative samples. The current study's measures of Agreeableness-Antagonism were self-reported and could be usefully supplemented in future research by peer reports or clinician ratings. Also, this work would have further benefited from the inclusion of a general intelligence measure such as the Wechsler Adult Intelligence Scale or International Cognitive Ability Resource (Condon & Revelle, 2014; Wechsler, 2008), as including such measures would allow us to better parse associations between personality and social cognition without the influence of general cognitive ability; such measures could also allow researchers to directly test how individual differences in general cognitive ability might be associated with Agreeableness-Antagonism and its constituent subfactors. These limitations could be usefully addressed by recruiting additional large samples with extensive, high-quality measures of personality, social cognition, general cognition, and real-world social functioning including peer-report and experience sampling data.

Additionally, although the current set of questionnaires captured a broad range of Agreeableness-Antagonism facets, future research could also incorporate measures from the HEXACO and Dark Triad literatures (Collison et al., 2018; Jonason & Webster, 2010; Jones & Paulhus, 2014; Lee & Ashton, 2004; Ashton & Lee, 2007). This approach could directly test whether the constructs represented in those measures (i.e., Honesty-Humility and Machiavellianism) map onto the Honesty-Manipulativeness dimension revealed in the current study using measures based on the Five Factor Model, and whether Honesty-Humility and Machiavellianism show corresponding associations with ToM. Future work could also intentionally recruit participants with clinically significant levels of Antagonism and other traits related to social cognitive deficits; this could be a particularly fruitful line of work if the tools used to assess social cognitive ability and personality domains such as Agreeableness could eventually be used to foster more effective early detection and intervention for forms of psychopathology related to interpersonal dysfunction.

Conclusion

Agreeableness-Antagonism is robustly related to life outcomes, including victimization, relationship satisfaction, aggression, and a variety of psychiatric disorders (Gore & Widiger, 2013; Lynam & Miller, 2019). Despite its enormous consequences, however, Agreeableness-Antagonism is arguably the least studied dimension of the Big Five and their pathological counterparts (Gore & Widiger, 2013; Lynam & Miller, 2019). The current research improves the scientific understanding of Agreeableness-Antagonism, replicating and extending work that suggests differential relations of Agreeableness-Antagonism subfactors with social cognition. Our findings suggest ToM abilities might facilitate individual differences in most traits related to Agreeableness, with a distinctly negative association with specific honesty-related tendencies. This paradox adds to a set of interesting similar patterns where the correlates or outcomes of personality traits diverge at levels below the Big Five, further underscoring the importance of facet-level research and parsing the subfactors of broad personality domains. Future research should more thoroughly examine situational, motivational, and relationship-specific factors to determine their potential moderating role in the associations between personality and social cognitive abilities.

Schizophrenia Bulletin vol. 46 no. 3 pp. 540–551, 2020
doi:10.1093/schbul/sbz103
Advance Access publication 11 October 2019

Toward a Neural Model of the Openness-Psychoticism Dimension: Functional Connectivity in the Default and Frontoparietal Control Networks

Scott D. Blain[1,*,], Rachael G. Grazioplene[2], Yizhou Ma[1], and Colin G. DeYoung[1]

[1]Department of Psychology, University of Minnesota Twin Cities, Minneapolis, MN; [2]Department of Psychology, Yale University, New Haven, CT

*To whom correspondence should be addressed; University of Minnesota Twin Cities, N616 Elliot Hall, 75 East River Parkway, Minneapolis, MN 55455; tel: (270) 287-8688, e-mail: blain058@umn.edu

Psychosis proneness has been linked to heightened Openness to Experience and to cognitive deficits. Openness and psychotic disorders are associated with the default and frontoparietal networks, and the latter network is also robustly associated with intelligence. We tested the hypothesis that functional connectivity of the default and frontoparietal networks is a neural correlate of the openness-psychoticism dimension. Participants in the Human Connectome Project (*N* = 1003) completed measures of psychoticism, openness, and intelligence. Resting state functional magnetic resonance imaging was used to identify intrinsic connectivity networks. Structural equation modeling revealed relations among personality, intelligence, and network coherence. Psychoticism, openness, and especially their shared variance were related positively to default network coherence and negatively to frontoparietal coherence. These associations remained after controlling for intelligence. Intelligence was positively related to frontoparietal coherence. Research suggests that psychoticism and openness are linked in part through their association with connectivity in networks involving experiential simulation and cognitive control. We propose a model of psychosis risk that highlights roles of the default and frontoparietal networks. Findings echo research on functional connectivity in psychosis patients, suggesting shared mechanisms across the personality–psychopathology continuum.

Key words: schizotypy/personality/fMRI/intelligence/Human Connectome Project

Introduction

Psychosis refers to a set of symptoms marked by loss of contact with reality, which are present in a range of disorders, most prominently schizophrenia. Schizophrenia symptoms can be divided into positive (hallucinations and delusions), negative (anhedonia and social withdrawal), and disorganized (confused thought, speech, and behavior) clusters.[1,2] In addition to studying psychosis in its severe manifestations, one can also study symptoms in the general population. Psychotic-like experiences (hallucinations and delusions) occur throughout the general population, even in the absence of disorder, and are distributed dimensionally.[3,4] Persistent psychotic-like characteristics, in the absence of severe mental illness, are often described as schizotypy, which has been conceptualized as part of schizophrenia's extended phenotype.[5] Schizotypy can also be divided into positive, negative, and disorganized traits. Positive schizotypy presents a promising phenotype for studying mechanisms of risk for psychosis and has been referred to as psychoticism.[6] (This construct should not be confused with Eysenck's Psychoticism, a misleadingly named trait reflecting impulsive nonconformity, low agreeableness, and low conscientiousness, but not psychosis proneness.[7,8]) Advantages of studying psychoticism include sampling a broader range of the variables of interest, greater capacity of the target population to participate, and fewer confounds stemming from comorbidity and medications.

Studying normal personality traits in relation to psychoticism can give us a fuller picture of mechanisms and risk factors for psychosis. One promising approach integrates psychoticism with the Five Factor Model or Big Five, a well-established, empirically based model describing personality in terms of Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness/Intellect.[9] Major dimensions of risk for psychopathology correspond structurally and conceptually to the Big Five, suggesting that mental disorders typically involve extreme and maladaptive forms of normal personality traits.[10,11] Psychoticism has historically been the psychopathology dimension most difficult to locate within the Big Five,[12–16] but recent research demonstrates that this is due to differential associations with the subfactors of Openness/Intellect. Psychoticism is specifically associated with openness but unrelated or even negatively related to intellect.[17–19] Although often also positively related to Neuroticism, psychoticism loads onto a separate factor with openness, especially when measured separately from intellect.[20–23] Openness encompasses artistic and aesthetic interests, fantasy proneness, and individual differences in perceptual engagement, whereas Intellect reflects intellectual confidence and engagement with abstract or semantic information and is positively associated with IQ.[18,19] The positive association between psychoticism and openness may arise from the fact that both involve sensitivity of pattern detection, with features like unusual perceptual experiences and magical ideation representing a tendency toward false positives, also known as "apophenia."[18,24,25] As a maladaptive form of openness, apophenia (and hence psychoticism) reflects "openness to implausible patterns."[18] Although evidence is beginning to emerge that the neurocognitive mechanisms of psychosis risk overlap with those of openness,[24,25] further research is critical. In the current work, we investigated patterns of functional connectivity associated with psychoticism and openness.

A substantial literature relates psychotic symptoms to abnormal patterns of structural and functional connectivity.[26,27] Whereas structural connectivity refers to the state of the brain's white matter pathways, functional connectivity refers to patterns of temporal synchrony among brain regions.[28–30] Measuring patterns of functional connectivity in concert with clinical or subclinical assessments may provide a useful mechanistic approach for understanding both full-blown psychosis and psychosis risk throughout the population.

Many studies examining functional connectivity in the psychosis spectrum report atypical connectivity in the default network (DN). Central hubs of the DN are in medial prefrontal cortex and posterior cingulate cortex, with additional nodes in the hippocampus and the parietal and temporal cortices.[31] The DN is involved in many cognitive operations, including episodic memory, future-directed thought, and understanding the mental states of others—in short, anything that requires simulation of experience rather than attention to sensory input. Increased connectivity and activity of the DN have been observed in schizophrenia and in people at high risk for psychosis, in resting-state designs[32–37] and tasks.[36,38–40] Other studies, however, have reported decreased DN connectivity in psychosis.[41–43] Nonetheless, comprehensive reviews suggest that a majority of studies report increased connectivity.[44]

Importantly, the relation of schizophrenia to DN connectivity appears to be specifically linked to positive symptoms.[36,38] Relatives of those with schizophrenia show increased DN connectivity,[33] which is also seen among individuals who report higher levels of mind-wandering[45] and creativity.[46,47] Perhaps not surprisingly, then, DN connectivity is also positively related to multiple facets of openness, even after controlling for intelligence.[25,48,49] This positive association has been found for global efficiency and connectivity of the DN,[25,48,49] but more specifically with connectivity in its core subsystem, consisting of posterior cingulate and medial prefrontal cortex.[49,50]

Although the DN is perhaps the most studied network in relation to psychosis and openness, portions of the network originally conceptualized as its counterpart—the task positive network—have also been investigated. One neural network is particularly associated with cognitive control and intelligence and is known as the frontoparietal control network (FPCN). The FPCN has primary nodes in dorsolateral prefrontal cortex, lateral parietal cortex, and dorsal anterior cingulate cortex, appears to be responsible for voluntary control of attention, and exhibits reduced connectivity in psychosis.[51] A more substantial literature has linked psychosis to disrupted function of the dorsolateral prefrontal cortex.[52–54] Intelligence shows positive relations to FPCN connectivity[55,56] and prominent frameworks for conceptualizing the neural basis of intelligence, such as Parieto-Frontal Integration Theory, underscore the importance of the FPCN.[57] Although intellect has been linked to greater performance-related activity in the FPCN,[58] possible relations with openness remain unclear.

When studying the neural correlates of traits involved in risk for disorder, it is important to consider intelligence, so as to ensure patterns of connectivity are not merely corresponding to broader deficits.[25] Low intelligence is a common risk factor for psychopathology, including psychosis.[59,60] As noted above, psychoticism also tends to be weakly inversely associated with intelligence and specific cognitive domains (e.g., social cognition, attention, and working memory).[18,19,61–67] Working memory, the ability to manipulate information in short-term memory, is thought to be a core mechanism underlying intelligence, a theory supported both by their strong correlation and shared neural correlates.[68,69] In the current research, we examined whether the neural correlates of intelligence were related to those of psychoticism and openness.

A large number of studies have identified the DN and FPCN as implicated in psychosis. Other research has connected psychoticism to openness, but few studies have investigated psychoticism and normal-range personality together with their neural correlates. The current study examined functional networks associated with psychoticism, openness, and intelligence, using resting state fMRI data from the Human Connectome Project (HCP). We hypothesized that psychoticism and openness would be associated positively with DN coherence and negatively with FPCN coherence. Furthermore, we anticipated intelligence would be positively associated with FPCN coherence.

Methods and Materials

Participants

Our sample included 1003 participants (534 females), from the HCP, between ages 22 and 37 (M = 28.7, *SD* = 3.7). Participants completed self-report measures and underwent four resting state fMRI scans (for 1-h total scan time). Exclusion criteria included a history of severe psychiatric, neurological, or medical disorders. However, participants were not excluded on the basis of mild psychopathology. Given population estimates, 15%–20% of participants would likely warrant a DSM-5 diagnosis.[70] Informed consent was obtained,[71] and study protocols were approved by the Institutional Review Board of Washington University in St. Louis (IRB # 201204036; "Mapping the Human Connectome: Structure, Function, and Heritability").

Self-report Measures

NEO Five-Factor Inventory. The NEO Five-Factor Inventory (NEO-FFI) is a measure of the Big Five. It consists of 60 items taken from the longer NEO Personality Inventory, Revised (NEO PI-R)[72] and uses a five-point Likert scale. The NEO-FFI does not include subscales for openness and intellect; in order to create an openness aspect scale, correlations of items from the NEO-FFI Openness to Experience scale were examined in relation to openness and intellect from the Big Five Aspect Scales (BFAS).[73] Previous work has been done to extract a similar openness aspect scale using the full NEO PI-R, based on item-associations with BFAS in three samples[74]; this latter scale has been used in previous research examining relations of openness and intellect with psychoticism.[17] Items from this NEO PI-R openness scale that are also included in the NEO-FFI were selected to create our FFI openness scale. Items included "Sometimes when I am reading poetry or looking at a work of art, I feel a chill or wave of excitement," "I am intrigued by the patterns I find in art and nature," "I don't like to waste my time daydreaming (reversed)," and "Poetry has little or no effect on me (reversed)." In the current study, validation was done using the Eugene Springfield Community Sample, where all items in our FFI openness scale had a correlation with BFAS openness greater than 0.30 and this correlation was at least 0.15 greater than for Intellect. There was a very strong positive correlation between latent variables indicated by FFI and BFAS openness items ($r = 1.0$, $P < .001$), and the latent correlation between FFI openness and BFAS intellect was significantly smaller ($r = .41$, $P = < .001$).

At the request of reviewers, who were concerned about our use of a nonstandard openness scale, we additionally ran all models involving openness using an Openness latent variable that included all 12 items from the NEO-FFI. Because the NEO-FFI Openness scale is tilted toward openness rather than intellect to begin with, this latent variable should still be informative for our hypotheses. Direction and significance of all effects remained the same in these analyses, indicating that our measurement model was not unduly influencing results.

Achenbach Self-report. [75] To measure psychoticism, participants were administered 123 items from the Achenbach self-report (ASR), an instrument used to assess dimensions of psychopathology, which uses a three-point Likert scale. Items used in the current study include those corresponding to psychoticism[74]: "I hear sounds or voices that other people think aren't there," "I see things that other people think aren't there," "I do things that other people think are strange," and "I have thoughts that other people would think are strange."

Despite limited breadth, these items are thought to provide an adequate measure of psychotic-like experiences in the general population and have previously been linked to increased psychosis risk, altered functional connectivity, cannabis use, and suicidal ideation.[76–81] In the current data set, these ASR items predict family history of schizophrenia diagnosis ($\beta = .15$, $P = < .001$), social cognitive deficits, and behavioral metrics of apophenia—false positive cognitions and perceptions[22]—above variance in these constructs explained by IQ.

Intelligence Measures

Intelligence measures were taken from two batteries: the NIH Toolbox[82] and Penn Computerized Neurocognitive Battery.[83] In a Picture Vocabulary test, participants selected which picture from multiple-choices most closely matched the meaning of a presented word. In a Matrix Reasoning test, participants completed a set of visual patterns (matrices) using one image from a set of multiple-choice options. For a List Sorting working memory test, participants were required to remember and sort a list of items presented verbally or visually.

Neuroimaging and Derivation of Networks

Resting state fMRI data were collected using a 3T Siemens Skyra scanner. After preprocessing,[84] artifacts were removed using ICA+FIX.[85,86] Mean head movement was also calculated and included as a covariate, as this index can correlate with variables of interest.[87,88] Interparticipant registration of the cortex was carried out using areal feature-based alignment and the Multimodal Surface Matching algorithm.[89,90] Each data set was temporally demeaned and underwent variance normalization.[91] Data were entered into a group-PCA,[91,92] the results of which were fed into group-ICA using FSL's MELODIC,[91,93] applying spatial-ICA at a dimensionality of 50.

After components were derived, we examined their association with canonical networks, by computing percentage of overlap with networks derived by Yeo,[30] focusing on the DN and FPCN. Of note, ICA components can overlap, whereas Yeo's maps were nonoverlapping parcellations. This provides a more realistic depiction of brain organization, as many regions are involved in multiple networks.[94] Despite showing larger local extent due to overlap, our components were smaller than Yeo's due to the higher dimensionality (50 vs. 7). Components with the highest overlap were visually inspected (further described in our supplement) to make sure they were centered on the correct networks. This yielded five FPCN and seven DN components (Figure 1).

Computation of Network Coherence Variables

After identifying networks, we used two-stage dual regression to compute network coherence for each ICA component, for each participant.[95,96] Coherence quantifies connectivity within each ICA component and corresponds to the average correlation of each voxel within a network with the time series of that network. Previous studies have linked coherence to personality traits (both pathological and normal-range).[97–99]

First, node-time series were estimated using the standard "dual-regression stage-1" approach, in which the full set of ICA maps were used as spatial regressors against participants' full time series data, estimating one time series per ICA map.[95] Next, in stage-2 dual regression, participant-specific time series were used as temporal regressors onto each participant's resting state data. This allowed us to derive a set of participant-specific spatial maps. Network coherence was then computed for each ICA component: group-level component maps (thresholded at $z_{max} > .30$) were binarized and applied as masks to each participant-level map,[96] and the mean correlation between each voxel within each participant-level spatial map and the mean time-series for all voxels in that component was calculated for each participant, for each component, giving coherence values for each of our 12 components. Finally, overall coherence scores for FPCN and DN were computed by averaging scores from corresponding components (further explained in our supplement).

Statistical Approach

Descriptive statistics were calculated for self-report measures and intelligence. Variables with skewness > 2.0 (which included only the psychoticism items)

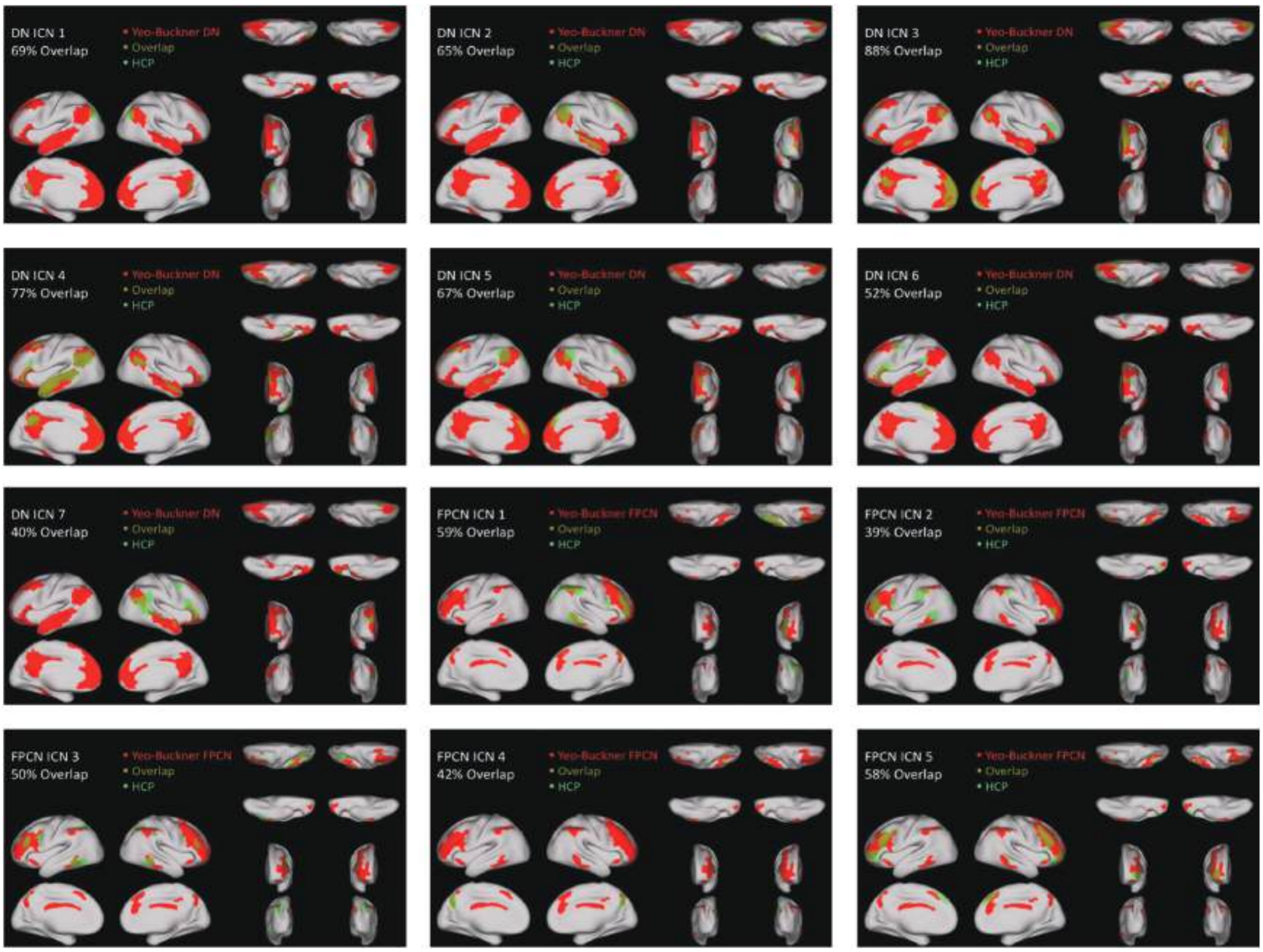


Figure 1. Visualizations of default and frontoparietal network components.

were logarithmically transformed. Structural equation modeling (SEM) was used to assess relations among latent factors. In each model, corresponding items or tasks were allowed to load onto latent variables for Psychoticism, Openness, and Intelligence. Due to the semantic similarity of two psychoticism items—"I do things that other people think are strange" and "I have thoughts that other people would think are strange"—and their high degree of correlation (r = .61, P < .001), we made an a priori decision to allow their residuals to correlate, which significantly improved model fit ($\Delta\chi^2$ = 265.1, P < .001).

SEMs with full information maximum likelihood estimation were used to test associations of Psychoticism, Openness, and Intelligence with DN and FPCN coherence. First, a model was fit to examine the relation between Openness and Psychoticism. For subsequent models, behavioral or self-report latent variables were used as criterion variables that were predicted by a set of observed variables: gender, age, head movement, DN coherence, and FPCN coherence. Predictors were allowed to correlate. Five models were created to examine effects of DN and FPCN coherence on (1) Psychoticism, (2) Openness, (3) shared variance of Openness and Psychoticism, and (4) Intelligence. For the first three models, a second iteration was fit with latent Intelligence as a covariate. Finally, scatter plots were used to visualize the relations of our Openness-psychoticism latent variable with DN and FPCN coherence (residualized by regression on coherence of the other network).

Results

Descriptive statistics are presented in Supplementary Table S1 and zero-order correlations are presented in Supplementary Table S2. SEM fit indices are presented in Table 1 and model results are presented in Table 2. For all models, all manifest variables had significant loadings on corresponding latent variables. Latent Psychoticism was significantly positively correlated with latent Openness (r = .61, P < .001).

In our first neural model, Psychoticism was negatively related to FPCN coherence but positively related to DN coherence. Similarly, DN coherence positively predicted and FPCN coherence negatively predicted Openness. These associations remained significant when controlling for Intelligence. The same patterns of association appeared, with even stronger regression weights, when testing associations of coherence with shared variance of openness and psychoticism (Figure 2). Scatter plots showing relations of latent Openness-psychoticism to residualized coherence variables are presented in Figure 3. Intelligence was positively associated with FPCN coherence but not associated with DN coherence. All regression paths from DN and FPCN coherence variables to criterion variables remained significant after estimation using robust (Huber-White) standard errors[100] and false-discovery-rate corrections.[101]

Discussion

As hypothesized, openness, psychoticism, and their shared variance were associated with increased DN coherence and decreased FPCN coherence. Intelligence showed a positive relation to FPCN coherence but no relation to DN coherence. Controlling for intelligence did not eliminate significant associations of coherence with openness and psychoticism. These findings suggest that similar biological mechanisms may underlie psychosis–proneness across traditional risk indicators and associated normal-range personality traits. Findings are in line with schizophrenia research, as a number of studies have demonstrated increased DN activity and connectivity in psychosis patients,[44] as well as decreased FPCN function.[51] To the best of our knowledge, however, this is the first study to examine associations between network coherence, psychoticism, and openness in a large community sample.

Although our results are in line with patterns of functional connectivity reported in the existing literature, it tional connectivity reported in the existing literature, it

Table 1. Fit statistics for structural equation models

Models	χ^2	P	RMSEA	95% C.I.	TLI	CFI
Model 1(Psychoticism and O)	71.2	< .001	.050	[.038, .062]	.960	.974
Model 2a (Psychoticism, DN, FPCN)	49.8	< .001	.048	[.035, .061]	.962	.983
Model 2b (Psychoticism, DN, FPCN, IQ)	131.5	< .001	.046	[.038, .055]	.948	.971
Model 3a (O, DN, FPCN)	75.5	< .001	.053	[.041, .066]	.955	.979
Model 3b (O, DN, FPCN, IQ)	244.5	< .001	.067	[.059, .075]	.898	.941
Model 4a (O-Psychoticism, DN, FPCN)	186.2	< .001	.046	[.039, .054]	.947	.964
Model 4b (O-Psychoticism, DN, FPCN, IQ)	384.7	< .001	.054	[.049, .060]	.907	.934
Model 5 (IQ, DN, FPCN)	55.7	< .001	.062	[.046, .078]	.943	.980

Notes: O = Openness, O-Psychoticism = Shared variance of Openness and Psychoticism, DN = Functional coherence of the default network, FPCN = Functional coherence of the frontoparietal control network.

Table 2. Structural equation model analyses—aggregate DN and FPCN coherence variables

	Model 1			Model 2		
Models	*z*	β	*P*	*z*	β	*P*
Psychoticism						
Gender	0.4	.01	.715	0.9	.04	.385
Age	−2.3	−.09	.023	−2.1	−.08	.039
Head Movement	−0.4	−.02	.681	−0.8	−.04	.422
DN Coherence	2.8	.25	.006	2.2	.24	.031
FPCN Coherence	−3.6	−.32	< .001	−2.7	−.29	.007
Intelligence				−2.3	−.13	.021
Openness						
Gender	−2.2	−.08	.030	−2.5	−.10	.013
Age	−1.7	−.06	.089	−1.9	−.07	.053
Head Movement	−1.2	−.04	.245	−0.8	−.03	.415
DN Coherence	2.7	.21	.008	2.7	.21	.006
FPCN Coherence	−3.0	−.25	.003	−3.2	−.28	.001
Intelligence				1.9	.09	.056
Psychoticism-Openness (Shared Variance)						
Gender	−1.4	−.10	.176	−0.6	−.07	.578
Age	−2.3	−.20	.019	−2.3	−.20	.019
Head Movement	−1.1	−.09	.263	−1.0	−.09	.310
DN Coherence	3.0	.60	.003	2.4	.61	.015
FPCN Coherence	−3.3	−.75	.001	−2.8	−.76	.005
Intelligence				−0.3	−.06	.762
Intelligence						
Gender	5.7	.21	< .001			
Age	2.6	.10	.010			
Head Movement	−3.4	−.14	.001			
DN Coherence	−0.8	−.07	.410			
FPCN Coherence	3.7	.27	< .001			

is worth discussing them in the context of two preliminary studies using HCP data.[76,79] Using subsamples of 229 and 468 participants, respectively, one study found increased visual connectivity and decreased DN connectivity among participants endorsing psychoticism items,[76] and the other found psychoticism was negatively associated with global efficiency of the cingulo-opercular network and DN.[79] Our finding of positive association between DN and psychoticism may at first seem contradictory, but several factors may account for this discrepancy. One explanation is that these earlier studies did not control for variance shared among networks, which is likely artifactual.[87] Indeed, our results did not show a significant zero-order relation between DN coherence and psychoticism, but controlling for FPCN coherence revealed a significant positive association. This is an example of statistical suppression,[102] and it can be interpreted as meaning that if one examined individuals with equivalent levels of FPCN coherence, one would expect to find that individuals higher in Openness or psychoticism would have higher DN coherence. Furthermore, our use of a larger sample and latent variable modeling will have improved our statistical power, relative to those earlier studies.

Our findings suggest that psychoticism and openness reflect heightened coherence in the DN but lower levels of FPCN coherence, possibly reflecting a tendency toward spontaneous self-generated thought coupled with reduced cognitive control. Such patterns are also reflected in studies showing that DN connectivity is positively related to positive symptoms, among patients with psychosis.[43] In terms of FPCN function, psychosis is not only linked specifically to dysfunction of prefrontal cortex, but it is also associated with broad cognitive deficits that have directly been tied to FPCN dysfunction.[103] Taken together, these neurocognitive associations with psychosis–proneness may suggest an increased default-network tendency toward erroneous thoughts and perceptions (false positives), coupled with diminished reality testing to screen out false positives. Future research including more comprehensive measurement of psychotic symptoms would be useful to confirm these speculations.

Importantly, the pattern of FPCN coherence that emerged for psychoticism and openness was also present, in the opposite direction, for intelligence. Given the prevalence of intelligence and working memory deficits in psychosis and its extended phenotype, it is not surprising similar patterns of connectivity would underlie these constructs. Mind wandering and related processes associated with openness, psychoticism, and the DN may directly compete with demands of intelligence and working

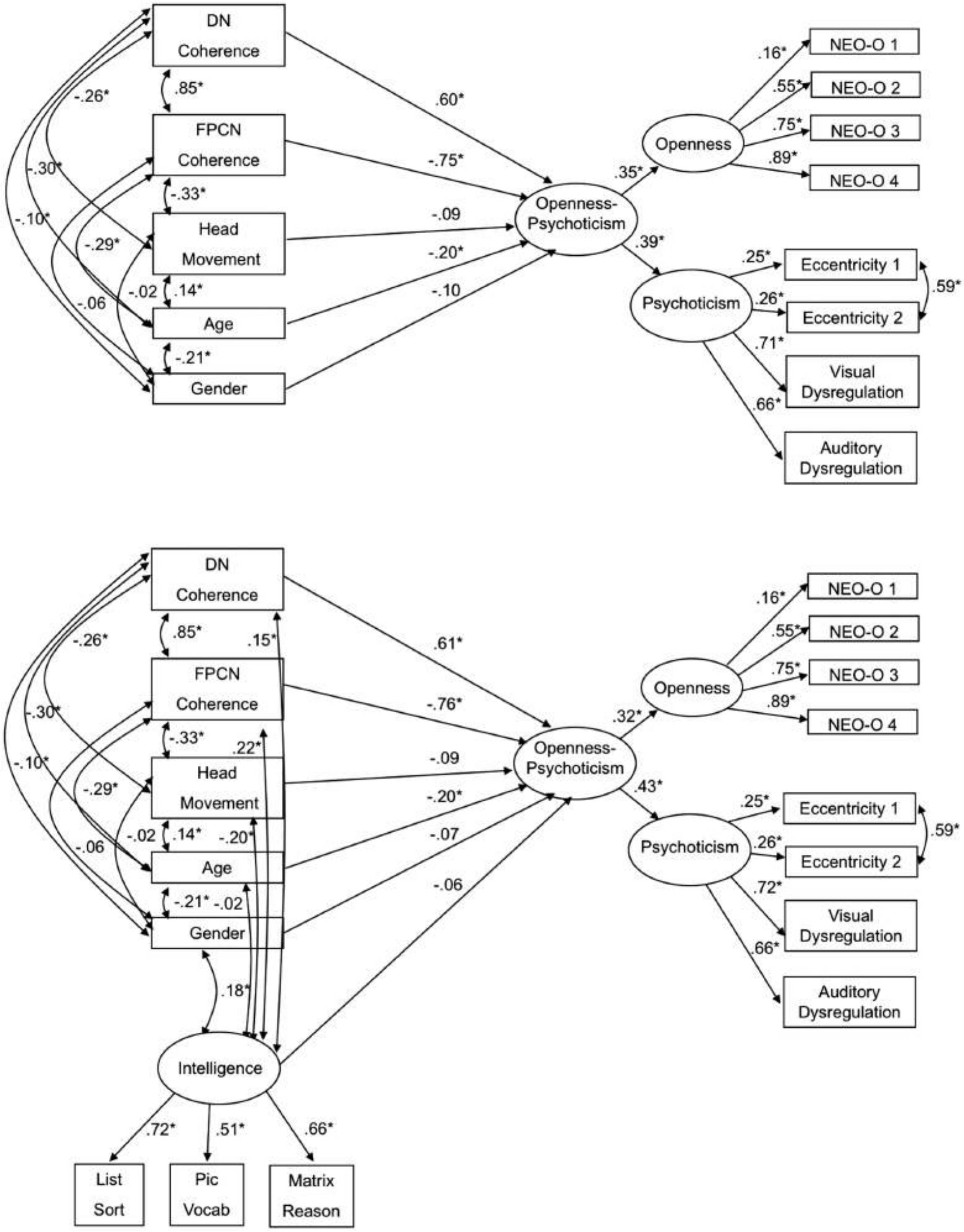


Figure 2. Structural equation models of Openness-psychoticism predicted by DN and FPCN coherence.

memory, leading to an inverse association with FPCN coherence. Indeed, research suggests that disrupted connectivity of the FPCN is associated with positive symptoms and working memory deficits among psychosis patients.[104] In short, impaired cognitive function and associated neural correlates may play an important role in the broader mechanisms of psychosis and its extended phenotype.

The current study provides important novel insights into the measurement of neurobehavioral characteristics of the openness-psychoticism dimension, a key step in better characterizing possible mechanisms underlying

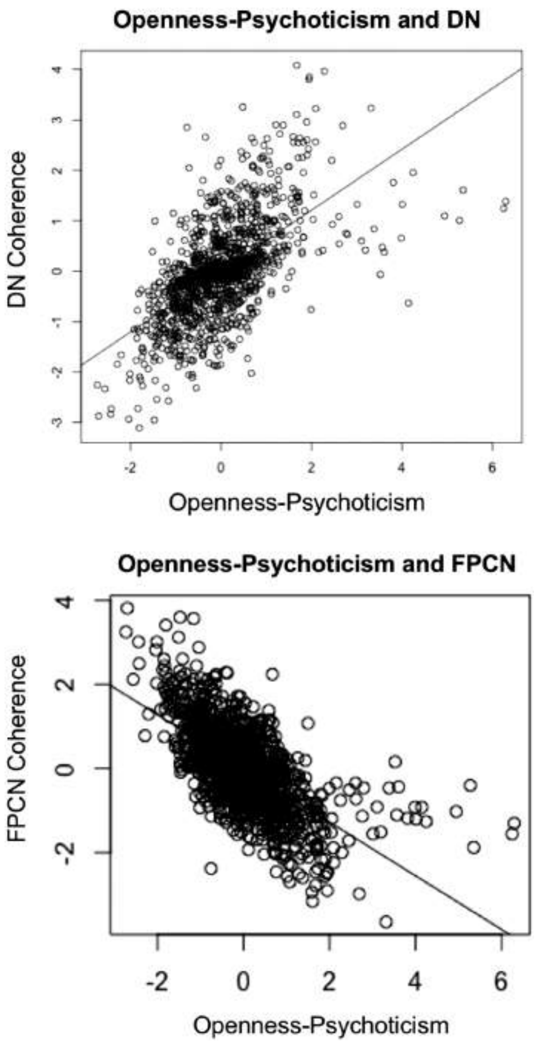


Figure 3. Scatter plots of Openness-psychoticism and residualized coherence variables.

risk factors for psychosis. Because psychotic symptoms are transdiagnostic features seen across a number of disorders, elucidating their neural mechanisms may eventually help facilitate more effective methods for assessment and treatment. Such an approach is in line with the National Institute of Mental Health's Research Domain Criteria initiative,[105] the Hierarchical Taxonomy of Psychopathology's conceptualization of psychiatric illness,[11] and a new theory that psychopathology is typically caused by extremity in normal personality mechanisms that interferes with goal-directed functioning.[10] Investigation of psychoticism, openness, and their relation to individual differences in functional brain networks and cognitive symptoms is a promising avenue for continued research into the etiology of risk for psychosis. Of particular relevance are theories positing psychosis and autism as diametrical disorders, involving divergent patterns of FPCN-DN coordination and associated cognitive processes.[106–108]

Limitations

Although the current study had multiple strengths, some limitations are worth noting. First, although the Achenbach Self Report is a reliable, well-established measure, its items measuring psychoticism are limited. Use of more extensive measures of psychoticism—such as the Personality Inventory for DSM-5—could allow more robust testing of associations with neural variables. Second, our measures of openness and psychoticism were self-reported and could be usefully supplemented by clinician ratings or peer-reports. Finally, although the current study demonstrates links between neural and behavioral metrics in a population without severe mental illness, we cannot tell how well they would generalize to clinical populations. Further research should be undertaken to examine the roles of functional connectivity and cognitive deficits in those with active psychosis.

Conclusion

Current findings suggest that psychoticism and openness are linked in part through their association with altered connectivity in neural networks associated with experiential simulation and cognitive control. It is increasingly recognized that risk for psychosis is distributed throughout the population dimensionally and that understanding subclinical indicators of risk in the general population is crucial to developing better models of psychosis etiology.[3,4,109] Assessing the coherence of intrinsic connectivity networks may provide a useful transdiagnostic approach to elucidating cognitive and neural mechanisms involved in psychosis and related phenomena. Our results advance understanding of the neural mechanisms of psychoticism that are shared with openness, a personality trait that appears to contribute specific risk for psychosis. This study adds to a growing body of research characterizing the underlying biology of transdiagnostic psychiatric features through the use of large, nonpatient samples and also underscores the continuity between normal personality variation and risk for psychopathology.

Supplementary Material

Supplementary material is available at https://academic.oup.com/schizophreniabulletin/.

Funding

This work was supported by the National Institutes for Health (1U54MH09165 [to D.V.E. and K.U.]) and the McDonnell Center for Systems Neuroscience at Washington University. Scott D. Blain was supported by the National Science Foundation Graduate Research Fellowship (1839286).