

# The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative

Monica E. Calkins,<sup>1</sup> Kathleen R. Merikangas,<sup>2</sup> Tyler M. Moore,<sup>1</sup> Marcy Burstein,<sup>2</sup> Meckenzie A. Behr,<sup>3</sup> Theodore D. Satterthwaite,<sup>1</sup> Kosha Ruparel,<sup>1</sup> Daniel H. Wolf,<sup>1</sup> David R. Roalf,<sup>1</sup> Frank D. Mentch,<sup>3</sup> Haijun Qiu,<sup>3</sup> Rosetta Chiavacci,<sup>3</sup> John J. Connolly,<sup>3</sup> Patrick M.A. Sleiman,<sup>4</sup> Ruben C. Gur,<sup>1</sup> Hakon Hakonarson,<sup>3,\*</sup> and Raquel E. Gur<sup>1,\*</sup>

<sup>1</sup>Department of Psychiatry, Neuropsychiatry Section, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; <sup>2</sup>National Institute of Mental Health, Genetic Epidemiology Research Branch, Intramural Research Program, Bethesda, MD; <sup>3</sup>Center for Applied Genomics at The Children's Hospital of Philadelphia, Philadelphia, PA; <sup>4</sup>Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

**Background:** An integrative multidisciplinary approach is required to elucidate the multiple factors that shape neurodevelopmental trajectories of mental disorders. The Philadelphia Neurodevelopmental Cohort (PNC), funded by the National Institute of Mental Health Grand Opportunity (GO) mechanism of the American Recovery and Reinvestment Act, was designed to characterize clinical and neurobehavioral phenotypes of genotyped youths. Data generated, which are recently available through the NIMH Database of Genotypes and Phenotypes (dbGaP), have garnered considerable interest. We provide an overview of PNC recruitment and clinical assessment methods to allow informed use and interpretation of the PNC resource by the scientific community. We also evaluate the structure of the assessment tools and their criterion validity. **Methods:** Participants were recruited from a large pool of youths ( $n = 13,958$ ) previously identified and genotyped at The Children's Hospital of Philadelphia. A comprehensive computerized tool for structured evaluation of psychopathology domains (GOASSESS) was constructed. We administered GOASSESS to all participants and used factor analysis to evaluate its structure. **Results:** A total of 9,498 youths (aged 8–21; mean age = 14.2; European American = 55.8%; African American = 32.9%; Other = 11.4%) were enrolled. Factor analysis revealed a strong general psychopathology factor, and specific 'anxious-misery', 'fear', and 'behavior' factors. The 'behavior' factor had a small negative correlation ( $-0.21$ ) with overall accuracy of neurocognitive performance, particularly in tests of executive and complex reasoning. Being female had a high association with the 'anxious-misery' and low association with the 'behavior' factors. The psychosis spectrum was also best characterized by a general factor and three specific factors: ideas about 'special abilities/persecution,' 'unusual thoughts/perceptions', and 'negative/disorganized' symptoms. **Conclusions:** The PNC assessment mechanism yielded psychopathology data with strong factorial validity in a large diverse community cohort of genotyped youths. Factor scores should be useful for dimensional integration with other modalities (neuroimaging, genomics). Thus, PNC public domain resources can advance understanding of complex inter-relationships among genes, cognition, brain, and behavior involved in neurodevelopment of common mental disorders. **Keywords:** Community cohort, children, adolescents, young adults, psychopathology, mood, anxiety, behavior, psychosis, comorbidity, structure, genomics, neuroimaging, neurocognition, public domain.

## Introduction

Mental illnesses are common disorders that emerge during childhood and adolescence and often persist into adulthood, with debilitating consequences. Both biologic and environmental risk factors underlie the complex phenotypes manifested in mental illnesses (Rapoport, Giedd, & Gogtay, 2012). Quantitative phenotypic brain-behavior measures enable rigorous research that can bridge molecular biology with disease phenomenology (Insel, 2014). Linking disease phenotypes to intermediate variables that modulate disease manifestations will help elucidate how these factors contribute to shape the neurodevelopment of brain systems that underlie complex behavior. To achieve this linking, large phenotypically characterized samples of youth are required for

integration with genomic methods that can potentiate targeted and effective interventions. The Philadelphia Neurodevelopmental Cohort (PNC) is the first such sample in the United States (U.S.).

The PNC is a collaboration between the Children's Hospital of Philadelphia (CHOP) and the University of Pennsylvania that capitalized on an unprecedented opportunity: an already recruited and genotyped large sample of youths who had assented or consented to be contacted for further research. The primary aim of the project was to assess behavioral dimensions reflecting vulnerability to major mental illnesses by characterizing clinical and neurobehavioral phenotypes in a cohort of approximately 10,000 genotyped youths. The phenotypic dimensions included clinical assessment and neurobehavioral measures of cognitive and emotion processing related to neural systems vulnerable to neurodevelopmental aberrations (Gur et al., 2012, 2014). The

\*These authors contributed equally.

Conflict of interest statement: No conflicts declared.

clinical assessment comprehensively measured psychopathology domains including anxiety, mood, eating, behavior disorders, and a broad 'spectrum' of psychosis-relevant experiences. The psychosis spectrum ranges from subtle and subclinical (positive, negative or disorganized) symptoms, which would not qualify for diagnosable disorders, to clinically salient positive symptoms (threshold hallucinations and/or delusions) that fulfill criteria for serious psychotic disorders (Binbay et al., 2012). Behavior disorders were the most frequent mental disorder in this cohort (18.2%), followed by attention deficit hyperactivity disorder (17.6%), mood disorders (12.4%), and anxiety disorders (11.7%) (Merikangas et al., 2015). Threshold positive-psychosis symptoms were reported by 3.7%, an additional 12.3% reported significant subpsychotic positive symptoms, and a minority of youths (2.3%) endorsed subclinical negative/disorganized symptoms in the absence of positive symptoms (Calkins et al., 2014).

A subsample underwent neuroimaging to establish neural substrates of behavioral phenotypic trajectories (Satterthwaite, Elliott et al., 2014). The PNC provides a public domain resource for investigations and characterization of gene networks underlying neuronal vulnerability leading to mental disorders. Data generated are available through the National Institute of Mental Health's Database of Genotypes and Phenotypes (dbGaP) (dbGaP, 2014), and there were 70 approved data requests as of Feb 2015, suggesting great scientific interest in this resource.

Here, we present an overview of PNC recruitment and clinical assessment methods. We detail the methods used to assess psychopathology to enable informed use of these data by the scientific community, and present the results of factor analysis that was employed to organize the common processes underlying psychopathology (Beesdo-Baum et al., 2009). Prior factor-analytic studies in adults suggest that three factors underlie comorbidity among mental disorders, with an externalizing factor and two correlated factors, 'anxious-misery' and 'fear', forming a second-order 'internalizing' factor (Krueger, 1999). The three-factor structure, but without the higher-order internalizing factor, has been replicated in a large prospective study of children aged 14–24 (Beesdo-Baum et al., 2009). However, this structure may not replicate when the number of disorders increases beyond the core disorder domains included in most investigations (Wittchen et al., 2009). In addition, recent evidence suggests that psychiatric disorders are best explained by 'internalizing', 'externalizing', 'thought disorder' liabilities, and a single higher-order 'general psychopathology' dimension (Caspi et al., 2014). Finally, many factor-analytic studies omitted psychotic symptoms (Krueger, 1999), evaluated psychosis in relation to few other disorder domains, or restricted analyses to psychiatric samples (Kotov, Chang et al., 2011; Kotov, Ruggero et al., 2011). We recently reported

that psychosis spectrum features in the PNC are common and associated with comorbid psychopathology, distress, and neurocognitive and functional impairment (Calkins et al., 2014; Gur et al., 2014). Despite growing evidence for the clinical salience of early psychotic-like experiences, to our knowledge, no prior factor-analytic investigation in youth has included measures of the broader psychosis spectrum, nor has the structure of the psychosis spectrum been examined in a large systematic sample of U.S. youth. We therefore investigated whether the previously reported three-factor structure of psychopathology fit the PNC data including the psychosis spectrum measures, and examined the structure of these psychosis spectrum tools.

## Methods

### Recruitment

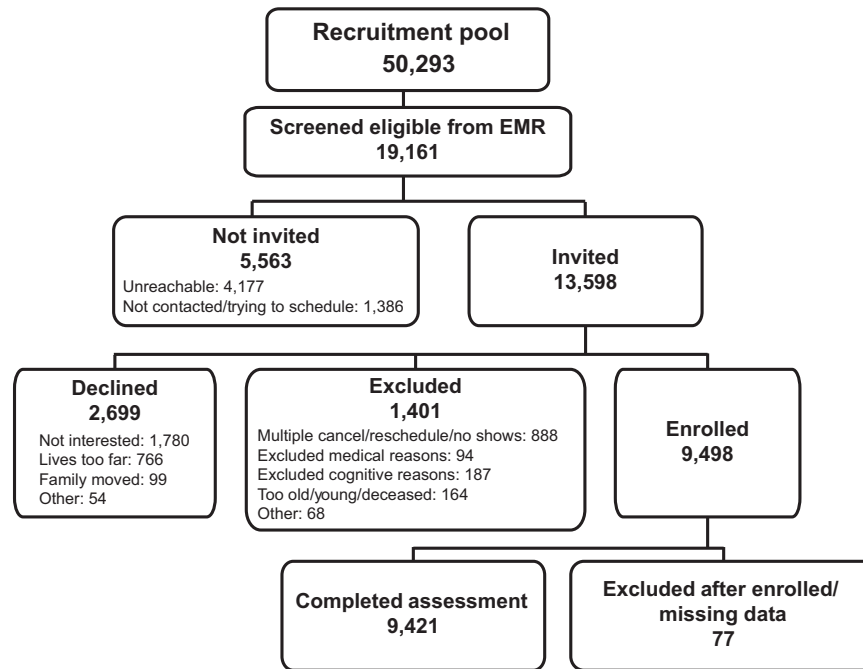
Between 2006 and 2012, 50,293 youths were recruited by the Center for Applied Genomics at CHOP through a pediatric healthcare network of over 30 clinical community sites in the tristate area of Pennsylvania, New Jersey, and Delaware. When undergoing blood work, patients were approached for participation in the recruitment pool. The percentage of patients undergoing blood work across recruitment sites varied from 11 to 53%, with a mean of 36%. Participants provided a blood sample for genomic studies and access to Electronic Medical Records (EMRs).

The EMR of each participant was screened for preliminary eligibility for PNC participation. Potential participants were included if they were between the ages of 8–21, had provided written informed consent/assent to be re-contacted for future studies, were proficient in English, and did not appear to have significant developmental delays or physical conditions that would interfere with their ability to complete study procedures. This screening yielded a pool of 19,161 eligible individuals, released to the recruitment team in weekly waves primarily over the 2-year period from November 2009 to December 2011 (see Figure 1). Recruitment for genomic samples was supervised by CHOP faculty and staff. Recruitment for phenotyping was conducted by a centralized team under the direction of the first author, with the assistance of CHOP staff. Potential probands (ages 18–21) or proband's caregivers/legal guardians (ages 8–17) were sent a letter introducing the study, and then contacted by phone using standardized scripts to explain the study, gauge interest, verify study eligibility, and schedule appointments. Recruitment processes and strategies are detailed in Appendix S1, available online.

### Procedures

After complete description of the study, written informed consent was obtained from participants aged  $\geq 18$ , and written assent and parental permission were obtained from children aged  $<18$  and their parents/legal guardian. The University of Pennsylvania and CHOP Institutional Review Boards approved all procedures.

**Psychopathology assessment.** Psychopathology was assessed using a computerized, structured interview (GOASSESS) that was administered to collaterals, who were caregivers or legal guardians (ages 8–10), probands and collaterals (ages 11–17), and probands (ages 18–21). Components of the interview by informant type are presented in Table S2. The user interface of GOASSESS mimicked the paper interview on



**Figure 1** Recruitment flow of the Philadelphia Neurodevelopmental Cohort. Among ineligible youths ( $n = 31,132$ ), approximately 31% had serious medical conditions precluding participation in the study procedures (severe developmental delay, significant hearing loss, limited mobility), 44% were outside the study age range or deceased, and 36% declined to be re-contacted for future studies

which assessors were trained rather than a data entry screen, incorporating reminders in the margins, and providing guided, as well as flexible, navigation between pages. Ample space was provided to record participant responses verbatim, along with observations and interpretative comments. Skip logic and error-checking ensured real-time and accurate data collection. Assessors uploaded data daily from their local database to the central GOASSESS database.

GOASSESS was designed to be highly structured, with screen-level symptom and episode information in order to: (a) allow rapid training and standardization across a large number of assessors conducting most assessments in the field; (b) provide for the brief administration necessary to allow assessment of 100–165 participants per week. The instrument is abbreviated and modified from the epidemiologic version of the NIMH Genetic Epidemiology Research Branch Kiddie-SADS (Merikangas, Avenevoli, Costello, Koretz, & Kessler, 2009). Modifications included: (a) addition of the diagnostic screening questions from the adolescent version of the World Health Organization Composite International Diagnostic Interview (CIDI), inserted at the beginning of each psychopathology domain section (Merikangas et al., 2009); (b) less restrictive criteria to complete the diagnostic modules (e.g. symptom counts or duration were not required to meet DSM-IV episodes in order to complete each section), and; and (c) derivation of dimensional ratings of distress and impairment associated with symptoms in each diagnostic section. Components of the assessment referenced in the current article (with additional components described in the Appendix S1) were as follows:

1. Demographics and medical history. This section, administered to collaterals and adult probands, inquired about school and work history, developmental and medical history, and current medications. To assign a medical rating to complement up-to-date findings in the participant's EMR, we included GOASSESS medical history data to ensure that we had the most recent update for children who visited CHOP for tertiary care (further detailed in the Appendix S1).
2. Psychopathology screen. The screen assessed psychiatric and psychological treatment history, and lifetime

occurrence of major domains of psychopathology including mood (major depressive episode, manic episode), anxiety (generalized anxiety disorder, separation anxiety disorder, specific phobia, social phobia, panic disorder, agoraphobia, obsessive-compulsive disorder, post-traumatic stress disorder), attention deficit/hyperactivity disorder (ADHD), behavioral (oppositional defiant disorder, conduct disorders), eating disorders (anorexia, bulimia), and suicidal thinking and behavior. Each section included a screen for relevant symptoms and additional DSM-IV criteria, including symptom frequency, duration of symptoms and episodes, and onset and offset of symptoms and episodes. For all symptoms endorsed that met threshold to continue within the section, the participant was asked to rate associated distress and impairment on separate 11-point scales ranging from 0 (no bother/problems) to 10 (extremely serious bother/problems). The specific phobia section is provided as an illustrative example in Figure S1. All sections are available by request to the authors, and all individual GOASSESS items are available through dbGaP with approved requests.

Three screening tools to assess the psychosis spectrum were embedded within the psychopathology screen as described recently (Calkins et al., 2014): (a) positive subpsychotic symptoms in the past year were assessed using the 12-item, assessor administered, PRIME Screen-Revised [PS-R (Kobayashi et al., 2008; Miller, Cicchetti, Markovich, McGlashan, & Woods, 2004)]; (b) positive threshold psychotic symptoms (lifetime hallucinations and/or delusions) were assessed with the K-SADS psychosis screen, modified as for the other psychopathology domains described above; (c) negative/disorganized symptoms were assessed using six assessor-rated subscales of the Scale of Prodromal Symptoms (SOPS) from the Structured Interview for Prodromal Syndromes (SIPS) (McGlashan et al., 2003): Avolition; Expression of Emotion; Experience of Emotions and Self; Occupational Functioning; Trouble with Focus and Attention; Disorganized Communication. The SOPS Focus/Attention scale was rated based on responses and interviewer observations following administration of the ADHD items; Expression of Emotion based on interviewer

observations; and the remaining four scales based on selected SIPS questions.

**Psychopathology summary variables.** Computerized algorithms assigned rankings for DSM-IV indices of each psychopathology domain based on endorsement of contributing items by either the proband or collateral (behavioral domains), or by the proband (all other domains). For the current analyses, psychopathology domains were considered significant if sufficient symptoms were endorsed with frequency and duration approximating DSM-IV disorder or episode criteria, accompanied by significant distress or impairment rated  $\geq 5$ . Comparison of the diagnostic algorithms from this interview with the full criteria using data from the National Comorbidity Survey-Adolescent (Merikangas et al., 2009) yielded fair (i.e. eating disorders) to excellent (i.e. ADHD) area under the receiver-operating characteristic curve values for the major classes of disorders (data available upon request).

Psychosis Spectrum criteria were as follows:

1. Positive-Subpsychosis: Because age effects were observed on PRIME Screen-Revised total scores (Calkins et al., 2014), an Age Deviant index was derived using proband total score, normalized within age year, identifying children with extreme total scores ( $z \geq 2$ ) compared with age mates. Additionally, because psychosis risk may not be linearly related to total scores such that endorsement of even one symptom at a severe level (Definitely Agree = 6) may be indicative of psychosis risk, and to allow comparison with prior studies with the PRIME Screen, an Extreme Agreement index was also calculated based on traditional criteria [ $\geq 1$  item rated 6 or  $\geq 3$  items rated 5 (Somewhat Agree) (Miller et al., 2004)].
2. Positive-psychosis: Possible or definite hallucinations or delusions based on K-SADS screen, with duration  $\geq 1$  day, occurring outside the context of substance, illness and medicines, and accompanied by significant impairment or distress (rating  $\geq 5$ ).
3. Negative/disorganized: An Age Deviant index of negative/disorganized symptoms was generated using Scale of Prodromal Symptoms z-scores. Specifically, SOPS total scores were normalized within age year;  $z \geq 2$  cutoff reflected extreme ratings of negative or disorganized symptoms for age cohort.

Individuals evidencing any one of these three criteria were classified as having significant 'psychosis spectrum' symptoms.

**Assessor training and quality assurance.** Bachelor's and Master's level assessors underwent a common 25-hr training protocol developed and implemented by the first author that included didactic sessions, assigned readings, and supervised pairwise practice. They were certified for independent assessments through a standardized procedure requiring observation by a certified clinical observer who rated the proficiency of the assessor on a 60-item checklist of interview procedures ( $n$  certified assessors = 55). Additionally, responses coded in GOASSESS by the assessor were required to correspond to responses coded by a certified clinical observer. Assessors who did not achieve these standards were required to undergo repeat observation until passing criteria were met. Assessor drift was monitored and corrected through periodic review of audio-recordings of real interviews, and re-training and re-certification conducted at data collection mid-point. Assessors were assigned a maximum of 10 interviews a week, with the goal of completing 5–7 interviews per week. To assure the quality of interview data, each assessment underwent a computerized error-checking algorithm that identified areas requiring assessor's attention, and a standardized postadministration review process by certified clinical

reviewers. Results were reported to assessors and supervisors. A computerized chart review module provided management tools for the comprehensive review process for supervisors, reviewers, and assessors, as well as an automated check to ensure that all steps were completed successfully. Data were checked and corrected prior to final inclusion in the dataset.

**Other PNC procedures.** All enrolled probands were also administered a Computerized Neurocognitive Battery (CNB), described in detail elsewhere (Gur et al., 2012, 2014). Several additional computerized measures were administered at the time of the CNB. Substance use history was assessed by an abbreviated version of a widely used self-report measure (Han, McGue, & Iacono, 1999) computerized locally. Current pubertal status was measured by an abbreviated, computerized, and privately self-administered (Satterthwaite, Vandekar et al., 2014) version of a self-report measure of pubertal status (Morris & Udry, 1980). The Wide Range Achievement Test (WRAT-4) Reading subscale (Wilkinson & Robertson, 2006) provided an estimate of IQ. At the conclusion of each appointment, all participants were given a flyer with information about the neuroimaging study. The neuroimaging team then proceeded with recruitment of eligible probands as described elsewhere (Satterthwaite, Elliott et al., 2014); a subset of 1,445 participants completed the imaging procedures.

**Statistical analyses.** First, we summarized the recruitment outcome and demographic characteristics of the final cohort. Second, to investigate the factor structure of the GOASSESS, we used dichotomized GOASSESS summary scores (significant vs. not significant) from all assessed psychopathology domains to run four exploratory factor analyses (EFA's) ranging from 1 (unidimensional) to four factors. We then ran a confirmatory factor analysis based on the literature suggesting 3-factors ('anxious-misery', 'fear', and 'externalizing'). Third, to investigate dimensions of the psychosis spectrum in young people, we conducted factor analysis of psychosis spectrum screening tools. Bifactor modeling (Holzinger & Swineford, 1937) is a method to model data that have both a strong underlying factor (indicated by all variables) and several individual factors (each indicated by only some variables). The bifactor model is unique insofar as each variable loads on two factors simultaneously, such that the two factors 'compete' for the explainable variance in the variable (Reise, Moore, & Haviland, 2010). We aimed to extract three factors based on previous literature (Beesdo-Baum et al., 2009; Krueger, 1999). All analyses were estimated using mean- and variance-adjusted weighted least squares (wlsmv) in Mplus (version 7) (Muthén & Muthén, 2013).

## Results

Recruitment data are shown in Figure 1. From the recruitment pool, 13,598 participants were invited (71.0% of EMR eligible), 9,498 were enrolled (64.3% of invited), and 9,421 completed the assessment. Demographic data of the recruitment pool are provided in Table S1. The proportion of males to females was comparable across *enrolled*, *not invited*, *declined*, and *excluded* youths. *Enrolled* participants were on average slightly younger than those *not enrolled*; however, the mean difference in age among the groups was  $<1$  year, with small corresponding effect sizes ranging from 0.12 to 0.21. A disproportionate number of *enrolled* and *declined* individuals were European American, whereas those *not invited* and *excluded* were disproportionately other racial

groups. Among those *not invited*, a comparatively higher number of non-European American’s were *unreachable*. Figure 2A presents the number of assessments conducted each year, showing that the majority of assessments were completed over a 2-year period. The distribution of assessments by date and participant age is shown in Figure 2B; following intentional recruitment of older probands during the start-up period, recruitment was distributed across the age range for the remainder of the recruitment period. The majority of evaluations ( $n = 6,228$ ; 65.6%) occurred in participant’s homes,

with the remainder conducted in CHOP or Penn offices ( $n = 3,110$ ; 32.7%) or other locations (e.g. libraries;  $n = 160$ ; 1.7%).

*Demographics of the Philadelphia Neurodevelopmental Cohort*

Demographic characteristics of the final PNC sample are presented in Table 1. The mean age of the cohort is 14.2 years ( $SD = 3.7$ ), and females represent slightly more than half of the total sample (51.7%), with the proportion of females increasing significantly

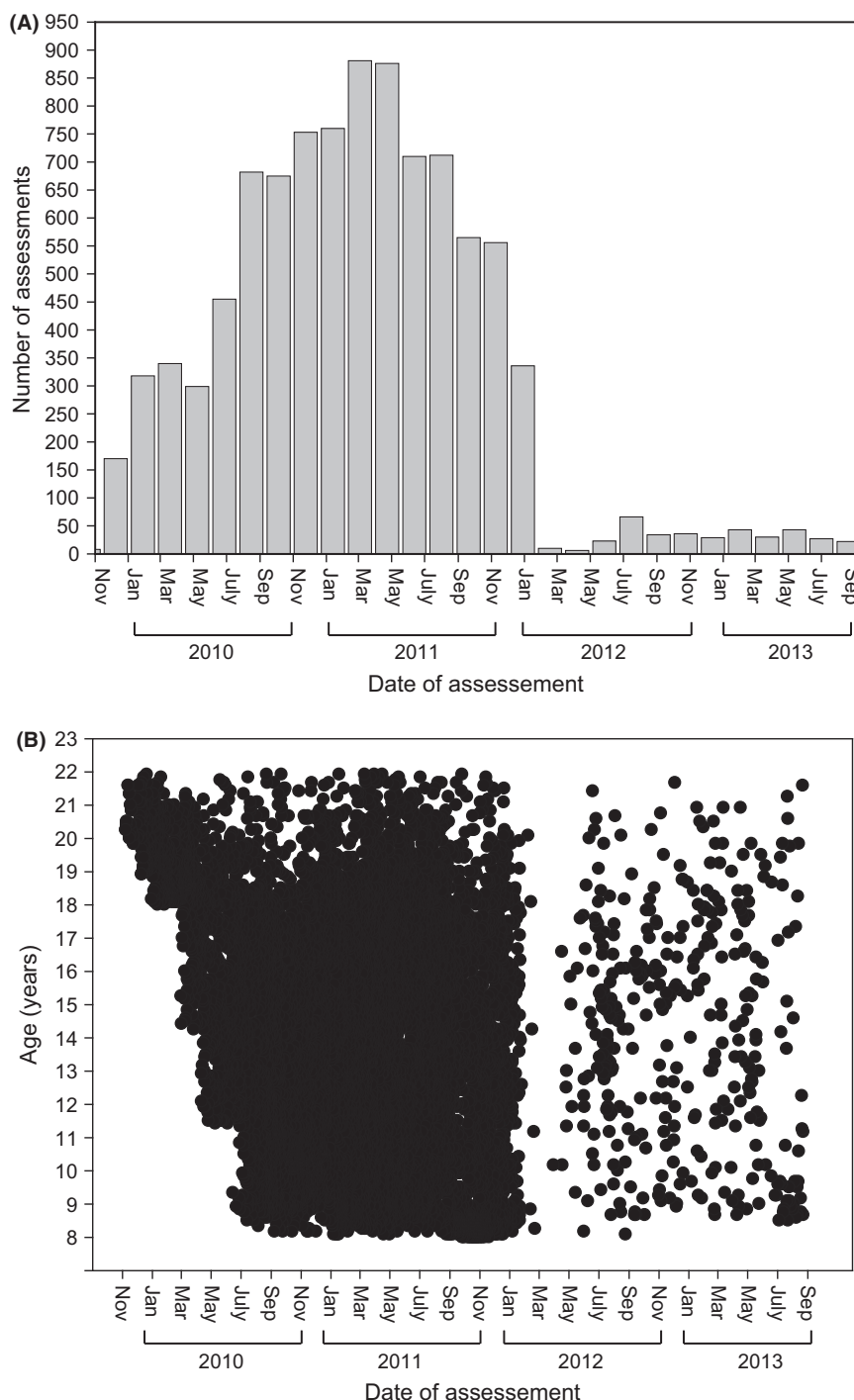


Figure 2 (A) Number of clinical assessments by date of assessment. (B) Date of clinical assessment by proband age

**Table 1** Philadelphia Neurodevelopmental Cohort demographic characteristics

	Total sample	Proband age group		
		8–10	11–17	18–21
<i>n</i>	9498	2392	5214	1892
Age				
Mean (years)	14.2	9.4	14.5	19.3
SD	3.7	0.8	2.0	1.0
Male:Female	4592:4906	1321:1071	2500:2714	771:1121
% Female	51.7	44.8	52.1	59.2
Race				
European American %	55.8	54.9	56.8	54.0
African American %	32.9	32.3	32.0	36.1
Other %	11.4	12.7	11.2	10.0
Maternal Education				
Mean (years)	14.5	14.7	14.5	14.2
SD	2.4	2.5	2.4	2.4
Paternal Education				
Mean (years)	14.2	14.3	14.2	14.1
SD	2.7	2.8	2.7	2.8

Race Other = Asian, Native American, Hawaiian Pacific Islander, and Multi-Racial. SD = standard deviation. The proportion of females in each proband age group increases with age (Overall Pearson chi-square = 89.4,  $df = 2$ ,  $p < 0.001$ ; pairwise chi-squares, all  $p$ 's  $< 0.001$ ). The proband age groups are balanced by race (Overall Pearson chi-Square = 5.6,  $df = 2$ , n.s.). Overall ANOVAs testing the significance of parental education were significant (maternal  $F = 4.6$ ,  $p < 0.01$ ; paternal  $F = 22.6$ ,  $p < 0.001$ ). However, inspection of the means shows that all the parental groups have a mean of 14 years of education and that the differences between the groups are within a few decimal places.

with age. Notably, the sample includes a large number of African American (32.9%) and other non-European American (11.4%) youths, and the age groups are balanced by race. Youths in the 11–17 age group constitute more than half of the sample (54.9%;  $n = 5,214$ ). Although the age groups differed in both maternal and paternal education, inspection of the means shows that all the parental groups have a mean of 14 years of education, and that the differences between the groups are within a few decimal points (likely significant due to the large sample size).

**Factor analysis of the GOASSESS.** *Exploratory factor analysis:* Table 2a shows the unidimensional, 2-, 3-, and 4-factor exploratory factor analysis solutions for the GOASSESS. To evaluate the fit of the models, we applied recommended cutoff values (Hu & Bentler, 1999) as follows: comparative fit index (CFI) = 0.95; root-mean-square error of approximation (RMSEA) = 0.06; and standardized root-mean residual (SRMR) = 0.08. Missing data were dealt with by pairwise deletion (final sample = 9,361). The fit of the unidimensional model was moderate to poor. Although its RMSEA was good ( $0.040 \pm 0.002$ ), the unidimensional model's CFI and SRMR were not quite acceptable (0.89 and 0.098, respectively) (Hu & Bentler, 1999). This borderline fit suggested that a multifactorial solution was necessary to adequately describe the latent structure of the GOASSESS. However, large loadings of the unidimensional model (mean = 0.57) suggest a single strong latent factor underlying the structure, and the 2-factor model in Table 2a was consistent with this suggestion. When an additional

factor was extracted, the behavioral disorders disaggregated to form their own factor, while the mood and anxiety disorders remained together. Oppositional defiant disorder exerted an extreme influence on determining the 'behavior' factor. The fit of the 2-factor model was within the acceptable range (CFI = 0.98; RMSEA =  $0.021 \pm 0.002$ ; SRMR = 0.051), suggesting additional factors might be unnecessary to model GOASSESS. However, if 'anxious-misery' and 'fear' are distinct phenomena, extraction of an additional factor would be required.

The 3-factor model (Table 2a) largely confirms that three factors are unnecessary to adequately explain the correlations among disorders. The fit of the model was excellent (CFI = 0.99; RMSEA =  $0.016 \pm 0.003$ ; SRMR = 0.042), but one of the goals of using fit indices in exploratory factor analysis is to determine the minimum number of factors necessary to explain covariance. By that standard, three factors were too many. However, examination of the factor loadings weakly supports the existence of a third, fear-related factor: the three phobias (specific, social, and agoraphobia) disaggregated to form their own factor. Social phobia and agoraphobia both had salient cross-loadings on the 'anxious-misery' factor. Finally, Factor 3 clearly remained a well-defined 'behavior' factor. Due to the possible overextraction of the 3-factor model (indicated by the acceptable fit of the 2-factor model), conclusions cannot be drawn from the 4-factor model.

Notably, the psychosis spectrum had numerous cross-loadings in multifactorial solutions. We therefore re-ran the analyses after decomposing the

**Table 2** (a) Unidimensional, 2-, 3-, and 4-factor exploratory solutions of GOASSESS (with oblique rotation). Unidimensional solution and exploratory 2-, 3-, and 4-factor oblique solutions of the psychosis spectrum screening tools

Scale	Uni	2-Factor		3-Factor		4-Factor				
		F1	F2	F1	F2	F1	F2	F3	F4	
(a)										
Post-traumatic stress disorder	.55	.53		.51		.34	.20			.23
Mania	.57	.49		.40		.30	.25			
Major depressive episode	.67	.73		.78		.68				
Eating disorders	.48	.59		.67		.54				-.21
Generalized anxiety disorder	.59	.66		.61		.82				
Separation anxiety disorder	.40	.49		.37	.20	.43	.54	.21		
Specific phobia	.40	.39		.39	.57					
Social phobia	.51	.52		.29	.38					
Panic disorder	.59	.75		.69		.55	.23			
Agoraphobia	.63	.60		.33	.45		.78			
Obsessive-compulsive disorder	.64	.67		.63		.58				
Attention deficit hyperactivity disorder	.47		.57		.23			.75		
Oppositional defiant disorder	.70		.91					.23		
Conduct disorder	.70		.78							.67
Psychosis spectrum	.63	.38	.36	.31		.24	.21			.32

GOASSESS Factor correlations (phi matrices)

Summary measure	Tool	Item	2-Factor				3-Factor				4-Factor				
			Uni	F1	F2	F3	Uni	F1	F2	F3	Uni	F1	F2	F3	F4
(b)															
Auditory Perceptions	PS-R	I have had the experience of hearing faint or clear sounds of people mumbling/talking when no one is nearby	.77	.68	.78										
Reality Confusion	PS-R	I think that I may get confused at times whether something I experience or perceive may be real or may just be part of my imagination or dreams	.79	.75	.74										
Mind Tricks	PS-R	I think that I might feel that mind is 'playing tricks' on me	.79	.73	.73										
Hallucination Screen	K-SADS		.55	.39	.28										
Odd/Unusual Thoughts	PS-R	I think that I have felt that there are odd or unusual things going on that I can't explain	.73	.66	.68										
Audible Thoughts	PS-R	I think that I may hear my own thoughts being said out loud	.75	.76	.59										
Delusion Screen	K-SADS		.45	.31	.24										
Superstitions	PS-R	I have had the experience of doing something differently because of my superstitions	.72	.74	.50										

**Table 2** (continued)

Summary measure	Tool	Item	2-Factor		3-Factor			4-Factor				
			Uni	F1	F2	F1	F2	F3	F1	F2	F3	F4
Thought Control	PS-R	I may have felt there is something interrupting or controlling my thoughts, feelings, or actions	<b>.77</b>	<b>.75</b>	<b>.48</b>	.34			.37	<b>.42</b>		
Suspicious	PS-R	I wonder if people may be planning to hurt me or even may be about to hurt me	<b>.72</b>	<b>.65</b>	<b>.47</b>	.25			.31	<b>.32</b>		.20
Predict Future	PS-R	I think that I might be able to predict the future	<b>.67</b>	<b>.76</b>		<b>.75</b>				<b>.83</b>		
Grandiosity	PS-R	I believe that I have special natural or supernatural gifts beyond my talents and natural strengths	<b>.72</b>	<b>.82</b>		<b>.70</b>				<b>.86</b>		
Mind Reading	PS-R	I have thought that it might be possible that other people can read my mind or that I can read others' minds	<b>.71</b>	<b>.78</b>	.29	<b>.52</b>				<b>.62</b>		
Occupational Functioning	SOPS	-	<b>.42</b>		<b>.69</b>					<b>.73</b>		<b>.69</b>
Expression of Emotion	SOPS	-	<b>.35</b>		<b>.60</b>					<b>.65</b>		<b>.67</b>
Experience of Emotions/Self	SOPS	-	<b>.50</b>		<b>.63</b>				.24	<b>.59</b>		<b>.69</b>
Trouble with Focus/Attention	SOPS	-	<b>.41</b>		<b>.55</b>			.26	-.24	<b>.53</b>		<b>.48</b>
Disorganized Communication	SOPS	-	<b>.49</b>		<b>.48</b>			.22		<b>.46</b>		<b>.43</b>

Psychosis spectrum factor intercorrelations (Phi matrices)

	F1	F2	F1	F2	F3	F1	F2	F3	F4
F1	1								
F2	.48	1							
F3	.51	.26	1						
F4	.47	.25	.37	1					

*N* = 9,361. Loadings <.20 removed; rotation = oblimin; F = Factor; Uni = unidimensional; maximin loading (per item, per solution) bolded; solutions estimated using the mean- and variance-adjusted weighted least squares (wlsmv) estimator in Mplus (version 7)(Muthén & Muthén, 2013). K-SADS = Kiddie Schedule for Affective Disorders and Schizophrenia. SOPS = Scale of Prodromal Symptoms. PS-R = PRIME Screen Revised. All PS-R items are self-rated from 0 = Definitely Disagree to 6 = Definitely Agree. For part a, unidimensional (1-factor) comparative fit index (CFI) = 0.89, standardized root-mean residual (SRMR) = 0.098, root-mean-square error (RMSEA) = 0.040; 2-factor CFI = 0.98, SRMR = 0.051, RMSEA = 0.021; 3-factor CFI = 0.99, SRMR = 0.042, RMSEA = 0.016; 4-factor fit not assessed due to likely overextraction; for part b, unidimensional (1-factor) CFI = 0.94, SRMR = 0.074, RMSEA = 0.064; 2-factor CFI = 0.98, SRMR = 0.039, RMSEA = 0.027, RMSEA = 0.029; 3-factor CFI = 0.99, SRMR = 0.027, RMSEA = 0.029; 4-factor CFI = 1.00, SRMR = 0.019, RMSEA = 0.02. All chi-square difference tests comparing the 1-, 2-, 3-, and 4-factor models were significant at the *p* < 0.005 level (see Table S5).



psychosis summary variable into its constituent positive-psychosis, positive-subpsychosis and negative/disorganized variables. Negative/disorganized symptoms loaded with behavior disorders, whereas the positive-psychosis and positive-subpsychosis variables both loaded exclusively with a factor composed of mania, the phobias, and post-traumatic stress disorder (see Table S3). However, there were more cross-loadings than in the model shown in Table 2a, rendering interpretation slightly more difficult.

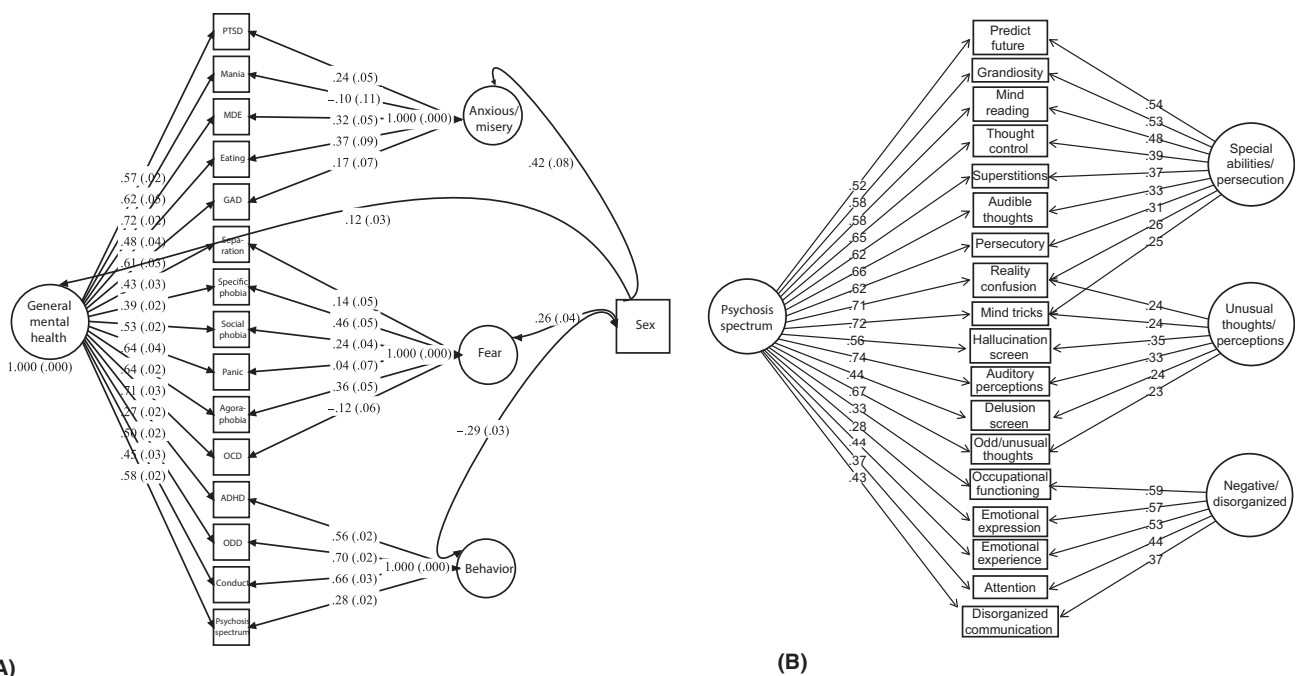
Finally, to further address criterion validity, we calculated factor scores using the 3-factor model shown in Table 2a and correlated them with neurocognitive summary measures (Moore, Reise, Gur, Hakonarson, & Gur, 2014). The ‘behavior’ factor had a small-to-moderate negative correlation (−0.21) with overall accuracy of neurocognitive performance, particularly in tests of executive and complex reasoning (−0.21). Smaller negative correlations across neurocognitive domains were also observed for the ‘anxious-misery’ and ‘fear’ factors (range = −0.03 to −0.19) (see Table S4).

**Confirmatory bifactor analysis:** Figure 3A shows the 3-factor bifactor solution of GOASSESS, including sex as a covariate; the fit of the model was excellent (CFI = 0.98; RMSEA = 0.019 ± 0.002). The general factor was very strong (mean loading = 0.54), with major depressive episode and obsessive-compulsive disorder having the highest loadings (0.72

and 0.71, respectively), and ADHD and specific phobia having the lowest loading (0.27 and 0.39, respectively). After controlling for the general factor, some disorders were no longer associated with their specific factors: mania no longer loaded on the ‘anxious-misery’ factor, panic disorder no longer loaded on the ‘fear’ factor, and obsessive-compulsive disorder actually had a significant *negative* loading on the ‘fear’ factor. The ‘behavior factor’ remained well determined (mean loading = 0.55) after controlling for general ‘mental health.’ Psychosis spectrum weakly loaded on this ‘behavior’ factor. There was a significant decrease in loadings on the ‘anxious/misery’ and ‘fear’ factors in Figure 3A (compared to Table 2a), suggesting they are central to the general construct of mental health. There were strong associations with participant sex: being female had a high association with ‘anxious-misery’ and low association with ‘behavior’.

**Factor analysis of psychosis spectrum screening tools**

Table 2b presents unidimensional, 2-, 3-, and 4-factor solutions of the psychosis spectrum screening tools. The SOPS Avolition scale was dropped from the factor analysis because it resulted in a Heywood case (communality >1.0; Heywood, 1931). There were two clear factors for symptom type (positive vs. negative/disorganized symptoms) in all solutions except the unidimensional, and 3-factor solution revealed a clear second positive symptom factor



**Figure 3** (A) Confirmatory bifactor analysis of the GOASSESS with sex as a covariate ( $n = 9,361$ ). PTSD = post-traumatic stress; MDE = major depressive episode; eating = anorexia/bulimia; GAD = generalized anxiety; separation = separation anxiety; panic = panic; OCD = obsessive compulsive; ADHD = attention deficit/hyperactivity; ODD = oppositional defiant. (B) Exploratory Bifactor (Schmid-Leiman) rotation of the psychosis spectrum screening tools in adolescents and young adults ( $n = 6,963$ ). See Table 2 for items/tools associated with each variable

determined by 'ideas about unusual abilities'. When the 3-factor solution was expanded to a fourth factor, interpretation of the item groupings became less clear. Factor 2 was made up entirely of weak loadings, and only one of those nine loadings (odd/unusual thoughts) was the dominant loading for that item, indicating that the 4-factor solution was likely overextracted (i.e. Factor 2 is a factor of cross-loadings). Finally, the correlations among the factors ( $\phi$  matrices) were moderate to high, reflecting a general factor explaining correlations among all items.

The results suggest that: (a) a 3-factor model is likely the most appropriate for the psychosis spectrum measures and (b) a general factor should be included to account for overall psychosis spectrum. We therefore estimated a bifactor model, shown in Figure 3B as an exploratory bifactor (Schmid-Leiman) orthogonalization (Schmid & Leiman, 1957) where each item loads on a specific factor and a general factor (with all factors orthogonal). This shows the direct effects of the highest-order construct (the general factor – 'psychosis spectrum') on each item, allowing more accurate judgments about the quality of each item. One example is 'auditory perceptions', which in the unidimensional factor model (Table 2b) had a high loading of 0.77, but 'reality confusion' and 'mind tricks' had higher loadings. In the bifactor model (Figure 3B), however, 'auditory perceptions' had the highest loading on the general factor. This indicated that 'reality confusion' and 'mind tricks' had inflated loadings in the unidimensional model (Table 2), and revealed auditory perceptions as the item with the highest discriminability on the scale. After accounting for the general psychosis spectrum factor, there were two positive symptom factors – ideas about 'special abilities/persecution'; 'unusual thoughts/perceptions' – and a 'negative/disorganized factor' (Figure 3B).

## Discussion

This overview of the recruitment and clinical assessment of the Philadelphia Neurodevelopmental Cohort (PNC) elucidates a strategy for accruing and psychiatrically assessing a large U.S. community sample of young people. The PNC was ascertained over a 2-year period in response to an unprecedented opportunity to characterize clinical and neurobehavioral phenotypes in previously genotyped youths from the community. There are several noteworthy features of the PNC. First, the distribution of participants across a wide age range (age 8–21, mean age = 14.2) allows powerful cross-sectional comparisons of age-group-related differences in neurobehavioral domains (Gur et al., 2012, 2014). Second, being drawn from the greater Philadelphia area, the sample is racially diverse (56% European American; 44% African American and

Other) and includes youths from surrounding urban, suburban, and rural areas. Third, youths were initially recruited from a wide range of pediatric clinics and thus include those who are healthy and those with a wide range of medical conditions, which allows comprehensive evaluation of associations among various medical conditions and mental disorders (Merikangas et al., 2015). Notably, participants were not recruited from psychiatric services.

In the PNC, we implemented a structured assessment tool with the aim of screening for a comprehensive array of psychopathology domains, including psychotic and subpsychotic symptoms. Several characteristics of the structure of GOASSESS emerged. First, there is a strong general 'mental health' factor, on which major depressive episode and OCD have the highest loadings, and ADHD and specific phobia have the lowest. Such a strong general factor of psychopathology is consistent with prior studies reporting high inter-factor correlations or strong loadings on a higher-order factor, or both (Beesdo-Baum et al., 2009). Moreover, it is consistent with Caspi et al. (2014) recent findings of a general psychopathology factor reflecting comorbidity among all disorders above and beyond their specific disorder strata. This suggests that each individual can be assigned a general 'mental health' score based on GOASSESS. Second, after controlling for the general factor, there are specific 'anxious-misery', 'fear', and 'behavior' factors as previously reported (Kessler et al., 2012). However, some disorders lose their associations with their specific factors, including mania, panic disorder and obsessive-compulsive disorder. Conversely, the 'behavior' factor remains fairly well determined. Overall, the results are consistent with distinct dimensions of mental disorders that are sometimes overwhelmed by a general 'mental health' factor. A partial exception, however, might be the psychosis spectrum, indicated by its numerous cross-loadings in multifactorial solutions (Table 2a). Prior studies of adults using confirmatory factor analysis and a greater number of psychosis variables have suggested that psychosis may be a distinguishable behavioral dimension from the internalizing and externalizing dimensions (Kotov, Chang et al., 2011; Kotov, Ruggero et al., 2011). Our model accounting for the general 'mental health' factor shows a summary psychosis spectrum measure in youths loading only weakly on the behavior factor, and strongly on the general mental health factor. The associations of the specific disorders with participant sex are consistent with existing literature showing higher rates of externalizing ('behavior') disorders in males, and internalizing ('anxious-misery' and 'fear') disorders in females (Merikangas et al., 2010). Finally, psychopathology factor scores are negatively correlated with dimensions of neurocognitive functioning.

The large community surveys in the United States are built on a tradition of ascertaining DSM diagnostic criteria for mental disorders. Therefore, the diagnostic modules closely adhere to current diagnostic nomenclature and do not include assessment of subthreshold psychosis spectrum symptoms (Merikangas et al., 2009). The assessment of these symptoms in the PNC therefore provides a unique dataset. In the current analyses, a bifactor model revealed a well-defined general psychosis spectrum factor and three specific factors: ideas about 'special abilities/persecution'; 'unusual thoughts/perceptions'; and 'negative/disorganized' symptoms. Notably, the loading of 'auditory perceptions' with hallucinations and delusions suggests that it taps a similar dimension of the psychosis spectrum as threshold hallucinations. Moreover, its highest loading on the general factor is consistent with reports that unusual auditory perceptions are generally predictive of psychotic symptoms (Kelleher et al., 2012; Laurens, Hobbs, Sunderland, Green, & Mould, 2012). Negative symptoms, understudied in young people (Dominguez, Saka, Lieb, Wittchen, & van Os, 2010), formed a discrete factor with disorganized symptoms. While the differentiation of positive from negative/disorganization symptoms is commensurate with prior studies of psychotic features in schizophrenia (Stefanis et al., 2002) and psychosis proneness (Therman et al., 2011), the lack of distinct negative and disorganized factors here may reflect the small number of contributing items. The results nevertheless suggest that there are separable yet related dimensions of the psychosis spectrum in this young cohort. As suggested by the literature on schizotypy in young people, these dimensions may be differentially associated with varying developmental trajectories and outcomes (Kwapil, Gross, Silvia, & Barrantes-Vidal, 2013).

Overall, it appears that the GOASSESS and psychosis spectrum screening tools have good structural validity, as established by exploratory and confirmatory factor analyses. Moreover, the correlations of the bifactor group factors with sex establish some nomological validity by suggesting GOASSESS is associated in the theoretically predicted way with measures of different but related constructs.

### Limitations

Although we made efforts through our recruitment approach to enhance the representativeness of the cohort, initial ascertainment through pediatric clinics may affect the generalizability of the findings to the youths in the United States or elsewhere. Moreover, our results suggest that the enrolled group is representative of the original recruitment pool in age and sex, but not race. There are several possible reasons for this disparity including factors associated with living in an urban

environment (e.g. decreased stability of housing/telephone numbers), which are not independent of race in the PNC. A further analysis of environmental contributors is beyond the scope of this paper, but is currently underway. We are also currently examining geocoded data to check the extent to which participants are representative of the greater Philadelphia area, as well as the extent to which sample demographics differ from those of the general U.S. population.

There are some limitations of the psychopathology assessment that should be considered when interpreting findings, or using data, from the PNC. First, the highly structured nature of the GOASSESS was necessary given the time frame and scope of the study. Although factor-analytic results appear consistent with prior findings, the structured and abbreviated format may have reduced sensitivity to clinically significant symptoms or conversely, yielded a high level of false positives for particular symptoms or disorders. Similarly, episode level information for some domains was gathered in the absence of information about their temporal relationships with other domains, but knowledge of these relationships is required for differential diagnosis among many disorders (e.g. mood disorders vs. psychotic disorders). Thus, the results may not be comparable to studies adhering to strict DSM-IV disorder criteria.

Second, while both collateral and proband interviews were conducted for probands aged 11–17, only collateral and only proband interviews are available for the 8- to 10- and 18- to 21-year-old probands, respectively. Our analyses used just one of many possible strategies to analyze data from different informants, but clearly the proband and collateral assessment data in the 11- to 17-year-old age range should be a rich source for alternative and informative parent–child comparisons (Hudziak, Achenbach, Althoff, & Pine, 2007). Any analyses of differences among various age groups (e.g. 8–10, 11–17, 18–21) must be conducted with recognition of the corresponding differences in informants across these groups.

Finally, the cross-sectional nature of the assessment imposes inherent limitations on developmental inferences about psychopathology and its intra-individual trajectories. Cross-sectional designs make it difficult to differentiate true age-related changes from apparent age differences that are actually due to other factors, such as cohort factors. In the PNC, the recruitment over a very brief time span may limit the influence of intra-age group effects, but across age group cohort effects may exist. Additional clinical information obtained through longitudinal follow-up will allow further evaluation of clinical relevance and stability of the current findings.

There are also some limitations of our current analyses. First, because we aimed for consistency with studies that used DSM criteria for major

domains of psychopathology, we required the 'significant' cutoff level to include significant distress or impairment reported by the proband (or collateral). However, there are multiple ways the data could be analyzed and many possible cutoffs; classification choices will depend on the questions being asked. Similarly, we selected factor analysis, but a number of data reduction methods exist and can be applied to the GOASSESS data. The computerization of GOASSESS afforded the opportunity to capture and archive individual item responses that will permit alternative approaches to diagnostic classification and analyses. Combined with PNC neurocognitive and neuroimaging data, they also offer a rich dataset for dimensional approaches such as those embodied in the NIMH Research Domain Criteria (RDoC) (Insel, 2014). Second, in factor analysis, it is difficult to specify fit-index cutoff values that apply across a wide range of models. We opted to follow the recommendation in the most commonly cited simulation study of fit indices (Hu & Bentler, 1999), but alternative cutoffs are possible (Marsh, Hau, & Wen, 2004; Sivo, Fan, Witta, & Willse, 2006) and can be applied in the future. In addition, given evidence that the factor structure of psychopathology may vary across age groups (Wittchen et al., 2009), future analyses can focus on the developmental stability of observed comorbidities.

## Conclusion

Advances in genomics are revolutionizing medicine with discoveries that help elucidate mechanisms and design novel treatments. For mental illnesses to benefit from genomics, data are needed linking behavior to brain function in large prospective samples. In a diverse cohort of nearly 9,500 genotyped youths, we have taken the first steps to characterize neurodevelopmental phenotypes – clinical, neurocognitive and, in a subsample, neuroimaging – creating a landmark dataset that we hope will propel understanding and treatment of developmental neuropsychiatric disorders and allow their genomic underpinnings to be dissected in future studies. Factor-analytic results of the clinical data are consistent with the suggestion that studies of the relationships among psychopathology, genomics, and biobehavioral measures (CNB, imaging) may increase efficiency by focusing on groups of comorbid disorders with shared general factors, rather than on individual disorders (Kendler, Prescott, Myers, & Neale, 2003). Collaborative efforts to analyze the rich, accessible, PNC

resource will accelerate research on the complex inter-relationships among genes, cognition, brain, and behavior involved in the neurodevelopment of common mental disorders (National Institute of Health, 2014).

## Supporting information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1.** PNC recruitment and psychopathology assessment.

**Table S1.** Demographic characteristics of recruitment pool ( $n = 19,161$ ) by enrollment status.

**Table S2.** GOASSESS sections and informants.

**Table S3.** Unidimensional, 2-, 3-, and 4-factor models of the GOASSESS with the summary psychosis variable decomposed into positive (Psychosis; subpsychosis) and negative/disorganized symptoms.

**Table S4.** Correlations of exploratory clinical factor scores with neurocognitive factor scores.

**Table S5.** Chi-Square-based model comparison.

**Figure S1a.** Sample GOASSESS proband screener section.

**Figure S1b.** Stimulus cards shown during GOASSESS specific phobia section.

**Figure S1c.** Distress/impairment cards shown during GOASSESS.

## Acknowledgements

This work was supported by the National Institute of Mental Health (R.E.G. and H.H., RC2 grant numbers MH089983 and MH089924), (M.E.C., grant number K08MH079364), and (T.D.S., grant number K23MH098130); the Dowshen Program for Neuroscience; and the Marc Rapport Family Investigator grant through the Brain and Behavior Foundation (T.D.S.). The authors thank the participants of this study, and all the members of the Recruitment, Assessment, and Data Teams whose individual contributions collectively made this work possible. Also William G. Iacono and Genevieve Ryczek (University of Minnesota Center for Twin and Family Research; MCFTR) for their consultation while developing our recruitment strategy, and Iacono for allowing use, modification, and computerization of the MCFTR substance use measure. The authors have declared that they have no competing or potential conflicts of interest.

## Correspondence

Monica E. Calkins, Department of Psychiatry, Neuropsychiatry Section, Perelman School of Medicine at the University of Pennsylvania, 9 Maloney, 3600 Spruce Street, Philadelphia, PA 19104, USA; Email: mcalkins@upenn.edu

## Key points

- The Philadelphia Neurodevelopmental Cohort (PNC) of nearly 9,500 young people (ages 8–21) was established as a public domain resource of clinical and neurobehavioral phenotypes from a community cohort of previously genotyped youths.
- This paper provided an overview of PNC recruitment and clinical assessment methods to allow informed use and interpretation of PNC data for the scientific community.
- Results of factor analysis employed to organize the common processes underlying psychopathology yielded psychopathology data with strong factorial validity.
- The PNC resource should enable advancement of understanding of the inter-relationships among genes, cognition, brain, and behavior involved in the neurodevelopment of common mental disorders.

## References

- Beesdo-Baum, K., Hofler, M., Gloster, A.T., Klotsche, J., Lieb, R., Beauducel, A., ... & Wittchen, H.U. (2009). The structure of common mental disorders: A replication study in a community sample of adolescents and young adults. [Research Support, Non-U.S. Gov't]. *International Journal of Methods in Psychiatric Research*, 18, 204–220.
- Binbay, T., Drukker, M., Elbi, H., Tanik, F.A., Ozkinay, F., Onay, H., ... & Alptekin, K. (2012). Testing the psychosis continuum: Differential impact of genetic and nongenetic risk factors and comorbid psychopathology across the entire spectrum of psychosis. [Research Support, Non-U.S. Gov't]. *Schizophrenia Bulletin*, 38, 992–1002.
- Calkins, M.E., Moore, T.M., Merikangas, K.R., Burstein, M., Satterthwaite, T., Bilker, W.B., ... & Gur, R.E. (2014). The psychosis spectrum in a young community sample: Findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry*, 13, 296–305.
- Caspi, A., Houts, R., Belsky, D.W., Goldman-Mellor, S., Harrington, H.L., Israel, S., ... & Moffitt, T.E. (2014). The 'p factor': One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, 2, 119–137.
- Dominguez, M.D., Saka, M.C., Lieb, R., Wittchen, H.U., & van Os, J. (2010). Early expression of negative/disorganized symptoms predicting psychotic experiences and subsequent clinical psychosis: A 10-year study. Research Support, Non-U.S. Gov't]. *American Journal of Psychiatry*, 167, 1075–1082.
- dbGaP. (2014). Neurodevelopmental Genomics: Trajectories of Complex Phenotypes. Retrieved February 6, 2015, from [http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study\\_id=phs000607.v1.p1](http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v1.p1)
- Gur, R.C., Calkins, M.E., Satterthwaite, T.D., Ruparel, K., Bilker, W.B., Moore, T.M., ... & Gur, R.E. (2014). Neurocognitive growth charting in psychosis spectrum youths. *JAMA Psychiatry*, 71, 366–374.
- Gur, R.C., Richard, J., Calkins, M.E., Chiavacci, R., Hansen, J.A., Bilker, W.B., ... & Gur, R.E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology*, 26, 251–265.
- Han, C., McGue, M.K., & Iacono, W.G. (1999). Lifetime tobacco, alcohol and other substance use in adolescent Minnesota twins: Univariate and multivariate behavioral genetic analyses. *Addiction*, 94, 981–993.
- Heywood, H.B. (1931). On finite sequences of real numbers. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 134, 486–501.
- Holzinger, K.J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2, 41–54.
- Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Hudziak, J.J., Achenbach, T.M., Althoff, R.R., & Pine, D.S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research*, 16(Suppl. 1), S16–S23.
- Insel, T.R. (2014). The NIMH Research Domain Criteria (RDoC) project: Precision medicine for psychiatry. *The American Journal of Psychiatry*, 171, 395–397.
- Kelleher, I., Connor, D., Clarke, M.C., Devlin, N., Harley, M., & Cannon, M. (2012). Prevalence of psychotic symptoms in childhood and adolescence: A systematic review and meta-analysis of population-based studies. *Psychological Medicine*, 42, 1857–1863.
- Kendler, K.S., Prescott, C.A., Myers, J., & Neale, M.C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, 60, 929–937.
- Kessler, R.C., Avenevoli, S., McLaughlin, K.A., Green, J.G., Lakoma, M.D., Petukhova, M., ... & Merikangas, K.R. (2012). Lifetime co-morbidity of DSM-IV disorders in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A). *Psychological Medicine*, 42, 1997–2010.
- Kobayashi, H., Nemoto, T., Koshikawa, H., Osono, Y., Yamazawa, R., Murakami, M., ... & Mizuno, M. (2008). A self-reported instrument for prodromal symptoms of psychosis: Testing the clinical validity of the PRIME Screen-Revised (PS-R) in a Japanese population. *Schizophrenia Research*, 106, 356–362.
- Kotov, R., Chang, S.W., Fochtmann, L.J., Mojtabai, R., Carlson, G.A., Sedler, M.J., & Bromet, E.J. (2011). Schizophrenia in the internalizing-externalizing framework: A third dimension? *Schizophrenia Bulletin*, 37, 1168–1178.
- Kotov, R., Ruggero, C.J., Krueger, R.F., Watson, D., Yuan, Q., & Zimmerman, M. (2011). New dimensions in the quantitative classification of mental illness. *Archives of General Psychiatry*, 68, 1003–1011.
- Krueger, R.F. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, 56, 921–926.
- Kwapil, T.R., Gross, G.M., Silvia, P.J., & Barrantes-Vidal, N. (2013). Prediction of psychopathology and functional impairment by positive and negative schizotypy in the Chapmans' ten-year longitudinal study. *Journal of Abnormal Psychology*, 122, 807–815.
- Laurens, K.R., Hobbs, M.J., Sunderland, M., Green, M.J., & Mould, G.L. (2012). Psychotic-like experiences in a community sample of 8000 children aged 9 to 11 years: An item response theory analysis. *Psychological Medicine*, 42, 1495–1506.
- Marsh, H.W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting

- cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11, 320–341.
- McGlashan, T.H., Miller, T.J., Woods, S.W., Rosen, J.L., Hoffman, R.E., & Davidson, L. (2003). *Structured interview for prodromal syndromes, Version 4.0*. New Haven, CT: Prime Clinic Yale School of Medicine.
- Merikangas, K., Avenevoli, S., Costello, J., Koretz, D., & Kessler, R.C. (2009). National comorbidity survey replication adolescent supplement (NCS-A): I. Background and measures. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48, 367–369.
- Merikangas, K.R., Calkins, M.E., Burstein, M., He, J.P., Chiavacci, R., Lateef, T., ... & Gur, R.E. (2015). Comorbidity of physical and mental disorders in the Neurodevelopmental Genomics Cohort Study. *Pediatrics*, 135, e927–e938.
- Merikangas, K.R., He, J.P., Burstein, M., Swanson, S.A., Avenevoli, S., Cui, L., ... & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, 49, 980–989.
- Miller, T.J., Cicchetti, D., Markovich, P.J., McGlashan, T.H., & Woods, S.W. (2004). The SIPS screen: A brief self-report screen to detect the schizophrenia prodrome. *Schizophrenia Research*, 70(Suppl. 1), 78.
- Moore, T.M., Reise, S.P., Gur, R.E., Hakonarson, H., & Gur, R.C. (2014). Psychometric properties of the penn computerized neurocognitive battery. *Neuropsychology*, 29, 235–246.
- Morris, N.M., & Udry, J.R. (1980). Validation of a self-administered instrument to assess stage of adolescent development. *Journal of Youth and Adolescence*, 9, 271–280.
- Muthén, L.K., & Muthén, B.O. (2013). *Mplus user's guide* (7th edn). Los Angeles, CA: Muthén & Muthén.
- National Institute of Health (2014). RFA-MH-15-400: Leveraging a Recovery Act Resource to Accelerate Research on Neurodevelopment (R01) Retrieved June 5, 2014, from <http://grants.nih.gov/grants/guide/rfa-files/RFA-MH-15-400.html>
- Rapoport, J.L., Giedd, J.N., & Gogtay, N. (2012). Neurodevelopmental model of schizophrenia: Update 2012. [Review]. *Molecular psychiatry*, 17, 1228–1238.
- Reise, S.P., Moore, T.M., & Haviland, M.G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544–559.
- Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M.E., ... & Gur, R.E. (2014). Neuroimaging of the Philadelphia neurodevelopmental cohort. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. *NeuroImage*, 86, 544–553.
- Satterthwaite, T.D., Vandekar, S., Wolf, D.H., Ruparel, K., Roalf, D.R., Jackson, C., ... & Gur, R.C. (2014). Sex differences in the effect of puberty on hippocampal morphology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53, 341–350 e341.
- Schmid, J., & Leiman, J.M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53–61.
- Sivo, S.A., Fan, X., Witta, E.L., & Willse, J.T. (2006). The search for 'optimal' cutoff properties: Fit index criteria in structural equation modeling. *Journal of Experimental Education*, 74, 267–288.
- Stefanis, N.C., Hanssen, M., Smirnis, N.K., Avramopoulos, D.A., Evdokimidis, I.K., Stefanis, C.N., ... & Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine*, 32, 347–358.
- Therman, S., Heinimaa, M., Miettunen, J., Joukamaa, M., Moilanen, I., Maki, P., & Veijola, J. (2011). Symptoms associated with psychosis risk in an adolescent birth cohort: Improving questionnaire utility with a multidimensional approach. *Early Intervention in Psychiatry*, 5, 343–348.
- Wilkinson, G.S., & Robertson, G.J. (2006). *Wide range achievement test* (4th edn). Lutz, FL: Psychological Assessment Resources.
- Wittchen, H.U., Beesdo-Baum, K., Gloster, A.T., Hofler, M., Klotsche, J., Lieb, R., ... & Kessler, R.C. (2009). The structure of mental disorders re-examined: Is it developmentally stable and robust against additions? [Research Support, Non-U.S. Gov't]. *International Journal of Methods in Psychiatric Research*, 18, 189–203.

Accepted for publication: 20 February 2015

First published online: 8 April 2015