Journal of Abnormal Psychology

An Item Response Theory Analysis of the Prodromal Questionnaire-Brief Child Version: Developing a Screening Form That Informs Understanding of Self-Reported Psychotic-Like Experiences in Childhood

Nicole R. Karcher, Michael T. Perino, Deanna M. Barch

Funded by:

- National Institute of Mental Health
- National Institute on Drug Abuse Grant
- National Institutes of Health

An IRT Analysis of the Prodromal Questionnaire-Brief Child Version: Developing a Screening Form that Informs Understanding of Self-Reported Psychotic-Like Experiences in Childhood

Nicole R. Karcher*, Michael T. Perino*, Deanna M. Barch

*These authors contributed equally

Washington University

Author Note

Correspondence to: Nicole R. Karcher, Ph.D., Department of Psychiatry, One Brookings Drive, Campus Box 1125, Washington University School of Medicine, St. Louis, MO 63130 email: nkarcher@wustl.edu

The manuscript contains original work that has not been published or submitted for publication elsewhere. All procedures were approved by the institutional review board at each Adolescent Brain Cognitive Development collection site (WUSTL IRB number: 201708123; https://abcdstudy.org/study-sites/).

Funding/Support: This work was supported by National Institute on Drug Abuse grant U01 DA041120 (DMB), National Institute of Mental Health grants MH014677 (NRK) and 2T32MH100019-06 (MTP).

Manuscript Word Count: 8,969

Abstract

The Prodromal Questionnaire-Brief Child Version (PQ-BC) has been developed as a tool for identifying psychotic-like experiences (PLEs) in school-age children. The current study examined the psychometric properties of the PQ-BC, examined how well the PQ-BC estimates the latent construct of PLEs ($\hat{\theta}$), and began the process of developing a screening form informed by item response theory (IRT). Utilizing the baseline (n=11,129) sample from the Adolescent Brain Cognitive Brain (ABCD) study, we examined which PQ-BC items provide the most information and best discriminate individuals experiencing PLEs. Using hierarchical linear models (HLMs), we found that $\hat{\theta}$ scores were significantly associated with several previously identified predictors of psychosis spectrum symptoms (i.e., history of psychosis, internalizing symptoms, cognitive impairments, developmental milestone delays, and resting-state functional connectivity impairments) at baseline and year 1 (n=5,532). Using item level information and discrimination parameters of the PQ-BC from the baseline sample, we created a seven-item screening form. HLMs generally found significant associations between screening form scores for both baseline and year 1 with the aforementioned predictors. The analyses provide evidence for the validity of a screening form derived from the PQ-BC using IRT derived parameters. This screening form could prove useful when the full measure is not feasible.

Keywords: item response theory; psychotic-like experiences; psychosis spectrum; screening form; nomological network

General Scientific Summary: The current study examined a measure of unusual thoughts and unusual perceptual experiences in school-age children, specifically

looking at which items in the measure were the most informative and most likely to be endorsed by individuals with higher levels of these experiences. Using this information, the current study developed a screening form of this measure and found that this screening form showed the expected associations with predictors of these unusual experiences (e.g., family history of psychosis, impaired cognition, delays in attaining motor and speech milestones, and symptoms of depression and anxiety).

Since its development, the Prodromal Questionnaire has been an important tool to assist in identifying individuals at-risk for the development of psychosis spectrum disorders (Loewy, Bearden, Johnson, Raine, & Cannon, 2005; Savill, D'Ambrosio, Cannon, & Loewy, 2018). Abbreviated versions of the Prodromal Questionnaire have been developed to shorten administration time while providing valid and reliable estimates of psychotic-like experiences (PLEs) (Loewy, Pearson, Vinogradov, Bearden, & Cannon, 2011). PLEs are subclinical psychotic symptoms, including subthreshold unusual beliefs and perceptual distortions. PLEs in childhood can be considered as unspecific though potentially sub-clinical expression of psychosis (Rimvall et al., 2019; van Os & Reininghaus, 2016; Wusten et al., 2018) experienced by approximately 13-17% of all children (Kelleher et al., 2012; Laurens et al., 2007). PLEs are thought to arise as a result of a combination of genetic, environmental, and pathophysiological factors (e.g., disruptions in connectivity), factors that can be indexed using predictors of psychosis spectrum symptoms, including family history of psychosis, cognitive functioning (and especially working memory), developmental milestone delays and motor impairments, and impairments in resting state functional connectivity (Cannon et al., 2002; Linscott & van Os, 2013; Satterthwaite et al., 2015; van Os, Linscott, Myin-Germeys, Delespaul, & Krabbendam, 2009). The current study examined the extent to which the latent construct of PLEs in school-age children is associated with these psychosis spectrum symptom predictors.

The most widely used short form of the Prodromal Questionnaire is the Prodromal Questionnaire-Brief (Loewy et al., 2011). This questionnaire has been extensively validated in a variety of populations, including both clinical and non-clinical

samples, and was designed as part of a two-step screening process in conjunction with a clinical assessment. In addition, the Prodromal Questionnaire-Brief was adapted for a childhood sample, titled the Prodromal Questionnaire-Brief Child Version (PQ-BC) (Karcher et al., 2018). The PQ-BC is a 21-item scale validated for use with 9-10-year-old children and used to identify PLEs in this non-clinical population. However, future community-based assessment efforts will likely necessitate briefer measures for screening children with self-reported psychotic-like experiences. The current study aimed to better understand the latent construct of psychotic-like experiences, examine the psychometric properties of the PQ-BC, and begin the process of developing a screening form using item response theory (IRT).

IRT (Lawley, 1943) refers to psychometric modeling techniques (Lord, 1953; Rasch, 1960; Samejima, 1969) used to develop, improve, and scale assessments of individual differences. IRT offers researchers detailed information about the items used in psychological assessments and their relationship to the underlying trait. IRT differs from classical test theory (CTT) techniques in the assumption that individual test items reflect varying levels of an underlying trait (Hambleton, Swaminathan, & Rogers, 1991). By accounting for information about both individual response profiles and scalar properties, IRT can provide researchers with significant advantages for understanding how psychological instruments tap into the latent constructs. For example, IRT techniques can provide information about item discriminability (i.e. how well a test item delineates low trait individuals from high trait individuals), difficulty (i.e. how high on a trait one must be to respond affirmatively to a particular question), and total trait information (i.e. how responses across the whole scale should be translated to estimate

an individual's true trait score). IRT methodology has been used estimate trait levels within individuals (Hays, Morales, & Reise, 2000) and identify items for the creation of short-form measures (Maples-Keller et al., 2019), just to name a few applications.

IRT applications may be particularly advantageous for trait assessment of clinically relevant phenomena (Reise & Waller, 2009); rather than treating all symptoms of clinical conditions as equally important, IRT allows for symptom presence (via responses on specific test items) to be weighted based on individual and population responses. Participants responses are scored based on individual and test characteristics, which translate to a theoretically more accurate score along the measured trait. Given that psychopathological assessments often focus on constructs situated at extreme values (Reise & Waller, 2009), the greater distribution at extrema of trait situations that IRT approaches provide is a particularly important assessment concern from a psychometric standpoint. IRT is a methodology well-suited for determining how well the PQ-BC assesses across the range of latent trait (θ) of PLEs, including the range of the trait that the PQ-BC is measuring. Thus, the current study will utilize IRT in order to examine how well the PQ-BC estimates the latent trait ($\hat{\theta}$) of psychotic-like experiences, as well as associations between $\hat{\theta}$ and predictors of psychosis spectrum symptoms, providing valuable insight about psychotic-like experiences in school-age children. In particular, it will help inform key questions such as the age at which PLEs can manifest, whether predicted associations with external variables are present even at ages 9 and 10, and the distribution of this trait in a large population of 9-10-year-olds. In addition, IRT will be used to create a PQ-BC screening form for efficiently identifying PLEs in school-age children.

The current study aimed to better understand the latent construct of PLEs in school-age children, and is the first to examine psychometric properties of the PQ-BC using the entire baseline sample from the Adolescent Brain and Cognitive Development (ABCD) study, as well as the first to analyze the PQ-BC using IRT analyses to begin the process of developing a screening form, including examining the psychometric properties of the screening form. Our goals were to (a) examine which PQ-BC items provide the most information and discriminate individuals experiencing PLEs, (b) examine the associations between the latent construct of PLEs ($\hat{\theta}$) and predictors of psychosis spectrum symptoms (i.e., parent-rated PLEs, internalizing symptoms, cognitive deficits, developmental milestone delays, resting-state functional connectivity) at baseline and year 1 of the ABCD data, and (c) use the information and discrimination parameters to assess the PQ-BC and derive a screening form.

Methods

Participants

A sample of 11,874 individuals was obtained from the ABCD study (Data Release 2.0.1), a large-scale study tracking 9-10-years-olds recruited from 21 research sites across the United States (see Supplement for study-wide exclusion criteria) (Barch et al., 2018). Institutional review board approval was obtained for each site before data collection. All parents provided written informed consent and all children provided assent.

These data were accessed from the National Institutes of Mental Health Data Archive (see Acknowledgments). Participants were removed from analyses in the case of missing data (baseline: n=744; year 1: n=237). The final sample size for baseline

analyses was n=11,129 individuals (47.9% female; 52.6% White, 20.1% Hispanic, 14.6% African American, 2.2% Asian, and 10.5% Other; year 1: final sample size: n=5,532; 47.7% female; 59.3% White, 18.1% Hispanic, 10.2% African American, 2.3% Asian, and 10.0% Other; note, only a subset of individuals were included for imaging data analyses: individuals with complete imaging data at baseline: n=8,859, individuals with complete imaging data at year 1: n=4,283).

Measures

Prodromal Questionnaire-Brief Child Version (PQ-BC)

Participants completed the Prodromal Questionnaire-Brief Child Version (PQ-BC) (Loewy et al., 2011), a 21-item self-report questionnaire with a visual response scale included as a distress scale (see Supplement for additional questionnaire details; (Hirshfeld-Becker, 2006). As previously mentioned, the PQ-BC demonstrates validity in the ABCD sample (Karcher et al., 2018). Consistent with previous research (Karcher et al., 2018), Total and Distress scores were calculated. The Total score is the sum of endorsed questions (i.e., 0=no, 1=yes). The Distress score is the total number of endorsed questions weighted by level of distress [i.e., 0=no, 1=yes (but no distress), 2-6=yes (1+score on distress scale)]. Consistent with previous work (Karcher et al., 2018), we also examined log-transformed distress scores ((Howell, 2007); formula=LG10(X + 1)). 61.0% of the baseline sample endorsed at least one PQ-BC question (46.2% for year 1), with 42.9% reporting distress associated with at least one PQ-BC question (28.8% for year 1). Both the PQ-BC Total (α=.86; year 1: α=.87) and Distress (α=.87; year 1: α=.87) scores showed high internal reliability, which did not increase when any item was deleted. The correlation coefficients between each item and the PQ-BC Total

score were 0.35–0.52 (year 1:0.35-0.55), and for the Distress Score, 0.38–0.54 (year 1:0.37–0.60).

Symptom Measures

The parent and youth versions of the validated and computerized Kiddie-Structured Assessment for Affective Disorders and Schizophrenia (KSADS) for DSM-5 (Kaufman et al., 1997; Kobak, Kratochvil, Stanger, & Kaufman, 2013) were used in current analyses as measures of psychopathology (Barch et al., 2018). We examined baseline internalizing symptoms using a composite of summations of current child-rated depression and generalized anxiety disorder (GAD) symptoms. A measure of parent-rated child PLEs was created from four items from the CBCL (Achenbach, 2009; Ducharme et al., 2015; Sheffield, Kandala, Burgess, Harms, & Barch, 2016). These questions were: “Hears sounds or voices that aren’t there,” “Sees things that aren’t there,” “Strange behavior,” and “Strange ideas.” Each question was scored from 0=not true, 1=somewhat or sometimes true, and 2=very true or often true.

History of psychosis in first-degree relatives was assessed at baseline using the parent-rated Family History Assessment Module Screener (Rice et al., 1995), with each scored as either present or absent. A financial adversity index was computed as the sum of binary responses to seven parent-rated questions of financial adversity from a demographic questionnaire (e.g., “Were evicted from your home for not paying the rent or mortgage?”; see Supplement for all questions).

Neuropsychological Test Battery

At baseline, participants completed all tests within the National Institutes of Health Toolbox Cognitive Battery (NIHTB-CB) (Weintraub et al., 2013). The NIHTB-CB

consists of 7 tasks, grouped into two composite scores. The fluid composite consists of Flanker Inhibitory Control and Attention, List Sorting Working Memory, Dimensional Change Card Sort, Pattern Comparison Processing Speed and Picture Sequence Memory. The crystallized composite consists of Picture Vocabulary and Oral Reading Recognition test (see Weintraub et al. for descriptions of individual NIHTB-CB tests (Weintraub et al., 2013)). Consistent with previous research using the ABCD sample to examine associations with PQ-BC scores, the current study examined associations with the fluid and crystallized composite scores, as well as the List Sorting Working Memory task (Karcher et al., 2018). The current study utilized uncorrected NIHTB-CB scores, but all analyses include age and sex as covariates.

Developmental Milestones

The parent assessment battery included questions regarding motor and speech development rated at baseline (Achenbach, 2009; Kessler et al., 2009). The motor delays composite was coded as the summation of delays in attaining motor milestones [scored as 0=no delay or 1=delay; individual milestones: rolling over (delay=6 months or later), sitting (delay=after 9 months), and walking (delay=after 18 months)], parent-rated concern regarding motor development (0=earlier, 1=average, 2=later), and parent-rated current child clumsiness (0=not true, 1=somewhat or sometimes true, and 2=very true or often true). The speech delays composite was coded as the summation of a delay in speaking first word (scored as 0=no delay or 1=delay; delay=after 12 months) and parent-rated concern regarding speech development (0=earlier, 1=average, 2=later). Consistent with previous research using the ABCD sample to examine associations with PQ-BC scores, the current study examined associations with the motor and speech

delays composite scores, as well as the parent-rated current child clumsiness (Karcher et al., 2018).

Resting State Functional Connectivity (RSFC)

Consistent with previous research using the ABCD sample to examine associations with PQ-BC scores (Karcher, O'Brien, Kandala, & Barch, 2019), we examined cingulo-opercular (CON) within-network connectivity, cingulo-parietal (CPAR) within-network connectivity, default mode (DMN) within-network connectivity, CON to cerebellar connectivity, and CPAR to cerebellar connectivity. In terms of collecting RSFC data, all children were run on a 3T scanner (either Siemens or General Electric; note, we followed recent guidance from ABCD about excluding the data from the Philips scanners) with a 32-channel head coil and completed T1-weighted and T2-weighted structural scans (1mm isotropic). Participants also completed four 5-minute resting-state BOLD scans, with their eyes open and fixated on a crosshair. Resting state images were acquired in the axial plane using an EPI sequence. Other resting-state image parameters varied by 3T scanner and have been previously detailed (https://abcdstudy.org/images/Protocol_Imaging_Sequences.pdf) (Casey et al., 2018). The data analysis pipeline has also been detailed previously (Karcher et al., 2019). Briefly, the aforementioned RSFC metrics were derived from regions of interest (ROIs) within functionally-defined parcellations (i.e., Gordon networks; (Gordon et al., 2016)) and subcortical ROIs (i.e., cerebellum; (Fischl et al., 2002)). The Fisher Z-transform of the correlation values were examined.

IRT Methodology

IRT models assume the latent trait being measured is unidimensional (Hambleton et al., 1991) or has a dominant first factor (Drasgow & Hulin, 1990), which the PQ-BC had satisfied with its previously established single factor structure (PQ-BC; Karcher et al., 2018). This is consistent with evidence from large [i.e., between n=449 (Fonseca-Pedrero, Gooding, Ortuno-Sierra, & Paino, 2016) to n=15,000 (Fonseca-Pedrero, Inchausti, Perez-Albeniz, & Ortuno-Sierra, 2018)] community samples of adolescents and adults for a one factor structure for the PQ-B (Cicero, Krieg, & Martin, 2017; Fonseca-Pedrero et al., 2016; Fonseca-Pedrero et al., 2018). Confirmatory factor analyses results based on the larger baseline sample (n=11,129) again demonstrated a single factor solution [for Total scores: CFI=0.929, RMSEA=0.026; for binned Distress scores (see below for binning procedure): CFI=0.923, RMSEA=0.023; see Supplement for additional details about factor analyses (Brown, 2014; Muthén, 2010; Rosseel, 2012)].

After passing this requirement, we assessed item performance by applying a two parameter logistic (2PL) graded response model using the ltm R package (Rizopoulos, 2006). 2PL models create item characteristic curves (ICC) for each test item, which are logistic functions that represent how responses on said item relates to its underlying trait (i.e. the amount of information the item provides), denoted as $\hat{\theta}$ (Edelen & Reeve, 2007), as well as item response functions (IRF) which can be used to assess estimation and parameter information. The two parameters assessed in such models – difficulty (or location/threshold) parameters denoted by *b* and discrimination (or slope of the IRF at *b*) parameter denoted by *a* – can be utilized to assess how well (or poorly) an item performs at measuring the underlying trait of interest, as well as individual strengths or

weakness of particular items. Discrimination parameters are often used to identify items which more successfully discriminate between those who do (or do not) possess the underlying trait, with high *a* items indicating stronger relationships with the underlying trait compared to lower *a* items. Threshold parameters reflect the point at which the probability of endorsing the trait is chance; therefore, higher *b* parameters serve as markers of difficulty, where larger values suggest respondents must possess higher trait thresholds for endorsement.

While 2PL models were originally suited for binary response choices (Reise & Waller, 2009), graded response modeling (GRM) represents a generalization of this approach translatable to polytomous scale data (Samejima, 1969). In our data, the raw response data was translated into an ordinal scale using the procedure addressed in the prior scale validation (Karcher et al., 2018). To increase the efficiency of item parameter estimation (Cho, Drasgow, & Cao, 2015; Stark, Chernyshenko, & Drasgow, 2006) we binned the seven response categories into three categories, [0=symptom not present (no to trait); 1=symptom present without (or mild) distress (yes to trait but reported either no distress or distress=2 or 3); 2= symptom present with distress (yes to trait and distress=4, 5, or 6)], a strategy highlighted as useful when incorporating items with low endorsement rates (Chernyshenko, Stark, Drasgow, & Roberts, 2007).

Statistical Analyses

Following IRT analyses, the remainder of the analyses used hierarchical linear models (HLMs). To account for nonindependence of observations due to familial relatedness, family members were treated as clustered observations, as were the 21 ABCD research sites. All HLM analyses were conducted in R lme4 package [(Bates,

Mächler, Bolker, & Walker, 2014); multcomp package for multiple comparison analyses; (Hothorn, Bretz, & Westfall, 2008)] with family unit and research site modeled as random intercepts, and age, sex, financial adversity, family history of psychosis, and ethnicity included as covariates. First of all, we examined the longitudinal stability from baseline to year 1 for all PQ-BC scores, including $\hat{\theta}$ scores (and differences between significant correlations were examined using Meng's z-test procedures (Meng, Rosenthal, & Rubin, 1992). HLMs analyzed the associations between $\hat{\theta}$ scores for both baseline and year 1 and the following predictors previously identified as significantly associated with PLEs in a subset of the ABCD baseline sample (Karcher et al., 2018): 1) parent-rated PLEs, 2) history of psychosis, 3) internalizing symptom composite, 4) neuropsychological test performance (crystallized or fluid intelligence Toolbox composites, then the working memory subset) 5) motor and speech developmental milestone delays (composites, then the individual clumsiness items), 6) RSFC metrics (CON within-network connectivity, CPAR within-network connectivity, DMN within-network connectivity, CON to cerebellar connectivity, and CPAR to cerebellar connectivity). These estimates were then compared to the Total, Distress, log-transformed PQ-BC scores for baseline and year 1, respectively (and differences between significant correlations were examined using Meng's z-test procedures (Meng et al., 1992)). Then, based on item information and discrimination parameters from baseline ABCD data, we created a screening form with similar psychometric properties to the full PQ-BC. HLMs analyzed the associations between screening form Total, Distress, and log-transformed Distress scores for both baseline and year 1 with the

aforementioned predictors (i.e., items 1-5) to identify whether the screening form performed in the expected manner at both baseline and year 1.

Results

IRT Parameters and information

Item discrimination quantifies how well an item can discriminate individuals at varying levels of the latent trait, with higher discrimination parameter values indicating the item excels at discriminating low from high trait individuals. IRT item discrimination (*a)* ranged from 1.365 to 2.210 (*M*=1.819). The discrimination parameters for the PQ-BC are quite high, on average, which indicates that every item is highly discriminatory and that the measured construct (PLEs) is conceptually narrow (Reise, 2009). The means of the two difficulty parameter estimates (*b)* were 1.64 (range:1.14-2.3) and 2.75 (1.93-3.85), respectively. These estimates reflect the location of the inflection points of the item response functions. Both the *b*1 and *b2* threshold estimates are large, positive, and range restricted, which indicates that only people with high trait levels of PLEs are typically endorsing PQ-BC items (i.e., endorsing the item statement with significant distress). Finally, we calculated information provided by each test item, which utilizes both *a* and *b* parameters to estimate how much information a test item contributes to $\hat{\theta}$ (see Figure 1 for the overall test information function for the PQ-BC scores and screening form; see Supplemental Figures 1 and 2 for the individual IRFs; see Table 1 for item parameter estimates and information).

Associations of $\hat{\theta}$ and CTT Indices to Predictors of Psychosis Spectrum Symptoms

The Total and Distress factor scores significantly correlated with $\hat{\theta}$ scores (Total: .915; Distress: .884, *p*s<.001). Next, we examined the longitudinal stability of $\hat{\theta}$ scores from baseline to year 1. The longitudinal stability of $\hat{\theta}$ scores was β=0.43 (95% CI=0.41,0.45), *p*<.001, which was comparable to other PQ-BC scores [Total scores: β=0.43 (95% CI=0.41,0.46), *p*<.001, Distress: β=0.40 (95% CI=0.38,0.42), *p*<.001, log-transformed Distress: β=0.43 (95% CI=0.41,0.46), *p*<.001]. Next, to better understand the latent construct of PLEs, we also examined the associations between $\hat{\theta}$ scores and psychosis spectrum symptom predictors (i.e., parent-rated PLEs, family history of psychosis, internalizing symptoms, cognitive functioning, developmental milestone delays, RSFC metrics). Baseline $\hat{\theta}$ scores were significantly associated with almost all predictors (see Table 2, with the exception of several of the RSFC metrics, although notably none of the PQ-BC scores were significantly associated with CON-cerebellar or CPAR-cerebellar connectivity, and only baseline log-transformed distress scores were significantly associated with impaired CPAR connectivity; note, in terms of discriminant validity, as expected, the CBCL somatization index was not significantly associated with PQ-BC scores, β=-.016, *p*=.21). As can be seen in Table 2, year 1 $\hat{\theta}$ scores were also associated with almost all predictors (with the exception of the RSFC metrics, although notably most year 1 PQ-BC scores were not significantly associated with RSFC metrics, with the exception of Total, Distress, and $\hat{\theta}$ scores being significantly associated with impaired CON connectivity and Total scores being significantly associated with impaired CPAR connectivity). This provides evidence that assessments of the latent construct of PLEs showed the expected associations with predictors of psychosis spectrum symptoms.

Screening Form

Importantly, the IRT analyses enabled the examination of whether the total number of PQ-BC items could be shortened based on IRT-derived parameters. For the purpose of scale reduction, we prioritized both item level information and discrimination parameters from baseline ABCD data. Based on this information, we created a screening form comprised of items that ranked in the top 10 for both discrimination and information (see Table 1). This screening form consists of the following questions: 5,10,14,16,19,20,21 (see Table 1 for question content). This screening form showed fair internal consistency (screening form Total: α=.72, year 1: α=.71; screening form Distress: α=.72, year 1: α=.74, in comparison, for the overall PQ-BC Total: α=.86, year 1: α=.87; screening form Distress: α=.87, year 1: α=.87), which did not increase when any item was deleted. The correlation coefficients between each item and the screening form Total score were 0.33–0.48 (year 1: 0.32-0.52), and for the screening form Distress Score, 0.35–0.49 (year 1: 0.37–0.52). The longitudinal stability of the screening form for the Total Score was β=0.35 (95% CI=0.32,0.37), p<.001, Distress Score, β=0.32 (95% CI=0.29,0.34), p<.001, and log-transformed Distress score was β=0.34 (95% CI=0.32,0.36), p<.001.

We also examined associations between the screening form to predictors of psychosis spectrum symptoms (i.e., parent-rated PLEs, family history of psychosis, internalizing symptoms, cognitive functioning, developmental milestone delays, RSFC metrics). As can be seen in Table 3, the screening form was significantly correlated with most predictors for both baseline and Year 1 [with the exception of some of the RSFC metrics for both baseline and year 1 (specifically, none of the year 1 PQ-BC screening

form scores were associated with DMN within-network connectivity, furthermore, baseline and year 1 PQ-BC screening form scores were not significantly associated with either CON-cerebellar or CPAR-cerebellar connectivity). These findings are generally consistent with what was found for the full scale, as can be seen in Table 2.

Discussion

Our goals were to (1) examine item properties of the PQ-BC by utilizing IRT; (2) examine associations of the latent construct of PLEs ($\hat{\theta}$) with previously identified predictors of psychosis spectrum symptoms; (3) and use the information and discrimination parameters to assess the PQ-BC and derive a screening form. For the first aim, similar to many other clinical measures that have been assessed using IRT, analyses of the PQ-BC suggest each item provides a large amount of information and is highly discriminatory at high trait levels, but provides less reliable estimates at lower trait levels. Given that the measure was designed to measure a clinical phenomenon, this is not surprising; however, it does suggest the PQ-BC is less sensitive at assessing low trait presentation. As for the second aim, as predicted, the latent construct of PLEs showed significant associations in the expected direction with previously established predictors of psychosis spectrum symptoms. Finally, the 7-item screening form, derived from information and discrimination parameters, performed adequately within the ABCD dataset and may be beneficial for more general use when administration of the full measure is prohibitive.

Endorsement of any of the PQ-BC items was relatively rare; with that stated, IRT results suggest that endorsement of particular items may be more important for assessing syndrome severity than others. Importantly, our results replicate other recent

work (Phalen et al., 2018), which found item 20 to be particularly useful when assessing for PLEs. Examining the discrimination parameters provides convergent evidence, as the values reported for each item are quite high (Reise & Waller, 2009), suggesting that PQ-BC items were highly discriminatory at each *b* inflection point, though discrimination suffers at low levels of trait assessment. Other IRT examinations of clinical syndromes have reported similar parameter estimates (Aggen et al., 2005; Hays et al., 2007; Chan et al., 2004), suggesting this is not an issue specific to the PQ-BC but to clinical instrumentation broadly. Whereas the PQ-BC provides information at a restricted range of the trait (e.g. high levels of information at high theta), past scholarship using IRT in PLE assessment identified several items which provided less overall information but perhaps greater coverage across θ (Laurens et al., 2012). Although this restricted range is desirable for our screening form measure to serve as a quick screening tool, inclusion of other items may provide useful for more comprehensive assessment of PLEs at lower trait levels. Because the items in the PQ-BC are range-restricted to the highest levels of theta, the PQ-BC's association with other variables will be less reliable at lower theta levels. Overall, the IRT results are generally consistent with what would be expected for subclinical psychotic-like experiences, and support evidence of the feasibility and utility of IRT analyses for psychotic-like experiences measures (Fonseca-Pedrero, Paino, Ortuno-Sierra, Lemos-Giraldez, & Muniz, 2014; Levey et al., 2018; van Bebber et al., 2017).

Overall, $\hat{\theta}$ scores demonstrated expected associations with several key factors linked with PLEs, including cognitive impairments, developmental milestone delays, parent-rated PLEs, family history of psychosis, anxiety/depression, and some RSFC

impairments. These important findings indicate that the latent trait of PLEs showed analogous associations to those found in previous studies of psychosis spectrum symptoms and is therefore informative about the correlates of experiences across the psychosis spectrum (Cannon et al., 2002; Kelleher et al., 2013; Mollon & Reichenberg, 2017). Furthermore, examining the latent trait of PLEs is informative regarding the construct validity of the PQ-BC, as well as is more broadly informative of the nomological network of psychosis spectrum symptoms, including PLEs. IRT methodology includes stronger assumptions about assigning error variance at the item level than CTT methods, which proponents have argued makes $\hat{\theta}$ more accurately reflective of true trait scores (Reise & Henson, 2003). In our analyses, the $\hat{\theta}$ scores showed stronger associations with certain predictors of PLEs, such as fluid and crystallized intelligence impairments, which may suggest that some of these outcomes are more associated with psychotic-like experiences trait presentation than previously reported. Our results also suggest that prior examinations with CTT derived scores may have slightly inflated the associations of PLEs with several predictors (i.e., internalizing symptoms, working memory, speech developmental delays); with that stated, our results show that these associations are still significant, suggesting these continue to be important associations to consider during assessment.

As mentioned above, we used IRT derived parameters to aid in the creation of a screening form. This screening form consists of seven items along multiple domains [e.g. ideas of reference, auditory perceptual abnormalities, unusual beliefs (i.e., delusional confusion), other perceptual abnormalities (e.g., olfactory perceptual experiences), visual perceptual abnormalities, and disorganized speech]. For both

baseline and year 1, the screening form showed expected associations with the most risk factors (although note the screening form scores failed to show significant associations with several of the RSFC variables). The seven item screening form had fair internal consistency, which was expected (Nunnally, Bernstein, & Berge, 1967; Ziegler, Kemper, & Kruyen, 2014). Thus, one likely does not need to include all 21 items of the PQ-BC to assess elevations in PLEs since each item provides information at similar theta ranges. Our proposed screening form is adequately reliable, and provides wide, clinically relevant symptom coverage (more so than (Phalen et al., 2018)), and is comparable to the PQ-BC in terms of associations with known correlates of PLEs. The current study constructed a screening form starting from the full scale rather than directly replicating previous efforts (Phalen et al., 2018) due in part to several factors, including a) the previous study used the PQ-B rather than the PQ-BC, b) along these lines, the previous study used an adult population, and c) the current study decided to err on the side of retaining items which provided the most information and discrimination while incorporating a spectrum of PLEs experiences, including both delusional ideation and perceptual abnormalities.

This is the first time that the longitudinal stability and psychometric properties of the PQ-BC have been examined over time, information critical to understanding the nature of school-age PLEs. Importantly, all PQ-BC metrics (Total, Distress, log-transformed, and $\hat{\theta}$) showed evidence of longitudinal stability, including moderate positive correlations between baseline and year 1. This provides preliminary support for this measure assessing a stable trait (Fagerberg, Soderman, Petter Gustavsson, Agartz, & Jonsson, 2018). Furthermore, the psychometric properties of the PQ-BC

replicated in year 1 (e.g., good internal consistency), providing additional evidence that the PQ-BC is validly assessing school-age PLEs. However, there were some differences between baseline and year 1 in terms of associations between PQ-BC scores and risk factors. Interestingly, some of the associations between PQ-BC and RSFC variables were not significant at year 1. This may indicate that during this important developmental period, maturational changes in connectivity are occurring and therefore it will be important to examine associations between RSFC variables assessed later in development and year 1 PQ-BC scores. However, it could also indicate that some of the RSFC findings (e.g., CPAR connectivity) do not represent stable risk correlates of school-age PLEs.

Several questions still remain regarding the developmental nature of psychotic-like experiences, and how well the PQ-BC assesses these symptoms longitudinally. The screening form was developed using information and discrimination values from the baseline sample; future inquiries will need to examine developmental effects using the full sample at year 1. Examining measurement invariance at varying timepoints will be useful in revising which items are most appropriate for a screening-form measure. Additionally, while the ABCD sample is population-based and demographically diverse sample, these results are conducted on a non-help-seeking population. While IRT analyses are theoretically not sample dependent (Reise & Henson, 2003), future research may seek to develop PQ-BC clinical cut-offs for specific clinical settings. In addition, a number of screening forms could have been developed based on the information and discrimination properties of the items (e.g., top five items based on information or discrimination). The current study decided to use items that were in the

top ten most informative *and* most discriminative items in order to take into account the importance of retaining items which both a) were most informative of the PLE trait and b) most discriminating of those actually experiencing PLEs versus those who are not experiencing PLEs. Furthermore, the screening form was developed for early identification efforts, and therefore aims to capture greater severity of PLEs/higher scorers (as opposed to the lower end of PLE coverage, which is less likely to be clinically relevant; thus, while the PLE trait in the general population is often expressed as mild elevations in PQ-BC scores, it is expected that a screening measure should only identify those individuals at higher levels of the trait and therefore most likely to develop sustained PLEs of clinical relevance (Dominguez et al., 2011; Kalman et al., 2019). A future goal of developing a screening form is for use in large-scale screening efforts to identify children at-risk for increased psychotic-like experiences through schools, pediatrician offices, or other community settings. Overall, the current analyses provide initial evidence that a seven-item screening form of the PQ-BC validly assesses school-age PLEs.

Acknowledgments

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/Consortium_Members.pdf. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

The ABCD data repository grows and changes over time. The ABCD data used in this report came from DOI 10.15154/1503209.

Conflict of Interest Disclosures: Authors report no conflicts of interest.

Additional Contributions: We thank the families participating in the Adolescent Brain and Cognitive Development study.

References

Achenbach, T. M. (2009). *The Achenbach System of Emprically Based Assessment (ASEBA): Development, Findings, Theory and Applications*. Burlington, VT: University of Vermong Research Center for Children, Youth, and Families.

Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., . . . Sher, K. J. (2018). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Dev Cogn Neurosci, 32*, 55-66. doi:10.1016/j.dcn.2017.10.010

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*: Guilford Publications.

Cannon, M., Caspi, A., Moffitt, T. E., Harrington, H., Taylor, A., Murray, R. M., & Poulton, R. (2002). Evidence for early-childhood, pan-developmental impairment specific to schizophreniform disorder: results from a longitudinal birth cohort. *Arch Gen Psychiatry, 59*(5), 449-456. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11982449

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., . . . Dale, A. M. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci*. doi:10.1016/j.dcn.2018.03.001

Chernyshenko, O. S., Stark, S., Drasgow, F., & Roberts, B. W. (2007). Constructing personality scales under the assumptions of an ideal point response process: Toward increasing the flexibility of personality measures. *Psychological Assessment, 19*(1), 88.

Cho, S., Drasgow, F., & Cao, M. (2015). An investigation of emotional intelligence measures using item response theory. *Psychol Assess, 27*(4), 1241-1252. doi:10.1037/pas0000132

Cicero, D. C., Krieg, A., & Martin, E. A. (2017). Measurement Invariance of the Prodromal Questionnaire-Brief Among White, Asian, Hispanic, and Multiracial Populations. *Assessment*, 1073191116687391. doi:10.1177/1073191116687391

Dominguez, M. D., Wichers, M., Lieb, R., Wittchen, H. U., & van Os, J. (2011). Evidence that onset of clinical psychosis is an outcome of progressively more persistent subclinical psychotic experiences: an 8-year cohort study. *Schizophr Bull, 37*(1), 84-93. doi:10.1093/schbul/sbp022

Drasgow, F., & Hulin, C. L. (1990). Item response theory.

Ducharme, S., Albaugh, M. D., Nguyen, T. V., Hudziak, J. J., Mateos-Perez, J. M., Labbe, A., . . . Brain Development Cooperative, G. (2015). Trajectories of cortical surface area and cortical volume maturation in normal brain development. *Data Brief, 5*, 929-938. doi:10.1016/j.dib.2015.10.044

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Qual Life Res, 16 Suppl 1*, 5-18. doi:10.1007/s11136-007-9198-0

Fagerberg, T., Soderman, E., Petter Gustavsson, J., Agartz, I., & Jonsson, E. G. (2018). Stability of personality traits over a five-year period in Swedish patients with schizophrenia spectrum disorder and non-psychotic individuals: a study using the Swedish universities scales of personality. *BMC Psychiatry, 18*(1), 54. doi:10.1186/s12888-018-1617-y

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., . . . Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron, 33*(3), 341-355.

Fonseca-Pedrero, E., Gooding, D. C., Ortuno-Sierra, J., & Paino, M. (2016). Assessing self-reported clinical high risk symptoms in community-derived adolescents: A psychometric evaluation of the Prodromal Questionnaire-Brief. *Compr Psychiatry, 66*, 201-208. doi:10.1016/j.comppsych.2016.01.013

Fonseca-Pedrero, E., Inchausti, F., Perez-Albeniz, A., & Ortuno-Sierra, J. (2018). Validation of the Prodromal Questionnaire-Brief in a representative sample of adolescents: Internal structure, norms, reliability, and links with psychopathology. *Int J Methods Psychiatr Res, 27*(4), e1740. doi:10.1002/mpr.1740

Fonseca-Pedrero, E., Paino, M., Ortuno-Sierra, J., Lemos-Giraldez, S., & Muniz, J. (2014). The assessment of positive dimension of the psychosis phenotype in college students. *Compr Psychiatry, 55*(3), 699-707. doi:10.1016/j.comppsych.2013.09.013

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., & Petersen, S. E. (2016). Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb Cortex, 26*(1), 288-303. doi:10.1093/cercor/bhu239

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2): Sage.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Med Care, 38*(9 Suppl), Ii28-42.

Hirshfeld-Becker, D. R. (2006). *Being Brave: A Program for Coping with Anxiety for Young Children and Their Parents*. Massachusetts General Hospital.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical journal, 50*(3), 346-363.

Howell, D. (2007). Statistical methods for psychology Thomson Wadsworth. *Belmont, CA*, 1-739.

Karcher, N. R., Barch, D. M., Avenevoli, S., Savill, M., Huber, R. S., Simon, T. J., . . . Loewy, R. L. (2018). Assessment of the Prodromal Questionnaire-Brief Child Version for Measurement of Self-reported Psychoticlike Experiences in Childhood. *JAMA Psychiatry*. doi:10.1001/jamapsychiatry.2018.1334

Karcher, N. R., O'Brien, K. J., Kandala, S., & Barch, D. M. (2019). Resting-State Functional Connectivity and Psychotic-like Experiences in Childhood: Results From the Adolescent Brain Cognitive Development Study. *Biol Psychiatry*. doi:10.1016/j.biopsych.2019.01.013

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry, 36*(7), 980-988. doi:10.1097/00004583-199707000-00021

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. *Psychol Med, 42*(9), 1857-1863. doi:10.1017/S0033291711002960

Kelleher, I., Murtagh, A., Clarke, M. C., Murphy, J., Rawdon, C., & Cannon, M. (2013). Neurocognitive performance of a community-based sample of young people at putative ultra high risk for psychosis: support for the processing speed hypothesis. *Cogn Neuropsychiatry, 18*(1-2), 9-25. doi:10.1080/13546805.2012.682363

Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa, S., . . . Zaslavsky, A. M. (2009). Design and field procedures in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A). *Int J Methods Psychiatr Res, 18*(2), 69-83. doi:10.1002/mpr.279

Kobak, K. A., Kratochvil, C. J., Stanger, C., & Kaufman, J. (2013). *Computerized screening of comorbidity in adolescents with substance or psychiatric disorders.* Paper presented at the Anxiety Disorders and Depression, La Jolaa, CA.

Laurens, K. R., Hodgins, S., Maughan, B., Murray, R. M., Rutter, M. L., & Taylor, E. A. (2007). Community screening for psychotic-like experiences and other putative antecedents of schizophrenia in children aged 9-12 years. *Schizophr Res, 90*(1-3), 130-146. doi:10.1016/j.schres.2006.11.006

Lawley, D. N. (1943). XXIII.—On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 61*(3), 273-287.

Levey, E. J., Zhong, Q. Y., Rondon, M. B., Sanchez, S., Li, J., Williams, M. A., & Gelaye, B. (2018). The psychometric properties of the 16-item version of the Prodromal Questionnaire (PQ-16) as a screening instrument for perinatal psychosis. *Arch Womens Ment Health, 21*(5), 563-572. doi:10.1007/s00737-018-0833-2

Linscott, R. J., & van Os, J. (2013). An updated and conservative systematic review and meta-analysis of epidemiological evidence on psychotic experiences in children and adults: on the pathway from proneness to persistence to dimensional expression across mental disorders. *Psychol Med, 43*(6), 1133-1149. doi:10.1017/S0033291712001626

Loewy, R. L., Bearden, C. E., Johnson, J. K., Raine, A., & Cannon, T. D. (2005). The prodromal questionnaire (PQ): preliminary validation of a self-report screening measure for prodromal and psychotic syndromes. *Schizophr Res, 79*(1), 117-125. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16276559

Loewy, R. L., Pearson, R., Vinogradov, S., Bearden, C. E., & Cannon, T. D. (2011). Psychosis risk screening with the Prodromal Questionnaire--brief version (PQ-B). *Schizophr Res, 129*(1), 42-46. doi:10.1016/j.schres.2011.03.029

Lord, F. M. (1953). The relation of test score to the trait underlying the test. *Educational and Psychological Measurement, 13*(4), 517-549.

Maples-Keller, J. L., Williamson, R. L., Sleep, C. E., Carter, N. T., Campbell, W. K., & Miller, J. D. (2019). Using item response theory to develop a 60-item representation of the NEO PI–R using the International Personality Item Pool: development of the IPIP–NEO–60. *J Pers Assess, 101*(1), 4-15.

Meng, X.-L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychol Bull, 111*(1), 172.

Mollon, J., & Reichenberg, A. (2017). Cognitive development prior to onset of psychosis. *Psychol Med*, 1-12. doi:10.1017/s0033291717001970

Muthén, L. (2010). Mplus User's Guide, (Muthén & Muthén, Los Angeles). *Mplus User's Guide,(Muthén & Muthén, Los Angeles)*.

Nunnally, J. C., Bernstein, I. H., & Berge, J. M. t. (1967). *Psychometric theory* (Vol. 226): McGraw-hill New York.

Phalen, P. L., Rouhakhtar, P. R., Millman, Z. B., Thompson, E., DeVylder, J., Mittal, V., . . . Schiffman, J. (2018). Validity of a two-item screen for early psychosis. *Psychiatry Res, 270*, 861-868. doi:10.1016/j.psychres.2018.11.002

Rasch, G. (1960). Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.

Reise, S. P., & Henson, J. M. (2003). A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *J Pers Assess, 81*(2), 93-103. doi:10.1207/s15327752jpa8102_01

Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annu Rev Clin Psychol, 5*, 27-48. doi:10.1146/annurev.clinpsy.032408.153553

Rice, J. P., Reich, T., Bucholz, K. K., Neuman, R. J., Fishman, R., Rochberg, N., . . . Begleiter, H. (1995). Comparison of direct interview and family history diagnoses of alcohol dependence. *Alcohol Clin Exp Res, 19*(4), 1018-1023.

Rimvall, M. K., Gundersen, S., Clemmensen, L., Munkholm, A., Larsen, J. T., Skovgaard, A. M., . . . Jeppesen, P. (2019). Evidence that self-reported psychotic experiences in children are clinically relevant. *Schizophr Res, 204*, 415-416. doi:10.1016/j.schres.2018.08.003

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software, 17*(5), 1-25.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. Retrieved from <Go to ISI>://WOS:000305117100001

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika monograph supplement*.

Satterthwaite, T. D., Vandekar, S. N., Wolf, D. H., Bassett, D. S., Ruparel, K., Shehzad, Z., . . . Gur, R. E. (2015). Connectome-wide network analysis of youth with Psychosis-Spectrum symptoms. *Mol Psychiatry, 20*(12), 1508-1515. doi:10.1038/mp.2015.66

Savill, M., D'Ambrosio, J., Cannon, T. D., & Loewy, R. L. (2018). Psychosis risk screening in different populations using the Prodromal Questionnaire: A systematic review. *Early Interv Psychiatry, 12*(1), 3-14. doi:10.1111/eip.12446

Sheffield, J. M., Kandala, S., Burgess, G. C., Harms, M. P., & Barch, D. M. (2016). Cingulo-opercular network efficiency mediates the association between psychotic-like experiences and cognitive ability in the general population. *Biol Psychiatry Cogn Neurosci Neuroimaging, 1*(6), 498-506. doi:10.1016/j.bpsc.2016.03.009

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology, 91*(6), 1292.

van Bebber, J., Wigman, J. T. W., Meijer, R. R., Ising, H. K., van den Berg, D., Rietdijk, J., . . . Wunderink, L. (2017). The Prodromal Questionnaire: a case for IRT-based adaptive testing of psychotic experiences? *Int J Methods Psychiatr Res, 26*(2). doi:10.1002/mpr.1518

van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. (2009). A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychol Med, 39*(2), 179-195. doi:10.1017/S0033291708003814

van Os, J., & Reininghaus, U. (2016). Psychosis as a transdiagnostic and extended phenotype in the general population. *World Psychiatry, 15*(2), 118-124. doi:10.1002/wps.20310

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., . . . Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology, 80*(11 Suppl 3), S54-64. doi:10.1212/WNL.0b013e3182872ded

Wusten, C., Schlier, B., Jaya, E. S., Fonseca-Pedrero, E., Peters, E., Verdoux, H., . . . Lincoln, T. M. (2018). Psychotic Experiences and Related Distress: A Cross-national Comparison and Network Analysis Based on 7141 Participants From 13 Countries. *Schizophr Bull, 44*(6), 1185-1194. doi:10.1093/schbul/sby087

Ziegler, M., Kemper, C. J., & Kruyen, P. (2014). Short scales–Five misunderstandings and ways to overcome them. In: Hogrefe Publishing.

Table 1. Item Response Theory (IRT) Parameters for the Baseline ABCD Sample.

Item #	PQ-BC Question	*a*	*b*1	*b*2	*Information provided*	Discrimination Ranking	Information Ranking
1	Did places that you know well, such as your bedroom, or other rooms in your home, your classroom or school yard, suddenly seem weird, strange or confusing to you; like not the real world?	1.828	2.228	2.983	2.62	8	13
2	Did you hear strange sounds that you never noticed before like banging, clicking, hissing, clapping, or ringing in your ears?	1.660	1.235	2.376	2.6	18	14
3	Did things you looked at seem different than they usually do; like did they seem shinier or darker, larger or smaller or changed in some other way?	1.663	1.770	3.276	2.82	16	11
4	Did you feel like you had special, unusual powers like you could make things happen by magic, or that you could magically know what was inside another person's mind, or magically know what was going to happen in the future when other people could not?	1.742	1.731	3.583	3.14	14	8
5	Did you feel that someone else, who is not you, has taken control over the private, personal, thoughts or ideas inside your head?	**2.103**	**2.139**	**2.668**	**2.85**	**4**	**10**
6	Did you suddenly find it hard to figure out how to say something quickly and easily so that other people would understand what you meant?	1.538	1.351	3.119	2.66	20	12
7	Did you ever feel very certain that you have very special abilities or magical talents that other people do not have?	1.743	1.704	3.849	3.25	13	5
8	Did you suddenly feel that you could not trust other people because they seemed to be watching you or talking about you in an unfriendly way?	1.643	1.136	2.051	2.41	19	19
9	Did your skin or just beneath your skin suddenly start feeling strange, like bugs crawling?	1.661	1.540	2.376	2.38	17	20
10	Did you lose concentration because you noticed sounds in the distance that you usually don't hear?	**2.079**	**1.395**	**2.283**	**3.23**	**5**	**6**

11	Although you could not see anything or anyone, did you suddenly start to feel that an invisible energy, creature, or some person was around you?	1.738	1.224	1.931	2.42	15	18
12	Did you start to worry at times that your mind was trying to trick you or was not working right?	1.786	1.781	2.449	2.46	11	17
13	Did you feel that the world is not real, you are not real, or that you are dead?	1.795	2.119	2.889	2.57	10	15
14	Did you feel confused because something you experienced didn't seem real or it seemed imaginary to you?	**1.945**	**1.589**	**2.764**	**3.21**	**6**	**7**
15	Did you honestly believe in things that other people would say are unusual or weird?	1.365	1.953	3.293	2.11	21	21
16	Did you feel that parts of your body had suddenly changed or worked differently than before; like your legs had suddenly turned to something else or your nose could suddenly smell things you'd never actually smelled before?	**2.121**	**1.914**	**3.047**	**3.55**	**3**	**1**
17	Did you feel that sometimes your thoughts were so strong you could almost hear them, as if another person, NOT you, spoke them?	1.757	1.652	3.032	2.95	12	9
18	Did you feel that other people might want something bad to happen to you or that you could not trust other people?	1.798	1.470	2.186	2.53	9	16
19	Did you suddenly start to see unusual things that you never saw before like flashes, flames, blinding light, or shapes floating in front of you?	**2.140**	**1.498**	**2.327**	**3.29**	**2**	**3**
20	Did you suddenly start to be able to see things that other people could not see or they did not seem to see?	**2.210**	**1.637**	**2.529**	**3.5**	**1**	**2**
21	Did you suddenly start to notice that people sometimes had a hard time understanding what you were saying, even though they used to understand you well?	**1.887**	**1.384**	**2.780**	**3.24**	**7**	**4**

Abbreviations. *a* = discrimination parameter (i.e., how well an item can discriminate individuals at varying levels of the psychotic-like experiences). *b*1 and *b*2= two difficulty parameter estimates (i.e. how high on trait psychotic-like experiences one must be to respond affirmatively).

Table 2. Associations of PQ-BC Models for Baseline and Year 1 ABCD Samples.

	Total			Distress			Log-Transformed Distress			$\hat{\theta}$		
	β	*t*	*p*	*β*	*t*	*p*	*β*	*t*	*p*	*β*	*t*	*p*
Baseline (n=11,129)												
Parent-rated PLEs	0.089	9.55	<.001	0.095	10.11	<.001	0.082	8.86	<.001	0.086	9.28	<.001
Family History of Psychosis	0.035	3.71	<.001	0.036	3.86	<.001	0.035	3.70	<.001	0.032	3.44	.001
Internalizing Symptoms	0.323	37.79	<.001[a]	0.369	43.51	<.001[b]	0.283	32.65	<.001	0.286	33.28	<.001
Cognitive Functioning												
Fluid	-0.058	-5.42	<.001	-0.066	-6.20	<.001	-0.075	-7.14	<.001[c]	-0.069	-6.57	<.001[d]
Crystalized	-0.077	-6.942	<.001	-0.085	-7.58	<.001	-0.098	-8.90	<.001[c]	-0.091	-8.32	<.001[d]
Working Memory	-0.055	-5.10	<.001	-0.064	-5.88	<.001	-0.064	-6.04	<.001[c]	-0.060	-5.65	<.001
Developmental Milestone Delays												
Motor	0.036	3.89	<.001	0.035	3.71	<.001	0.040	4.35	<.001	0.039	4.19	<.001
Speech	0.044	4.63	<.001	0.046	4.80	<.001	0.054	5.74	<.001[c]	0.048	5.17	<.001
Clumsiness	0.03	3.43	.001	0.041	4.44	<.001	0.037	4.04	<.001	0.03	3.70	<.001

	1									4		
RSFC[e]												
CON	-0.025	-2.32	.02	-0.030	-2.75	.006	-0.026	-2.43	.02	-0.027	-2.60	.009
CPAR	-0.005	-0.49	.63	-0.009	-0.87	.39	-0.025	-2.39	.02[c]	-0.017	-1.669	.095[d]
DMN	-0.028	-2.68	.007	-0.030	-2.77	.006	-0.029	-2.71	.007	-0.029	-2.81	.005
CON-cerebellar	0.015	1.42	.16	0.016	1.56	.12	0.014	1.33	.18	0.015	1.45	.15
CPAR-cerebellar	-0.015	-1.52	.13	-0.012	-1.13	.26	-0.015	-1.51	.13	-0.016	-1.61	.11
Year 1 (n=5,532)												
Parent-rated PLEs	0.037	3.854	<.001	0.033	3.367	<.001	0.085	6.252	<.001	0.088	6.49	<.001[d, f]
Family History of Psychosis	0.063	4.673	<.001	0.069	5.082	<.001	0.059	4.432	<.001	0.060	4.47	<.001
Internalizing Symptoms	0.241	19.061	<.001[a]	0.234	18.257	<.001[b]	0.218	17.282	<.001	0.215	17.10	<.001
Cognitive Functioning												
Fluid	-0.052	-3.463	.001	-0.062	-4.108	<.001	-0.055	-3.685	<.001	-0.052	-3.51	<.001
Crystalized	-0.060	-3.831	<.001	-0.060	-3.816	<.001	-0.092	-5.953	<.001	-0.089	-5.74	<.001[d, f]

Working Memory	-0.059	-3.883	<.001	-0.052	-3.385	.001	-0.063	-4.246	<.001	-0.060	-4.03	<.001
Developmental Milestone Delays												
Motor	0.043	3.240	.001	0.043	3.148	.002	0.053	4.015	<.001	0.048	3.65	<.001
Speech	0.045	3.355	.001	0.042	3.072	.002	0.047	3.517	<.001	0.048	3.59	<.001
Clumsiness	0.051	3.895	<.001	0.051	3.875	<.001	0.051	3.925	<.001	0.049	3.80	<.001
RSFC[e]												
CON	-0.025	-2.13	.03	-0.026	-2.14	.03	-0.026	-1.63	.10	-0.037	-2.30	.02
CPAR	-0.028	-2.44	.01	-0.019	-1.62	.11	-0.013	-0.85	.40	-0.012	-0.80	.42
DMN	-0.012	-1.05	.29	0.006	0.48	.63	-0.013	-0.84	.40	-0.014	-0.92	.36
CON-cerebellar	0.008	0.68	.50	0.004	0.34	.73	0.000	0.02	.98	0.000	0.01	.99
CPAR-cerebellar	-0.012	-1.10	.27	-0.002	-0.22	.83	-0.003	-0.21	.84	0.001	0.07	.94

Abbreviations: PQ-BC: Prodromal Questionnaire Brief-Child Version; $\hat{\theta}$= scores that estimate how responses across the whole scale should be translated to rescore individual's placement along the trait continuum*; β*=standardized regression

coefficient; *t*=t-test test statistic; *p*=p-value; RSFC=resting-state functional connectivity; CON=cingulo-opercular within-network connectivity; CPAR=cingulo-parietal within-network connectivity; DMN=default mode within-network connectivity.
[a]Associations with Total scores were significantly stronger than $\hat{\theta}$ (p<.05).
[b]Associations with Distress scores were significantly stronger than $\hat{\theta}$ (p<.05).
[c]Associations with log-transformed Distress scores were significantly stronger than $\hat{\theta}$ scores (p<.05).
[d]Associations with $\hat{\theta}$ scores were significantly stronger than Total scores (p<.05).
[e]For RSFC, baseline n=8,859; year 1 n=4,283.
[f]Associations with $\hat{\theta}$ scores were significantly stronger than Distress scores (p<.05).

Table 3. Associations of Screening Form Total and Distress Scores for Baseline and Year 1.

	Screening Form Total			Screening Form Distress			Screening Form Log-Transformed Distress		
	β	*t*	*p*	*β*	*t*	*p*	*β*	*t*	*p*
Baseline									
Parent-rated PLEs	0.080	8.52	<.001	0.085	8.92	<.001	0.076	8.02	<.001
Family History of Psychosis	0.036	3.81	<.001	0.035	3.63	<.001	0.039	4.15	<.001
Internalizing Symptoms	0.298	33.95	<.001	0.336	38.53	<.001	0.297	33.91	<.001
Cognitive Functioning									
Fluid	-0.048	-4.44	<.001	-0.054	-4.95	<.001	-0.057	-5.26	<.001
Crystalized	-0.084	-7.55	<.001	-0.081	-7.13	<.001	-0.090	-8.03	<.001
Working Memory	-0.045	-4.17	<.001	-0.053	-4.81	<.001	-0.051	-4.74	<.001
Developmental Milestone Delays									
Motor	0.037	3.88	<.001	0.030	3.13	.002	0.036	3.82	<.001
Speech	0.039	4.16	<.001	0.040	4.22	<.001	0.048	5.07	<.001
Clumsiness	0.034	3.71	<.001	0.038	4.11	<.001	0.037	4.00	<.001
RSFC									
CON	-0.032	-2.92	.004	-0.035	-3.19	.001	-0.036	-3.36	<.001
CPAR	-0.001	-0.07	.95	-0.006	-0.59	.56	-0.051	-1.65	.10
DMN	-0.031	-2.87	.004	-0.036	-3.35	<.001	-0.040	-3.74	<.001
CON-cerebellar	0.012	1.16	.25	0.017	1.59	.11	0.012	1.11	.27
CPAR-cerebellar	-0.019	-1.88	.06	-0.014	-1.34	.18	-0.014	-1.39	.16
Year 1									
Parent-rated PLEs	0.073	5.26	<.001	0.070	5.03	<.001	0.071	5.14	<.001
Family History of Psychosis	0.052	3.91	<.001	0.052	3.70	<.001	0.047	3.43	.001

Internalizing Symptoms	0.223	17.33	<.001	0.200	15.57	<.001	0.202	15.96	<.001
Cognitive Functioning									
Fluid	-0.046	-3.08	.002	-0.052	-3.33	.001	-0.044	-2.88	.004
Crystalized	-0.066	-4.20	<.001	-0.064	-3.95	<.001	-0.087	-5.39	<.001
Working Memory	-0.050	-3.26	.001	-0.036	-2.19	.03	-0.054	-3.34	.001
Developmental Milestone Delays									
Motor	0.035	2.60	.009	0.034	2.47	.01	0.045	3.34	.001
Speech	0.046	3.38	.001	0.037	2.69	.007	0.039	2.87	.004
Clumsiness	0.041	3.08	.002	0.049	3.62	<.001	0.052	3.87	<.001
RSFC									
CON	-0.053	-3.22	.001	-0.049	-2.96	.003	-0.054	-3.29	.001
CPAR	-0.022	-1.46	.14	-0.017	-1.12	.26	-0.028	-1.83	.07
DMN	-0.022	-1.36	.17	-0.017	-1.04	.30	-0.023	-1.47	.14
CON-cerebellar	-0.009	-0.59	.56	-0.007	-0.42	.67	-0.007	-0.42	.67
CPAR-cerebellar	0.012	0.77	.44	0.008	0.49	.62	0.005	0.31	.76

Abbreviations: *β*=standardized regression coefficient; *t*=t-test test statistic; *p*=p-value. RSFC=resting-state functional connectivity; CON=cingulo-opercular within-network connectivity; CPAR=cingulo-parietal within-network connectivity; DMN=default mode within-network connectivity.

Figure 1. Test Information Functions for both PQ-BC scores and PQ-BC Screening Form scores.