Criterion validity of the Psychotic-Like Experiences Questionnaire for Children (PLEQ-C)

Tiffany P. Gutteridge^a, Cathryne P. Lang^{b,c}, Alison M. Turner^b, Brian W. Jacobs^{d,e}, Kristin R. Laurens^{a,f,*}

^a School of Psychology and Counselling, and Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia

^b School of Psychology, Australian Catholic University, Brisbane, QLD, Australia

^c Sunshine Coast Mind and Neuroscience - Thompson Institute, University of the Sunshine Coast, Birtinya, QLD, Australia

^d Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK

^e Child and Adolescent Mental Health Services, South London and Maudsley National Health Service Foundation Trust, London, UK

^f Department of Psychosis Studies, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK

* Corresponding author: Kristin R. Laurens, School of Psychology and Counselling, Queensland University of Technology, O Block B Wing Level 5, Victoria Park Road, Kelvin Grove QLD 4059, Australia. Email: kristin.laurens@qut.edu.au; Telephone: +61 (0)7 3138 4619; Facsimile: +61 (0)7 3138 0486.

Word Count: 3,963

Number of Tables: 4

Number of Figures: 1

Supplementary Online materials: 1 page

Abstract

Psychotic-like experiences (PLEs) are perceptual and thought disturbances that, although common among children, increase risk for future psychopathology, particularly if persistent. Clinical interviews are too time-consuming and costly to administer at a population level, but the criterion validity of a brief questionnaire for screening community samples of children as young as 9 years for PLEs has not been established. This study aimed to test the criterion (concurrent and predictive) validity of the Psychotic-Like Experiences Questionnaire for Children (PLEQ-C). The PLEQ-C (9-item self- and 10-item parent-report versions) was administered to 139 children aged 9-12 years and their caregivers recruited from Greater London, UK. Children additionally completed a diagnostic interview assessing hallucinations and delusions and three further PLEQ-C assessments at approximately 24-month intervals. Concordance of child- and caregiver-reports of PLEs on questionnaire (*PLE-Q*) was low. Self-reports of *any PLE-Q* demonstrated good sensitivity (73.3%), specificity (78.5%), positive and negative predictive values (PPV: 72.1%; NPV: 79.5%) for any PLE determined by interview (*PLE-I*), whereas caregiver-reports of *any PLE-Q* performed poorly (sensitivity 51.7%, specificity 78.5%, PPV 64.6%, NPV 68.1%). Multinomial regression analyses indicated that children reporting *any PLE-Q* at screening were at significantly increased risk of reporting PLEs on multiple assessments during adolescence relative to no PLEs, closely replicating the pattern and magnitude of effects (large-to-very large) obtained for children with *any PLE-I*. The PLEQ-C offers a valid, brief, feasible, and cost-effective means of community screening to identify children who present with PLEs and could be assessed with clinical interview.

Keywords: *psychometric properties; concurrent validity; predictive validity; cross-informant reports; psychopathology; psychotic symptoms*

1. Introduction

Psychotic-like experiences (PLEs), or subclinical hallucination and delusion symptoms, are commonly reported by young people. Meta-analyses indicate prevalence rates of 12.7-17.0% among children aged under 12 years and 7.5-12.4% among adolescents aged 13-17/18 years (Kelleher et al., 2012; Majer et al., 2018), relative to a mean lifetime prevalence of 5.8% in adults (Majer et al., 2018). These experiences may or may not be associated with distress, are typically transitory in nature, and are significantly more common than psychotic disorder (Dominguez et al., 2011; Linscott and van Os, 2013). However, some children who report PLEs are at higher risk of future psychotic and non-psychotic psychopathology (Kelleher et al., 2012, 2014; Nishida et al., 2010; Poulton et al., 2000; Welham et al., 2009a), with greater risk among those with PLEs that persist over time (Pedrero and Debbané, 2017). Of the children and adolescents who report PLEs, between one-third to one-half may experience persistent symptoms (Calkins et al., 2017; De Loore et al., 2011; Downs et al., 2013). As the likelihood of PLE persistence may be impacted by environmental exposures (Cougnard et al., 2007), being able to identify these children early may provide opportunities to mitigate or avoid cumulative environmental risk factors and lessen the risk of persistent PLEs and the associated negative outcomes. Feasible and cost-effective methods of identifying children in the community who present these experiences are imperative for preventative interventions.

Clinical interviews provide the gold standard for reliable and valid identification of psychotic symptoms, but they are too time-consuming and expensive to deliver at a population level. A number of brief questionnaires have been developed as a means of feasibly screening the general population for PLEs, including instruments designed for self-report by children and adolescents (Kelleher, 2015). Though rates of PLEs self-reported on questionnaires are typically higher than those determined by clinical interview, apparently “false positive” PLEs not confirmed at interview may nonetheless provide an important marker of possible future psychopathology (van der Steen et al., 2018). The instruments developed for self-report by young people have demonstrated variable psychometric properties, and typically, have been tested only for those aged 11 years or older. We previously demonstrated the construct validity of the self-report Psychotic-Like Experiences Questionnaire for Children (PLEQ; Laurens et al., 2007, 2012) in children aged 9-12 years. The availability of a valid tool for PLE self-report by children as young as 9 years is an important

provision for research and preventative intervention. Only low-to-moderate correspondence is typical between child- and parent-reports of other psychopathology, particularly for internalising (anxiety and depression) symptoms that are, like PLEs, not readily observable (De Los Reyes et al., 2015; van Der Ende et al., 2012). Thus, parents are likely to have limited capacity to provide valid reports of PLEs in population screening (Laurens et al., 2007; Welham et al., 2009b).

The present study thus aimed to test the criterion (concurrent) validity of the self- and parent-reported PLEQ-C (Laurens et al., 2007, 2012) against PLEs reported at clinical interview (*PLE-I*), using data from a community sample of children aged 9-12 years. We further sought to establish the predictive validity of the self-report PLEQ-C for later PLEs reported at multiple assessments conducted during a follow-up of approximately 4 years. Based on investigations of concurrent validity for similar self-report PLE instruments tested with children aged 11 years and older (Gundersen et al., 2019; Kelleher et al., 2011), we hypothesised that the self-report PLEQ-C would offer a valid means of identifying children aged 9-12 years who report *PLE-I*. By contrast, we anticipated that the parent-report PLEQ-C would show poor concurrent validity for child-reported *PLE-I*. We further hypothesised that children who reported PLEs on questionnaire when aged 9-12 years would be at increased risk of reporting PLEs at multiple questionnaire assessments conducted during adolescence (i.e., persistent PLEs) relative to reporting no PLEs at any later assessment.

2. Method

2.1 Participants

Data for this study were drawn from the London Child Health and Development Study (CHADS), a prospective, longitudinal investigation of children with premorbid risk markers for schizophrenia who were sampled from the general community and assessed biennially from the age of 9-12 years (Laurens and Cullen, 2016). The CHADS project follows the development of psychological, cognitive, and biological functioning in three groups of children: (i) individuals who presented with a triad of replicated antecedents of schizophrenia, including PLEs; (ii) individuals with a family history of schizophrenia or schizoaffective disorder in at least one first- or second-degree relative; and, (iii) typically developing children who presented no family history or antecedents. Exclusion criteria were a diagnosis of autism, Asperger's, or

neurological disorder, learning difficulties ($IQ < 70$), and prior use of antipsychotic medication. Children were recruited predominantly via school-based questionnaire screening for antecedents and family history of schizophrenia, which was conducted with children aged 9-12 years at 73 government and religious primary schools in Greater London, UK (Laurens et al., 2012). These children ($n=7966$; representing 95% of eligible children enrolled at participating schools) completed questionnaires independently and anonymously in class, with items read aloud by a researcher (4% of parents and 1% of children refused the child's participation, for reasons unknown). Caregivers ($n=1504$; 18.9%) completed corresponding questionnaires (matched by code number) at home and returned these via reply-paid mail. Among these, 799 families provided contact information and indicated willingness to be invited for further participation. A further 36 children aged 9-12 years who were relatives of patients with schizophrenia receiving treatment in the local health service were identified and approached following liaison with the patient's care worker.

From these 835 families, we invited 240 who met inclusion criteria for the CHADS project (the three groups described above), among whom 150 participated. For the present study, 139 children (mean age at questionnaire screening=10.4 years [SD 0.8]; 70 girls) had video recordings of self-reported diagnostic interviews available for PLE rating; recordings for the remaining children were unavailable due to technical problems and/or lack of consent. The ethnicity of the children (as reported by their caregivers within the Family Interview for Genetic Studies [FIGS]; Maxwell, 1992) included white British (30.9%), white other (22.3%), black African or African-Caribbean (28.8%), and other ethnicities (18.0%). Socio-economic status was indexed by the UK National Statistics Socioeconomic Classification of parental occupation, and included professional occupations (12.9%), managerial or technical occupations (54.0%), skilled occupations (22.6%), and unskilled occupations or unemployed (10.4%). Among the 139 children, 38 had a family history of a schizophrenia-spectrum disorder (including 13 recruited as child relatives of patients with schizophrenia rather than via community screening), and a further 8 had a family history of bipolar disorder; family history was confirmed via the FIGS assessment with the caregiver (Maxwell, 1992). Sensitivity analyses that excluded data from the 13 children recruited as relatives of patients with schizophrenia are reported in Supplementary Online materials, and indicate that results were unaffected when recruitment source was restricted to children identified via school-based screening.

2.2 Procedure

The measures completed at each assessment wave in the CHADS project are detailed in Laurens and Cullen (2016). Data from four waves were used in the present study: screening (T0), and assessments 1 (T1), 2 (T2), and 3 (T3). Children and their primary caregiver completed the PLEQ-C at T0 through T3, and a diagnostic interview at T1 only. The average lapse in time between screening (T0) and T1 assessments was 8 months (SD=5); T2 and T3 assessments were completed, on average, 24 months (SD=7) and 44 months (SD=9) following the T1 assessment. Ethical approval for the research was granted by the Joint South London and Maudsley National Health Service Foundation and the Institute of Psychiatry National Health Service Research Ethics Committee. Written informed consent and assent for participation in the study were obtained from caregivers and children, respectively.

2.3 Measures

2.3.1 PLEs from questionnaire (PLE-Q)

Child- and caregiver-reports of PLEs were gathered at T0-T3 using the self- and parent-report versions of the PLEQ-C respectively (Laurens et al., 2007, 2012, 2019). Children completed nine items assessing hallucination- and delusion-like experiences (Table 1), each rated on a 3-point scale (0=not true; 1=somewhat true; 2=certainly true). The PLEQ-C has excellent ordinal alpha reliability ($\alpha=.90$; Laurens et al., 2017) and demonstrated construct validity (all items load >0.5 on a single construct in the general population), with the auditory and visual hallucination items evidencing the best psychometric characteristics (item response theory discrimination and difficulty parameters; Laurens et al., 2012). Parents completed an additional item assessing bizarre thought/behaviour (Table 1).

Responses to all nine items of the self-report PLEQ-C were recoded into dichotomous variables indexing a high response threshold (where 0=not or somewhat true, and 1=certainly true). A total PLE questionnaire score (*total PLE-Q*; the sum of these nine dichotomous variables) was computed (range 0-9), and a dichotomous summary variable derived to indicate whether participants had experienced *any PLE-Q* (0=no PLE reported, 1=at least one PLE). Two further dichotomous variables, coded in the same manner, were created to indicate the presence of *any hallucination* (from items 4 and 9) or *any delusion* (items 1, 2, 3, 5, 6, 7, and 8).

Similar variables (*total PLE-Q score* [range 0-10], and the dichotomous variables *any PLE-Q*, *any hallucination*, *any delusion*) were constructed for caregiver responses to the 10-item parent-report PLEQ-C, but with a lower response threshold applied to indicate the presence of a PLE (i.e., where 0=not true, and 1=somewhat or certainly true).

To index the occurrence of PLEs over multiple waves of assessment (excluding screening [T0]), a four-level categorical variable was created from the child-reported PLEQ-C assessments at T1, T2, and T3. This variable was derived using the lower response threshold (where 0=not true, and 1=somewhat or certainly true), and coded the number of post-screening assessments at which a participant reported at least one PLE (0=no PLE at any assessment, 1=PLE present at any one assessment, 2=PLE present at any two assessments, 3=PLE present at all three assessments).

2.3.2 PLEs from diagnostic interview (PLE-I)

Children completed a semi-structured diagnostic interview, the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL), which assesses a range of psychopathology, including psychosis, in young people aged 6-18 years (Kaufman et al., 1997). Administration of the interview schedule was modified so that children responded not only to the hallucination and delusion screening prompts, but also to the entire psychotic disorders supplement that assesses a broad range of hallucination and delusion symptoms (supplement questions would normally be administered only following a positive response to either of the two screener items). In total, children reported on 10 hallucination and 13 delusion symptoms (Table 2). Interviews were conducted by psychiatrists and psychologists, and video-recorded for consensus rating. Approximately 15% of interviews were co-rated immediately from video by a child and adolescent psychiatrist (BWJ) in order to maintain fidelity of administration during the assessment period. For the present study, ratings of all 139 recordings were conducted independently by three raters with psychology training (CPL, AMT, and KRL), using a 3-point response scale for each item (0=PLE not present, 1=possible/weak PLE, 2=definite/strong PLE) employed previously (Kelleher et al., 2011). The clinical significance of definite/strong PLEs was agreed by all three raters, whereas possible/weak PLE ratings differentiated experiences of lower or less certain significance. The three raters achieved substantial inter-rater agreement ($\kappa=0.80$ for *any PLE-I*, 0.82 for

any hallucination, and 0.72 for *any delusion*), with any rating discrepancies resolved subsequently by consensus.

For analysis, each hallucination and delusion item was recoded into a dichotomous indicator defined at the lower symptom threshold (0=absent, 1=possible/weak or definite/strong PLE). A summary dichotomous variable was constructed to indicate the presence of *any PLE-I* (0=no PLE, 1=at least one PLE present at interview).

2.4 Statistical Analysis

All statistical analyses were conducted using IBM SPSS version 25.0 (IBM, 2017).

2.4.1 Criterion validity of questionnaire screening

Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and positive and negative likelihood ratios (LR+, LR-) were computed to determine the utility of the PLEQ-C for identifying PLEs at clinical interview. These indices were determined against *any PLE-I* for each of the nine child-reported PLEQ-C items individually, and for *any PLE-Q*, *any hallucination*, and *any delusion*. The LR reflects the balance of sensitivity and specificity independently of prevalence, such that LR+ (=sensitivity/1-specificity) indexes how much the likelihood of a PLE at interview increases given a PLE reported on questionnaire and LR- (=1-sensitivity/specificity) indexes how much the likelihood of PLE at interview decreases given no PLE on questionnaire.

A receiver operating characteristic (ROC) curve analysis determined the diagnostic capacity of the child-reported PLEQ-C at different discrimination thresholds, testing the *total PLE-Q score* obtained from the PLEQ-C against *any PLE-I*. We derived the Area Under the Curve (AUC), and computed Youden's Index [Sensitivity+Specificity-1] (Habibzadeh et al., 2016) to identify the optimal discrimination threshold for the PLEQ-C.

2.4.2 Caregiver-reports of PLEs

Concordance between child- and caregiver-ratings on the PLEQ-C was examined using Cohen's kappa, for each of the nine items reported by both informants, *any PLE-Q*, *any delusion*, and *any hallucination*. Inter-rater agreement was interpreted according to the definitions of Landis and Koch (1977): <0.00 Poor (no agreement); 0.00-0.20 Slight; 0.21-0.40 Fair; 0.41-0.60 Moderate; 0.61-0.80 Substantial; 0.81-1.00

Almost perfect.

We also examined the sensitivity, specificity, PPV, NPV, and likelihood ratios associated with caregiver-reports of *any PLE*, *any hallucination*, and *any delusion* on the PLEQ-C to child-reports of *any PLE-I*. ROC curve analysis (extracting AUC) determined the diagnostic capacity of the caregiver-reported PLEQ-C at different discrimination thresholds, with Youden's Index indicating the optimal discrimination threshold.

2.4.3 Predictive validity of PLE-Q and PLE-I for continuing PLEs

Separate multinomial logistic regression analyses examined the associations between child-reports of (a) *any PLE-Q* and (b) *any PLE-I* and the number of assessments at which children reported *any PLE-Q* at T1-T3 (i.e., 1, 2, or 3; relative to 0 reference). Effect sizes were determined as odds ratios (ORs), with ORs of 1.00–1.49 interpreted as small, 1.50–2.49 as medium, and 2.50–4.00 as large, and ≥ 10 as very large (Rosenthal, 1996).

3. Results

3.1 Prevalence of PLE-Q and PLE-I

Mean *total PLE-Q score* for child-reported PLEs was 1.20 (SD 1.79; range 0-9) and caregiver-reported PLEs was 0.86 (SD 1.58; range 0-8). Table 1 reports the prevalence of PLEs reported by children and caregivers on the PLEQ-C, with higher rates of positive endorsement by children than caregivers on most items, despite the higher response threshold adopted for child-reports. Rates of *any PLE-Q* were almost a third greater among children than caregivers and more than double for *any hallucination*, though only a fifth greater for *any delusion*. The hallucination items (*heard voices*, *seen things*) were the most prevalent PLEs reported by children, whereas for caregivers, mind reading delusions (*thoughts read*, *reads minds*) were most prevalent. Minimal inter-rater agreement between children and caregivers was apparent for *any PLE-Q*, *any delusion*, and each PLE item (Table 1); *any hallucination* and *heard voices* showed the best, though only "fair", agreement.

Table 2 summarises the prevalence of *any PLE-I* reported by children (43.1%). Hallucinations were almost twice as prevalent as delusions, with auditory hallucinations the most prevalent PLE. Persecution was the most common delusion.

3.2 Criterion Validity

Performance (specificity, sensitivity, PPV, NPV, LR+ and LR-) of the PLEQ-C relative to *any PLE-I* is detailed in Table 3. Approximately three quarters of *PLE-I* positive and negative cases were correctly identified by questionnaire. The predictive performance for self-reported *any PLE-Q* was superior to that apparent for any individual questionnaire item, with the criterion validity for *any hallucination* and *any delusion* approaching that for *any PLE-Q*.

ROC curve analysis (Figure 1) indicated a fair performance for the child-reported PLEQ-C in identifying *any PLE-I* (AUC=0.78, $p < 0.001$, 95% CI [0.70, 0.86]). According to Youden's Index, a single child-reported *certainly true* response on the PLEQ-C (i.e., *total PLE-Q* score=1) already achieved optimal discrimination in identifying children reporting *any PLE-I*.

Performance of the caregiver-reported PLEQ-C was notably poorer than the child-reported PLEQ-C (Table 3). Only half of children who self-reported *any PLE-I* were identified by the caregiver questionnaire. ROC curve analysis (Figure 1) indicated a poor performance for the caregiver-reported PLEQ-C in identifying *any PLE-I* (AUC=0.67, $p < 0.001$, 95% CI [0.58, 0.77]). Youden's Index indicated optimal discrimination already at single PLE item endorsement (somewhat or certainly true).

3.3 Predictive validity of any PLE-Q and any PLE-I for later PLEs

Multinomial regression analyses (Table 4) indicated that children who self-reported *any PLE-Q* at screening (T0) were at significantly increased odds of reporting PLEs at multiple subsequent assessments (i.e., significantly more likely to report PLEs at two and three later assessments); these effects were large to very large in magnitude (OR=4.00 for PLEs at two assessments, OR=16.63 at three assessments). Similarly, children presenting *any PLE-I* were also at significantly increased odds of persistent PLEs across the T1-T3 assessments; these effects were similarly large (OR=4.17 for PLEs at two assessments) to very large (OR=10.00 for PLEs at three assessments) in magnitude.

4. Discussion

The present study demonstrates the validity of the child-report PLEQ-C in population screening to identify children who self-report PLEs on diagnostic interview. The study further demonstrates low rates of

concordance between child- and caregiver-responses on the PLEQ-C, and poor capacity of caregiver-reported PLEQ-C data to identify children with *any PLE-I*. Children who reported *any PLE-Q*, similarly for children reporting *any PLE-I*, experienced significantly increased odds of reporting persistent PLEs (i.e., on at least two assessments) over a follow-up period of approximately 4 years, on average.

The PLEQ-C thus offers a sound, brief, and cost-effective means of screening community samples to identify 9- to 12-year-olds who may experience hallucinations or delusions. The screening tool does not replace clinical interviews as a means of confirming psychotic symptoms or determining their clinical significance, but is useful in identifying a subset of children who could feasibly undergo such comprehensive assessment. Previous evaluations of the performance of self-report PLE questionnaires against the K-SADS-PL interview have focused on children aged 11-13 years (Gundersen et al., 2019; Kelleher et al., 2011); we extend that work by demonstrating the valid use of the self-report PLEQ-C with children aged as young as 9 years. While most performance indices across studies were similar, our PPV of *any PLE-Q* for *any PLE-I* (72.1%) was higher than that reported by Gundersen and colleagues (26.8%), which might reflect the different style of questionnaire items employed in that study (i.e., brief vignettes providing an example of each experience for a child of the same gender) and/or sampling differences (e.g., such that the prevalence of *any PLE-Q* in that study was only 28.1% relative to 43.9% in the present study). Similarly to both Gundersen et al. (2019) and Kelleher et al. (2011), our study demonstrated only poor-to-modest performance of specific PLE items in identifying *any PLE-I* compared to the discriminative capacity of *any PLE-Q*; that is, individual PLE items were less likely to identify children who reported PLEs at interview, and more likely to falsely classify individuals as asymptomatic, than *any PLE-Q*. Also consistent with those previous investigations, the auditory and visual hallucination items performed better than individual delusion items in identifying children with *any PLE-I*. Nonetheless, two findings imply that assessing a range of PLE items in questionnaire-based screening of children in the community may offer the best prospect for identifying symptomatic individuals, namely: (1) the ROC curve analysis indicated that any single PLE report on questionnaire was already optimally capable of distinguishing children with *any PLE-I* from children without *any PLE-I*, and (2) the intermediate test performance of *any hallucination* and *any delusion* between that of *any PLE-Q* and the individual PLE items. Anecdotally, we observed that multiple

children in our sample reported no PLEs in response to the K-SADS-PL screener prompts, but subsequently revealed strong/definite PLEs when each of the supplement hallucination and delusion items were assessed. Thus, children aged 9-12 years may not always generalise from the screening prompts to their own experiences, potentially leading to psychotic symptoms being missed if a range of hallucinations and delusions are not assessed.

As predicted, caregiver reports of PLEs were a poor means of identifying children with *any PLE-I*. Only approximately one-half (51.7%) of caregivers identified PLEs that were self-reported by their child at interview. This might reflect their ignorance of experiences that are not readily observable, or that they ascribe another meaning to the experiences. The present findings align with those from adolescents (Nugent et al., 2013) in suggesting that caregivers are not a substitute informant for PLEs in children.

As hypothesised, a positive response to any item on the PLEQ-C at screening was indicative of an increased risk of PLEs at multiple reassessments spanning a period of ~4 years, with a dose-response type effect indicating strong likelihood of persistent PLEs. A similar pattern and magnitude of effects was obtained for the association of *any PLE-I* with reports of PLEs at two and three further assessments, which suggests that the pattern obtained for *any PLE-Q* does not simply reflect common measurement effects from the repeated administration of the instrument. The capacity of the PLEQ-C to identify children in the community who are at risk of PLEs that persist into adolescence offers the prospect of assisting them to manage or avoid environmental risk moderators, such as cannabis use, trauma, and urbanicity, that might lead to future psychopathology (Cougnard et al., 2007).

A strength of the present study is the availability of longitudinal data that permitted an assessment of predictive validity alongside concurrent validity of the PLEQ-C. The study extends previous research by demonstrating the utility of a self-report PLE screening questionnaire for children as young as 9 years, which is critical given that caregiver-reports do not provide an adequate substitute for self-reports. Nonetheless, several study limitations should be considered. First, the extended (average 8 months) and variable delay between questionnaire screening (T0) and diagnostic interview assessments (T1) may have introduced error, though this would be likely to underestimate the criterion validity of the PLEQ-C against interview. Second, the study oversampled children from inner-city London boroughs where rates of

schizophrenia/psychosis are elevated and self-reports of PLEs on the PLEQ-C are high relative to those obtained by similar screening questionnaires employed with other samples (Kelleher et al., 2012; Laurens et al., 2012, 2017), while also recruiting fewer children from lower socioeconomic backgrounds than are typical of those inner London communities. Despite this, our findings on criterion validity of the PLEQ-C relative to the K-SADS-PL diagnostic interview generally aligned well with investigations of similar questionnaires in other samples/locales (Gundersen et al., 2019; Kelleher et al., 2011). Third, the screening questionnaire was read aloud to children during the classroom administration, and whether the instrument would demonstrate similar psychometric properties if children were required to read items independently is unknown. Finally, we did not distinguish PLEs associated with distress or functional impairment from non-distressing/impairing PLEs for which interventions are not indicated. This should be considered in future research, given that the former may be more strongly associated with PLE persistence and a range of poor mental health outcomes, and have consequently been prioritised as a treatment target of intervention for young people with psychotic symptoms (Jolley et al., 2018).

The PLEQ-C offers a brief, valid, feasible, and cost-effective means of screening community samples of children aged as young as 9 years to identify individuals reporting PLEs. These children might be subsequently assessed by diagnostic interview to determine the clinical significance of their experiences and/or monitored over time to determine their persistence and impact, so that appropriate and timely psychological intervention may be offered. While caregivers provide an important perspective when evaluating their children's PLEs in clinical interview, they do not constitute a valid alternative informant for community-based screening for PLEs, as around half of children reporting PLEs at interview would be missed by parent-report. The brevity of the PLEQ-C means that it might feasibly be incorporated alongside instruments that assess internalising and externalising symptoms when screening for psychopathology in children.

6. References

- Calkins, M.E., Moore, T.M., Satterthwaite, T.D., Wolf, D.H., Turetsky, B.I., Roalf, D.R., Merikangas, K.R., Ruparel, K., Kohler, C.G., Gur, R.C., Gur, R.E., 2017. Persistence of psychosis spectrum symptoms in the Philadelphia Neurodevelopmental Cohort: A prospective two-year follow-up. *World Psychiatry* 16(1), 62-76.

- Cougnard, A., Marcelis, M., Myin-Germeys, I., De Graaf, R., Vollebergh, W., Krabbendam, L., Lieb, R., Wittchen, H.U., Henquet, C., Spauwen, J., Van Os, J., 2007. Does normal developmental expression of psychosis combine with environmental risk to cause persistence of psychosis? A psychosis proneness-persistence model. *Psychological medicine* 37(4), 513-527.
- De Loore, E., Gunther, N., Drukker, M., Feron, F., Sabbe, B., Deboutte, D., van Os, J., Myin-Germeys, I., 2011. Persistence and outcome of auditory hallucinations in adolescence: a longitudinal general population study of 1800 individuals. *Schizophrenia research* 127(1-3), 252-256.
- De Los Reyes, A., Augenstein, T.M., Wang, M., Thomas, S.A., Drabick, D.A.G., Burgers, D.E., Rabinowitz, J., 2015. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological bulletin* 141(4), 858.
- 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. *Schizophrenia bulletin* 37(1), 84-93.
- Downs, J.M., Cullen, A.E., Barragan, M., Laurens, K.R., 2013. Persisting psychotic-like experiences are associated with both externalising and internalising psychopathology in a longitudinal general population child cohort. *Schizophrenia research* 144(1-3), 99-104.
- Gundersen, S.V., Goodman, R., Clemmensen, L., Rimvall, M.K., Munkholm, A., Rask, C.U., Skovgaard, A.M., Van Os, J., Jeppesen, P., 2019. Concordance of child self-reported psychotic experiences with interview- and observer-based psychotic experiences. *Early intervention in psychiatry* 13(3), 619-626.
- Habibzadeh, F., Habibzadeh, P., Yadollahie, M., 2016. On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochemia Medica* 26(3), 297-307.
- IBM, 2017. IBM SPSS Statistics for Windows, 25.0 ed. IBM Corp., Armonk, NY.
- Jolley, S., Kuipers, E., Stewart, C., Browning, S., Bracegirdle, K., Basit, N., Gin, K., Hirsch, C., Corrigan, R., Banerjee, P., Turley, G., Stahl, D., Laurens, K.R., 2018. The Coping with Unusual Experiences for Children Study (CUES): A pilot randomized controlled evaluation of the acceptability and potential clinical utility of a cognitive behavioural intervention package for young people aged 8–14 years with unusual experiences and emotional symptoms. *British Journal of Clinical Psychology* 57(3), 328-350.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U.M.A., Flynn, C., Moreci, P., Williamson, D., 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. *Journal of the American Academy of Child & Adolescent Psychiatry* 36(7), 980-988.
- Kelleher, I., 2015. The assessment of psychosis: A reference book and rating scales for research and practice, in: Waters, F., Stephane, M. (Eds.). Routledge, New York, pp. 123-135.
- Kelleher, I., Cederlof, M., Lichtenstein, P., 2014. Psychotic experiences as a predictor of the natural course of suicidal ideation: a Swedish cohort study. *World Psychiatry* 13(2), 184-188.
- 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(9), 1857-1863.
- Kelleher, I., Harley, M., Murtagh, A., Cannon, M., 2011. Are screening instruments valid for psychotic-like experiences? A validation study of screening questions for psychotic-like experiences using in-depth clinical interview. *Schizophrenia bulletin* 37(2), 362-369.

- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1), 159.
- Laurens, K.R., Cullen, A.E., 2016. Toward earlier identification and preventative intervention in schizophrenia: Evidence from the London Child Health and Development Study. *Social Psychiatry and Psychiatric Epidemiology: The International Journal for Research in Social and Genetic Epidemiology and Mental Health Services* 51(4), 475-491.
- 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(7), 1495-1506.
- 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. *Schizophrenia research* 90(1-3), 130-146.
- Laurens, K.R., Tzoumakis, S., Dean, K., Brinkman, S.A., Bore, M., Lenroot, R.K., Smith, M., Holbrook, A., Robinson, K.M., Stevens, R., Harris, F., Carr, V.J., Green, M.J., 2017. The 2015 Middle Childhood Survey (MCS) of mental health and well-being at age 11 years in an Australian population cohort. *BMJ Open* 7(6), e016244.
- Laurens, K.R., Tzoumakis, S., Dean, K., Harris, F., Carr, V.J., Green, M.J., 2019. Population profiles of child-reported psychotic-like experiences and their differential association with other psychopathologies. *The British journal of clinical psychology*.
- 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. *Psychological medicine* 43(6), 1133-1149.
- Maijer, K., Begemann, M.J.H., Palmen, S., Leucht, S., Sommer, I.E.C., 2018. Auditory hallucinations across the lifespan: a systematic review and meta-analysis. *Psychological medicine* 48(6), 879-888.
- Maxwell, M.E.J.B., MD: Clinical Neurogenetics Branch, Intramural Research Program, National Institute of Mental Health, 1992. Family Interview for Genetic Studies (FIGS): a manual for FIGS.
- Nishida, A., Sasaki, T., Nishimura, Y., Tanii, H., Hara, N., Inoue, K., Yamada, T., Takami, T., Shimodera, S., Itokawa, M., Asukai, N., Okazaki, Y., 2010. Psychotic-like experiences are associated with suicidal feelings and deliberate self-harm behaviors in adolescents aged 12-15 years. *Acta psychiatrica Scandinavica* 121(4), 301-307.
- Nugent, K.L., Kline, E., Thompson, E., Reeves, G., Schiffman, J., 2013. Assessing psychotic-like symptoms using the BASC-2: Adolescent, parent and teacher agreement. *Early intervention in psychiatry* 7(4), 431-436.
- Office for National Statistics, 2010. The national statistics socio-economic classification. Office for National Statistics., Fareham.
- Pedrero, E.F., Debbané, M., 2017. Schizotypal traits and psychotic-like experiences during adolescence: An update. *Psicothema* 29(1), 5-17.
- Poulton, R., Caspi, A., Moffitt, T.E., Cannon, M., Murray, R., Harrington, H., 2000. Children's Self-Reported Psychotic Symptoms and Adult Schizophreniform Disorder: A 15-Year Longitudinal Study. *Archives of General Psychiatry* 57(11), 1053-1058.

- Rosenthal, J.A., 1996. Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21(4), 37-59.
- van Der Ende, J., Verhulst, F.C., Tiemeier, H., 2012. Agreement of informants on emotional and behavioral problems from childhood to adulthood. *Psychological assessment* 24(2), 293.
- van der Steen, Y., Myin-Germeys, I., van Nierop, M., Ten Have, M., de Graaf, R., van Dorsselaer, S., van Os, J., van Winkel, R., 2018. 'False-positive' self-reported psychotic experiences in the general population: an investigation of outcome, predictive factors and clinical relevance. *Epidemiol Psychiatr Sci*, 1-12.
- Welham, J., Isohanni, M., Jones, P., McGrath, J., 2009a. The Antecedents of Schizophrenia: A Review of Birth Cohort Studies. *Schizophrenia bulletin* 35(3), 603-623.
- Welham, J., Scott, J., Williams, G., Najman, J., Bor, W., O'Callaghan, M., McGrath, J., 2009b. Emotional and behavioural antecedents of young adults who screen positive for non-affective psychosis: a 21-year birth cohort study. *Psychological medicine* 39(4), 625-634.

Figure 1. Receiver operating characteristic curve illustrating the capacity of the child- and caregiver-report questionnaires at different discrimination thresholds to detect any psychotic-like experience reported by children on diagnostic interview

Table 1. The prevalence and inter-rater agreement of *certainly true* responses by children and *somewhat/certainly true* responses by caregivers to questionnaire items

Questionnaire Item	Abbreviation	Prevalence		Inter-rater agreement			
		Child-report (%)	Caregiver-report (%)	Kappa statistic	95% CI	<i>p</i>	Level of agreement
1. Some people believe that their thoughts can be read. Have other people ever read your thoughts?	Thoughts read	9.4	20.1	-0.04	(-0.17, 0.10)	0.64	Poor (no agreement)
2. Have you ever believed that you were being sent special messages through the television?	Special messages	5.0	5.8	-0.06	(-0.09, -0.03)	0.50	Poor (no agreement)
3. Have you ever thought that you were being followed or spied upon?	Spied upon	16.5	9.4	0.18	(-0.02, 0.38)	0.03	Slight
4. Have you ever heard voices that other people could not hear?	Heard voices	23.7	10.1	0.28	(0.09, 0.46)	<0.01	Fair
5. Have you ever felt that you were under the control of some special power?	Controlled	7.2	2.2	0.13	(-0.14, 0.39)	0.08	Slight
6. Have you ever known what another person was thinking even though that person wasn't speaking?	Reads minds	15.1	16.5	0.08	(-0.10, 0.27)	0.33	Slight
7. Have you ever felt as though your body had been changed in some way that you could not understand?	Body changed	12.2	2.9	-0.05	(-0.09, -0.01)	0.45	Poor (no agreement)
8. Do you have any special powers that other people don't have?	Special powers	12.9	5.8	0.00	(-0.15, 0.14)	0.96	Poor (no agreement)
9. Have you ever seen something or someone that other people could not see?	Seen things	18.0	5.8	0.11	(-0.07, 0.29)	0.13	Slight
10. Have you ever seriously worried that your child was thinking or acting in a bizarre way?	Bizarre		7.2				
<i>Any psychotic-like experience on questionnaire</i>	<i>Any PLE-Q</i>	43.9	34.5	0.15	(-0.01, 0.31)	0.08	Slight
<i>Any hallucination</i>		29.5	13.7	0.26	(0.09, 0.43)	<0.01	Fair
<i>Any delusion</i>		38.1	31.7	0.04	(-0.13, 0.20)	0.65	Slight

Table 2. Prevalence of child-reported psychotic-like experiences at diagnostic interview (*PLE-I*)

	Weak or strong symptom %	Weak symptom %	Strong symptom %
<i>Any PLE-I</i>	43.1	16.5	26.6
<i>Any hallucination</i>	40.3	15.8	24.5
Auditory hallucinations	30.2	8.6	21.6
Non-verbal sounds (e.g., music)	13.7	7.9	5.8
Command hallucinations	5.0	0	5.0
Running commentary (commenting voice)	5.0	1.4	3.6
Conversing voice	5.8	0	5.8
Thoughts aloud	8.0	2.2	5.8
Other verbal hallucinations	10.0	1.4	8.6
Visual hallucinations	17.9	1.4	16.5
Tactile hallucinations	9.3	7.9	1.4
Olfactory hallucinations	3.6	3.6	0
Illusions	15.8	15.1	0.7
<i>Any delusion</i>	22.3	12.9	9.4
Grandiosity	5.7	4.3	1.4
Guilt/Sin	0	0	0
Delusions of control	2.9	2.2	0.7
Somatic delusions	3.6	3.6	0
Nihilism	0.7	0.7	0
Thought broadcasting	2.9	0	2.9
Thought insertion	2.8	1.4	1.4
Thought withdrawal	2.9	0.7	2.2
Message from television/radio	2.9	2.2	0.7
Delusions of persecution	6.5	4.3	2.2
Delusions that others can read his/her mind	5.0	4.3	0.7
Delusions of reference	2.1	0.7	1.4
Other bizarre delusions	2.8	1.4	1.4
Read others' minds	2.9	2.2	0.7

Table 3. Sensitivity, specificity, positive and negative predictive values (PPV, NPV), and positive and negative likelihood ratios (LR+, LR-) of child- and caregiver-responses on questionnaire for child-reports of any psychotic-like experience reported at interview (*any PLE-I*)

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR-
Child-report:						
4. Heard voices	41.7	89.9	75.8	67.0	4.12	.65
9. Seen things	35.0	94.9	84.0	65.8	6.91	.69
1. Thoughts read	15.0	94.9	69.2	59.5	2.96	.90
2. Special messages	10.0	98.7	85.7	58.8	7.80	.91
3. Spied upon	26.7	91.1	69.6	62.1	3.01	.81
5. Controlled	13.3	97.5	80.0	59.7	5.27	.89
6. Reads minds	26.7	93.7	76.2	62.7	4.21	.78
7. Body changed	18.3	92.3	64.7	59.5	2.38	.89
8. Special powers	22.0	93.7	72.2	61.7	3.48	.83
<i>Any hallucination</i>	<i>55.0</i>	<i>89.9</i>	<i>80.5</i>	<i>72.4</i>	<i>5.43</i>	<i>.50</i>
<i>Any delusion</i>	<i>61.7</i>	<i>79.7</i>	<i>69.8</i>	<i>73.3</i>	<i>3.05</i>	<i>.48</i>
<i>Any PLE-Q</i>	<i>73.3</i>	<i>78.5</i>	<i>72.1</i>	<i>79.5</i>	<i>3.41</i>	<i>.34</i>
Caregiver-report:						
<i>Any hallucination</i>	<i>30.0</i>	<i>98.7</i>	<i>94.7</i>	<i>65.0</i>	<i>23.70</i>	<i>.71</i>
<i>Any delusion</i>	<i>46.7</i>	<i>79.7</i>	<i>63.6</i>	<i>66.3</i>	<i>2.30</i>	<i>.67</i>
<i>Any PLE-Q</i>	<i>51.7</i>	<i>78.5</i>	<i>64.6</i>	<i>68.1</i>	<i>2.40</i>	<i>.62</i>

Note: *Any PLE-Q* = at least one psychotic-like experience reported on questionnaire (at certainly-true level for child-reported PLEs, and at somewhat- or certainly-true level for caregiver-reported PLEs).

Table 4. Associations between (a) any child-reported psychotic-like experience on questionnaire (*any PLE-Q*) at time 0 (T0) and, (b) any psychotic-like experience on interview (*any PLE-I*), and the number of assessments at which at least one psychotic-like experience was self-reported by children (from time 1, time 2, or time 3).

Assessments with a PLE (%)		(a) <i>Any PLE-Q</i> (T0)		(b) <i>Any PLE-I</i>	
		OR (95% CI)	p	OR (95% CI)	p
No PLE at any assessment	(20.8)	(reference group)		(reference group)	
PLE at 1 assessment	(28.0)	3.50 (0.86 - 14.33)	.08	1.11 (0.28 - 4.47)	0.88
PLE at 2 assessments	(28.0)	4.00 (1.00 - 16.26)	.05	4.17 (1.17 - 14.89)	0.03
PLE at all 3 assessments	(23.0)	16.63 (3.84 - 71.93)	<.001	10.00 (2.62 - 38.15)	<0.01

Note: PLE = psychotic-like experience; OR = odds ratio; CI = confidence interval

Gutteridge et al.; *Criterion validity of the Psychotic-Like Experiences Questionnaire for Children*
(PLEQ-C)

Supplementary Online Materials

As a sensitivity analysis that considered the potential effects of recruitment source, the concurrent validity of the child- and caregiver-reports of *any psychotic-like experience on questionnaire (PLE-Q)* relative to *any PLE reported at interview (PLE-I)* was determined after excluding data from the 13 children recruited as relatives of patients with schizophrenia rather than via school-based screening.

Results remained consistent for both child-reported data (sensitivity 72.7%, specificity 80.3%, PPV 74.1%, NPV 79.2%, LR+ 3.69, and LR- .34) and caregiver-reported data (sensitivity 52.7%, specificity 78.9%, PPV 65.9%, NPV 68.3%, LR+ 2.50, and LR- .60).