Heuvelman, H., Nazroo, J., & Rai, D. (2018). Investigating ethnic variations in reporting of psychotic symptoms: a multiple-group confirmatory factor analysis of the Psychosis Screening Questionnaire. *Psychological Medicine*.
<https://doi.org/10.1017/S0033291718000399>

Peer reviewed version

Link to published version (if available):
[10.1017/S0033291718000399](https://doi.org/10.1017/S0033291718000399)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via CUP at <https://www.cambridge.org/core/journals/psychological-medicine/article/investigating-ethnic-variations-in-reporting-of-psychotic-symptoms-a-multiplegroup-confirmatory-factor-analysis-of-the-psychosis-screening-questionnaire/2015B151B1FD939F0D5CA1228FCD8CF5>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

For re-submission to Psychological Medicine:

Investigating ethnic variations in reporting of psychotic symptoms: a multiple-group confirmatory factor analysis of the Psychosis Screening Questionnaire

Hein Heuvelman PhD^{* 1, 2}, James Nazroo PhD², Dheeraj Rai PhD¹

¹ Centre for Academic Mental Health, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, BS8 2BN, Bristol, UK

² The Cathie Marsh Institute for Social Research (CMIST), School of Social Sciences, University of Manchester, Humanities Bridgeford Street Building, M13 9PL, Manchester, UK

* Correspondence to: Hein Heuvelman PhD, Centre for Academic Mental Health, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, BS8 2BN, Bristol, UK

Source of support: This study was carried out as part of an ESRC funded PhD studentship.

Conflict of interest: None.

Word count: 4,422 words, 1 figure, 3 tables, 8 supplements

Abstract

Background: Epidemiological evidence suggests risk for psychosis varies with ethnicity in Western countries. However, there is little evidence to date on the cross-cultural validity of screening instruments used for such comparisons.

Methods: Combining two existing UK population-based cohorts, we examined risk for reporting psychotic symptoms across White British (n=3,467), White Irish (n=851), Caribbean (n=1,899), Indian (n=2,590), Pakistani (n=1,956) and Bangladeshi groups (n=1,248). We assessed the psychometric properties of the Psychosis Screening Questionnaire with a multiple-group confirmatory factor analysis, assessing the equivalence of factor loadings, response thresholds and residual variances in an analysis of measurement non-invariance.

Results: Compared with prevalence among British Whites (5.4%), prevalence of self-reported psychotic symptoms was greater in the Caribbean group (12.7%, adjusted OR=2.38 [95% CI 1.84–3.07]). Prevalence was also increased among Pakistani individuals (8.3%, adjusted OR=1.36 [1.01–1.84]) although this difference was driven by a greater likelihood of reporting paranoid symptoms. PSQ items for thought interference, strange experience and hallucination were measured in equivalent ways across ethnic groups. However, our measurement models suggested that paranoid symptoms were measured less reliably among ethnic minorities than among British Whites and appeared to exaggerate latent differences between Pakistani and White British groups when measurement non-invariance was not accounted for.

Conclusions: Notwithstanding evidence for measurement non-invariance, the greater risk for reporting psychotic symptoms among Caribbean individuals is unlikely to be an artefact of measurement. Greater residual variance in the recording of paranoid symptoms among ethnic minority respondents warrants caution in using this item to investigate ethnic variation in psychosis risk.

Keywords

Psychosis Screening Questionnaire; self-reported psychotic symptoms; ethnicity; measurement non-invariance; multiple-group confirmatory factor analysis; population-based study.

1 Introduction

2 Psychosis is the defining characteristic of schizophrenia-spectrum disorder, although psychotic symptoms
3 are encountered across the wider range of major mental disorder (World Health Organisation, 2010) and in
4 the general population (Cohen and Marino, 2013, Johns *et al.*, 2002, King *et al.*, 2005, Nazroo, 1997, Nuevo
5 *et al.*, 2012, Vanheusden *et al.*, 2008). It is a mental state characterised by distortions in thought and
6 perception and inappropriate affect, which may involve hallucinations, delusions, excessive and unwanted
7 suspicions, or abnormalities of behaviour (World Health Organisation).

8 The expression of psychosis may vary with cultural background (Adebimpe *et al.*, 1981, Arnold *et al.*, 2004,
9 Barrio *et al.*, 2003, Bauer *et al.*, 2011, Chang *et al.*, 2011, Chu *et al.*, 1985, Littlewood and Lipsedge, 1981,
10 Maslowski *et al.*, 1998, Radhakrishnan *et al.*, 1983, Suhail and Cochrane, 2002, Thomas *et al.*, 2007, Weisman
11 *et al.*, 2000, Yamada *et al.*, 2006). There may be cultural differences in the content of psychotic symptoms
12 (Maslowski *et al.*, 1998, Viswanath and Chaturvedi, 2012, Weisman *et al.*, 2000, Yamada *et al.*, 2006) and/or
13 in the language with which these symptoms are expressed (Kleinman, 1987). In addition, the onset and
14 expression of psychosis may be influenced by social context, as suggested by the higher prevalence of
15 paranoid symptoms among ethnic minorities diagnosed with psychosis in Western countries (Barrio *et al.*,
16 2003, Fabrega *et al.*, 1994, Kendler, 1982, Littlewood and Lipsedge, 1981, Suhail and Cochrane, 2002, Veling
17 *et al.*, 2007, Whaley, 1998). These issues complicate the monitoring of psychosis risk as it becomes uncertain
18 whether group differences in symptom prevalence reflect underlying differences in levels of mental ill-
19 health, or whether they are a consequence of the way in which groups with different ethnic or cultural
20 backgrounds interact with diagnostic assessment or screening instruments.

21 Population-based studies of ethnic variation in psychosis risk have relied on the assumption that self-
22 reported symptoms reflect levels of mental ill-health in equivalent ways across groups (Cohen and Marino,
23 2013, Johns *et al.*, 2002, King *et al.*, 2005, Nazroo, 1997, Vanheusden *et al.*, 2008). This assumption, however,

is rarely tested explicitly, so much of the evidence pertaining to ethnic variation in psychosis risk therefore remains limited by the possibility of ethnic or cultural bias.

We examined data from two existing epidemiological studies to formally test the assumption of cross-ethnic equivalence through an analysis of measurement non-invariance in response provided to the Psychosis Screening Questionnaire (PSQ). Furthermore, we sought to compare ethnic differences based on observed PSQ response with differences based on latent variable means to illustrate the potential implications of assuming psychometric equivalence of the PSQ in a cross-ethnic context.

Our aims were to: (1) examine ethnic differences in the prevalence of self-reported psychotic symptoms; (2) compare across ethnic groups the measurement properties of the PSQ; and (3) evaluate ethnic differences in observed and latent response patterns, considering potential ethnic biases in PSQ screening response.

Methods

Study population

This study uses data from the Fourth National Survey of Ethnic Minorities (FNSEM: conducted during 1993/4 in England and Wales, n=8,207) (Modood *et al.*, 1997) and the Ethnic Minority Psychiatric Illness Rates in the Community study (EMPIRIC: conducted during 2000/1 in England, n=4,281) (Sproston and Nazroo, 2002, UK Data Archive). The FNSEM employed a multistage stratified random sample using the Postcode Address File (PAF) as a sampling frame and contains within it a boosted sample of ethnic minorities. EMPIRIC consists of a random subset of White British individuals recruited into the 1998 Health Survey for England (HSE) (King *et al.*, 2005) and all individuals recruited into the 1999 HSE (Erens *et al.*, 2001) who had agreed to be contacted for further interviewing (Sproston and Nazroo, 2002). The 1999 HSE contained a boosted sample of ethnic minorities, and both the 1998 and 1999 HSE employed similar sampling strategies to the FNSEM. Given similarities in design, coverage of the same groups and use of identical measures, we were able to combine FNSEM and EMPIRIC to obtain a uniquely powered population-based sample of White British and

ethnic minority groups in England and Wales. Further details on the construction of our study population are provided in an online supplement (Supplement 1).

We excluded individuals with missing data on ethnicity (n=263) and Chinese individuals (n=214) as a preliminary screening of the items suggested symptom prevalence in this group was too low for the purpose of our study. This left a sample of n=12,011 individuals, of whom n=3,467 identified as White British; n=851 as White Irish; n=1,899 as Caribbean; n=2,590 as Indian; n=1,956 as Pakistani; and n=1,248 as Bangladeshi.

Psychotic symptoms

We assessed the presence of five self-reported psychotic symptoms (mania, thought interference, paranoia, strange experience, hallucination) using the Psychosis Screening Questionnaire (PSQ) (Bebbington and Nayani, 1995). Each was assessed with a root question to assess the presence of psychosis-like experience, and one or two targeted questions to corroborate the experience as being symptomatic to psychosis (Table 1).

We derived a dichotomous measure for each of the five symptoms, capturing endorsement of the root and targeted questions. In addition to these five dichotomous measures, we constructed a composite screening measure capturing response across all five symptoms (0=negative on all; 1=positive on any). We excluded 156 individuals (1.3%) with missing values on PSQ screening variables.

Measurement model

Measurement models assume that an observed (or manifest) variable can be considered a proxy for an unobserved (or latent) variable that may be of primary interest to the researcher. Latent variables are assumed to *cause* manifest variables in the same way that having poor mental health (a latent trait) may cause one to report symptoms (a manifest response). As latent variables cannot be directly observed, they are instead inferred from patterns of correlation among the manifest variables. Given our use of categorical data, we employed a continuous latent response model (Millsap and Yun-Tein, 2004) assuming that observed dichotomous variable $[x]$ is determined by unobserved scores on a latent continuous response variable $[x^*]$,

which itself is determined by an unobserved score on latent construct $[\eta]$. This recognises that, while symptoms are typically recorded as dichotomous traits (i.e. absence versus presence), the underlying experience of symptoms may follow a continuous distribution, in which only a sufficient amount of latent continuous response $[\mathbf{x}^*]$ would result in positive response on the manifest variable $[\mathbf{x}]$. Furthermore, the specification of a latent variable $[\eta]$ in relation to multiple latent continuous response variables $[\mathbf{x}^*]$ allows for conceptual distinction between an underlying cause of mental ill-health (e.g. psychosis) and the symptoms ensuing from this cause (e.g. hallucination).

We illustrate this conceptual model in Figure 1a. Latent variable $[\eta_a]$ is assumed to cause three continuous latent response variables $[\mathbf{x}_{1a}^*, \mathbf{x}_{2a}^*, \mathbf{x}_{3a}^*]$, with the strength of these associations captured by factor loadings $[\lambda_{1a}, \lambda_{2a}, \lambda_{3a}]$. The continuous latent response variables then in turn cause positive response on manifest variables $[\mathbf{x}_{1a}, \mathbf{x}_{2a}, \mathbf{x}_{3a}]$ when level of latent continuous response exceeds the value of item-specific response thresholds $[\tau_{1a}, \tau_{2a}, \tau_{3a}]$. Furthermore, variation in the manifest variables that cannot be accounted for by the continuous latent response variables is captured as residual variance $[\theta_{1a}, \theta_{2a}, \theta_{3a}]$. It is worth emphasising that a factor loading may therefore be thought of as a measure of strength of association between the latent trait and a latent continuous response variable; the response threshold as the level of latent continuous response required for endorsement of the manifest variable; and the residual variance as a measure of the reliability with which a manifest variable captures the latent continuous response variable (i.e. with smaller residual variance equalling greater reliability).

Measurement non-invariance

The application of a continuous latent response model to a multiple-group setting enables the comparison of model parameters across groups (Figures 1a and 1b). Specifically, comparing factor loadings, response thresholds and residual variances across groups can inform the researcher about group differences in terms of i) the strength of association of the latent trait with the latent continuous response variables, ii) levels of latent continuous response at which the manifest variables are endorsed, and iii) the reliability with which the manifest variables capture the latent continuous response variables. Conversely, the equality of factor

loadings, response thresholds and residual variances implies that an item is psychometrically equivalent in the groups under study.

The procedure of testing for measurement non-invariance involves the specification of a hierarchical sequence of measurement models imposing increasingly strict equality constraints on model parameters across groups. More restrictive models are then compared with less restrictive counterparts in terms of fit to the data. We examined measurement non-invariance across three levels: (1) Configural invariance requires that the manifest variables measure the same latent trait in each of the groups, but does not require equality of factor loadings or response thresholds. For purposes of model identification, the factor means are fixed to zero and item residual variances fixed to one (Muthén and Muthén, 1998-2010); (2) Strong factorial invariance requires that the manifest variables measure the same latent trait in each of the groups, as well as equality of factor loadings and item response thresholds. For model identification, the factor mean is fixed to zero and residual variances are fixed to one in the reference group, while means and residual variances are allowed to vary in other groups (Muthén and Muthén, 1998-2010); (3) Strict factorial invariance poses an additional constraint on the residual variances (by fixing them to one across groups), while fixing the latent mean to zero in the reference group and allowing the latent means to vary in the other groups. Conditions of partial measurement non-invariance can exist where not all items meet a given set of invariance assumptions.

Our motivation to identify the strictest possible condition of measurement invariance (including conditions of partial invariance) was to minimize risk of bias in the comparison of latent means. Although there is a view that strict factorial invariance may not be necessary for the comparison of latent factor means, the debate around this issue is ongoing. There are conditions under which models assuming strong factorial invariance can produce biased latent mean estimates (Deshon, 2004, Lubke and Dolan, 2003, Wu *et al.*, 2007) and it has been recommended that researchers use their discretion on whether testing for strict factorial is necessary (Vandenberg and Lance, 2000).

Factor structure, goodness-of-fit and difference testing

To investigate the latent dimensionality of the PSQ, we used Stata 12.1/SE (StataCorp, 2011) to randomly divide our study population ($n=12,011$) into two subsets (Set 1: $n=6,040$ / Set 2: $n=5,971$) and checked that these were broadly similar in terms of ethnic composition. Using Mplus 6.12 (Muthén and Muthén, 1998-2010), we explored the latent structure of the PSQ in our study population using the Set 1 data, and then tested this latent structure with a confirmatory factor analysis (CFA) of the Set 2 data. The fit of these models was assessed with the chi-square test of overall model fit (good fit indicated by $0 \leq \chi^2 \leq 2df$) (Byrne, 1991); acceptable fit by $2df \leq \chi^2 \leq 3df$ (Carmines and McIver, 1981)); the root mean square error of approximation (good fit indicated by $RMSEA < 0.06$) (Hu and Bentler, 1999); and the comparative fit index (good fit indicated by $CFI \geq 0.95$; acceptable fit by $CFI \geq 0.90$) (Bentler, 1990, Hu and Bentler, 1999). To determine whether the identified factor structure was appropriate for the groups under study, we examined the fit of group-specific CFAs using the same fit statistics (Supplement 2).

We then performed a sequence of multiple-group CFAs (Jöreskog, 1971) (Mplus syntax provided in Supplement 3). First, we assessed the fit of a measurement model assuming configural invariance. A model assuming strong factorial invariance was then compared with the configural invariance model with a chi-square difference test to assess whether the imposed equality constraints on factor loadings and response thresholds had resulted in worse fit to the data. If strong factorial invariance was rejected (at $p < 0.05$), we identified the item with the largest modification indices for factor loadings and response thresholds across groups, released the equality constraints for these parameters, and compared the fit of a model assuming partial strong invariance with the configural invariance model. We repeated this procedure until a model with acceptable fit was identified. We then proceeded to test for strict factorial invariance against a model assuming (partial) strong invariance, again using a chi-square difference test to assess whether the additional constraints on residual variances had resulted in worse fit to the data. If strict factorial invariance was rejected in favour of (partial) strong invariance, we identified the item with the largest modification index for residual variances across groups, released the equality constraint on the residual variance parameter, and compared the fit of a model assuming partial strict invariance with the (partial) strong invariance model. Again, this procedure was repeated until a model with acceptable fit was identified.

147 Probability weighting

148 Both the FNS and EMPIRIC studies intentionally oversampled ethnic minority respondents to ensure
149 sufficient numbers for comparison. Using UK Census data, we scaled the FNS and EMPIRC probability weights
150 so that the ethnic composition of our study population was consistent with that of the population of England
151 and Wales in 1991 (for FNS respondents) and with the population in England in 2001 (for EMPIRIC
152 respondents). Details of our method are provided in an online supplement (Supplement 4).

153 Sensitivity analyses

154 To assess the robustness of our findings against our choice of psychometric approach, we also examined the
155 PSQ for measurement non-invariance using Multiple Indicators Multiple Causes (MIMIC) models
156 (Supplement 5). Rather than stratifying the study population by ethnicity and assessing the fit of a model
157 simultaneously in multiple groups, MIMIC models examine the influence of ethnic group membership on the
158 mean level of latent trait, which in turn is determined by response on PSQ items. The assessment of
159 measurement non-invariance is then performed by testing for the presence of direct effects of ethnicity on
160 PSQ items, holding constant group differences in latent factor means. Conceptually, this can be thought of
161 as an assessment of whether certain groups are more (or less) likely to endorse a given PSQ item, irrespective
162 of their tendency to report psychotic symptoms per se. Within this framework, significant direct effects of
163 ethnicity on observed PSQ items are interpreted as evidence for measurement non-invariance.

164 Results

165 Descriptive statistics

166 Compared with British whites (56% female), females were under-represented in Indian (51%), Pakistani
167 (51%) and Bangladeshi groups (49%). Furthermore, respondents from Caribbean (mean age=39.7 years [95%
168 CI=38.9–40.6]), Indian (39.1 [38.3–39.9]), Pakistani (34.6 [33.8–35.4]) and Bangladeshi groups (34.6

[33.7–35.6]) were on average younger than White British (44.7 [44.1–45.4]) and White Irish counterparts (43.5 [41.8–45.2]).

Ethnic differences in prevalence of self-reported symptoms

Approximately 5% of our study population reported psychotic experiences (Table 2). Prevalence was greatest among Caribbean individuals of whom 13% reported at least one psychotic symptom, compared with 5% of British Whites (age and sex adjusted odds ratio [aOR]=2.38 [1.84–3.07]). Prevalence in terms of the individual symptoms was also greatest among Caribbean individuals who, most notably, were more likely than British Whites to report paranoid symptoms (aOR=3.43 [2.01–5.86]). Prevalence was also increased among Pakistani individuals, of whom 8% reported psychotic symptoms (aOR=1.36 [1.01–1.84]). This difference, however, was largely driven by an excess in reporting paranoid symptoms (aOR=3.13 [1.77–5.55]) and was not consistently observed in relation to other PSQ items. Risk of reporting psychotic symptoms was moderately increased in the Indian (7%, aOR=1.21 [0.89–1.65]) and White Irish groups (7%, aOR=1.39 [0.92–2.12]) compared with British Whites. The overall risk of reporting psychotic symptoms was moderately lower in the Bangladeshi group than among British Whites (5%, aOR=0.70 [0.48–1.03]), despite being more likely to report paranoid symptoms (aOR=1.69 [0.87–3.25]).

Factor structure of the PSQ instrument

In an exploratory factor analysis (EFA) of the Set 1 data, we examined the scree plot and eigenvalues (Supplement 6) as well as fit statistics and parameter estimates for a 1-factor solution (Supplement 7). While the 1-factor solution provided good fit to the data (χ^2 [df, p] = 5.38 [5, 0.37]; RMSEA = 0.004; CFI = 0.999), the mania item was only weakly associated with the latent factor (rotated factor loading [SE] = 0.09 [0.11]). Furthermore, Mplus statistical output for these analyses contained a warning that the cross-tabulations of the mania item with other PSQ items contained zero cells. As this implies a tetrachoric correlation of one, the mania item was therefore rendered unsuitable for use in multiple-group analyses based on tetrachoric correlation matrices, such as those reported in our study (Muthen and Muthen).

We then conducted a CFA of the Set 2 data, comparing a 1-factor model which included the mania item with a 1-factor model which did not (Supplement 8). Again, the mania item was only weakly associated with the latent factor (standardised factor loading [SE] = 0.28 [0.14]), and when we excluded the mania item the fit of the model improved (CFA model *including* mania item: χ^2 [df, p] = 15.64 [5, 0.008], RMSEA = 0.019, CFI = 0.937; CFA model *excluding* mania item: χ^2 [df, p] = 2.41 [2, 0.299], RMSEA = 0.006, CFI = 0.997). We evaluated the fit of a 1-factor solution in each of the groups (Supplement 2) finding that the 1-factor model without the mania item yielded good fit in the White British, White Irish, Caribbean, Pakistani groups, and adequate fit in the Indian and Bangladeshi groups. However, we identified an additional zero cell in the cross-tabulation of the thought interference and hallucination items in analysis of the Bangladeshi group. Given the choice to exclude an additional PSQ item from further analysis or the Bangladeshi group in its entirety, we chose the latter as we felt this would least impact on our inference from these data.

Psychometric comparison

We then performed a sequence of multiple-group CFAs of four PSQ items (though interference, paranoia, strange experience and hallucination) across five groups (White British, White Irish, Caribbean, Indian and Pakistani) (Table 3). All multiple-group models provided good fit to the data. The chi-square test of overall model fit was consistently within the 0 to 2df range, RMSEA statistics were below 0.06 and CFI statistics were above 0.95. Difference test statistics suggested a model assuming strong factorial invariance did not result in worse fit compared with the configural invariance model (model 1.2 difference test p-value=0.25). Our results therefore suggested that the item factor loadings and response thresholds were invariant across groups. However, placing further constraints on residual variances, the fit of the model deteriorated (model 1.3 difference test p-value=0.002). Examining modification indices for residual variances across groups, we identified the item measuring paranoid symptoms to be the largest source of measurement non-invariance. When we released the constraints on the residual variances for this item, a model assuming partial strict invariance provided acceptable fit to the data compared with the strong factorial invariance model (model 1.4 difference test p-value=0.18). Compared with British Whites, the residual variance of the paranoia item

was larger in the Pakistani group. We note that an investigation of our data with a MIMIC model produced a consistent result (see Supplement 5).

Group comparisons based on composite screening measures and latent means

We examined the potential impact of measurement non-invariance by comparing group differences based on PSQ composite screening response with group differences based on latent means obtained from our measurement models (Table 3). Furthermore, we compared latent mean differences based on a model which (erroneously) assumed strict factorial invariance (i.e. model 1.3) with those based on a model in which the different residual variances for the paranoia item across groups had been accounted for (i.e. model 1.4). Based on observed PSQ screening response, Pakistani (Crude OR 1.60 [1.18 to 2.17]) and Caribbean groups (Crude OR 2.54 [1.95 to 3.30]) were at greater odds of reporting psychotic symptoms compared with British Whites. However, group comparisons in terms of latent means suggested only the Caribbean group was at greater risk of experiencing psychotic symptoms, and that latent differences between Pakistani and White British groups were exaggerated when measurement non-invariance in the paranoia item was not accounted for. These findings may therefore signal problems with the validity of self-reported paranoid symptoms in the context of cross-ethnic psychosis research. We note that consistent results were found when we investigated our data with a MIMIC model (Supplement 5).

Discussion

Our study examined ethnic differences in psychotic symptom prevalence and assessed the psychometric characteristics of the Psychosis Screening Questionnaire in a cross-ethnic context. There were four main findings. First, Caribbean and Pakistani individuals were more likely than British Whites to report psychotic symptoms. However, while this risk was apparent across all PSQ items in the Caribbean group, risk among Pakistani appeared to be driven mainly by a greater likelihood of reporting paranoid symptoms. Second, our exploratory and confirmatory factor analyses revealed that the mania item was weakly associated with the latent factor and was therefore psychometrically distinct from other PSQ items. However, the lack of

correlation between the mania item and other PSQ items may be unsurprising considering that the psychotic symptoms associated with mania are more likely to be congruent with elated mood and involve themes of grandeur rather than those of persecution or paranoia captured by the other PSQ items. Third, multiple-group CFAs suggested that the thought interference, strange experience, and hallucination items were measured in fully equivalent ways across the groups under study. The paranoia item, however, captured the latent factor with greater residual variance among Pakistani respondents than among British Whites. Fourth, notwithstanding evidence for measurement non-invariance, Caribbean individuals were at greater risk of reporting symptoms and had higher levels of latent trait. The higher prevalence of self-reported psychotic symptoms among Caribbean individuals in our study population is therefore unlikely to be an artefact of measurement (Morgan *et al.*, 2010).

Our study adds to a growing body of research investigating the psychometric characteristics of screening instruments for psychosis-like experiences in cross-cultural settings. We investigated the psychometric properties of the Psychosis Screening Questionnaire simultaneously across several ethnic groups in a large UK population-based sample. This allowed the direct comparison of factor model parameters for groups who self-identified as having White British, White Irish, Caribbean, Indian, or Pakistani backgrounds.

There were several limitations. First, despite the large size of our study population we were limited in the number of comparisons we were able to draw. Small group sizes and/or low symptom prevalence meant we were unable to test our models for Chinese and Bangladeshi groups, or to evaluate potential measurement non-invariance in the mania item. Second, given the low prevalence of psychotic symptoms in our study population, the normality assumption for the latent variables may have been violated. Simulation studies have shown that tetrachoric correlations may overestimate associations between the underlying latent continuous response variables when the normality assumption for the latent variables does not hold, suggesting that our results may have exaggerated the true association between these symptoms in the general population (Flora and Curran, 2004). Third, while noting that the clinical validity of the PSQ instrument has been evidenced for a subset of high-risk individuals within the FNSEM (Nazroo, 1997), our

data did not include diagnostic records and we were therefore unable to evaluate differences between ethnic groups in terms of the clinical relevance of PSQ items or latent factors. Fourth, we emphasise that our results are contingent on assuming the existence of a latent construct which caused the observed response on the PSQ in our study population. This assumption underlies all latent variable methods, including those used in psychometric assessment, and is untestable when objective markers for a disorder under investigation are not available, or do not yet exist. It has recently been argued that if one were to relinquish the latent variable assumption, it would become possible to view the symptoms as being constitutive of disorder, rather than consequential to it (Borsboom and Cramer, 2013, McNally, 2016). From this perspective, the paranoid symptoms reported by Pakistani respondents in our study population would have validity in their own right, rather than being merely a biased measure of an underlying latent construct. Further research is therefore needed to determine the clinical relevance of self-reported paranoid symptoms in the context of cross-ethnic psychosis research.

Our results are consistent with prior evidence in finding a higher prevalence of paranoid symptoms in ethnic minority groups compared with a majority population (Barrio *et al.*, 2003, Fabrega *et al.*, 1994, Kendler, 1982, Littlewood and Lipsedge, 1981, Suhail and Cochrane, 2002, Veling *et al.*, 2007, Whaley, 1998). While it is difficult to appraise our findings in a field where only few studies have been conducted to date, we propose that problems relating to the measurement of paranoid symptoms may arise specifically in the context of majority-minority relationships. Exposure to ethnic and racial inequalities may foster a “healthy” cultural mistrust (Whaley, 1998) which then becomes conflated with the expression of paranoid symptoms when one is screened for psychosis, thus introducing noise into measurement. Alternatively, it is possible that exposure to racial discrimination gives rise to true paranoid symptoms more readily than to other psychotic symptoms. In support of this, population-based evidence from the Netherlands suggests that exposure to discrimination at baseline is associated longitudinally with the first onset of delusional ideation, but not with the first onset of hallucinatory experience (Janssen *et al.*, 2003). This therefore provides an alternative explanation for the excess in paranoid symptoms among the non-white ethnic groups included in our study population, and potentially for their differential association with the latent factor when

compared with a White British reference population. For reasons discussed here, our findings warrant caution when assessing paranoid symptoms to proxy ethnic variation in psychosis risk within the context of major ethnic inequalities, although further research is needed investigate the generalisability of these findings to other populations and screening instruments for psychosis.

Conclusions

PSQ items capturing thought interference, strange experience and hallucination were measured in fully equivalent ways across the ethnic groups included in our psychometric assessment, and may therefore be used to proxy ethnic variation in psychosis risk in our study population. The role of the paranoia item was less clear-cut. From a psychometric point of view, it provided a less reliable measure of variation in latent trait between ethnic groups, which may signal problems with the validity of this item in a cross-ethnic context. Alternatively, it is possible that true paranoid symptoms are more prevalent than other psychotic symptoms in ethnic minority populations, which could explain their differential association with the latent trait under investigation in this study. Given our findings, we recommend that self-reported paranoid symptoms be investigated separately from other self-reported psychotic symptoms in future studies of ethnic variation in psychosis risk to appreciate more fully the social context in which these symptoms are reported.

309 **Table 1. The Psychosis Screening Questionnaire**

Symptom screened for	Type of question	Question content	Positive screening indicated by
Mania	Root	"Over the past year, have there been times when you felt very happy indeed without a break for days on end?"	"Yes"
	Targeted ^a :	"Was there an obvious reason for this?"	"No"
	Targeted ^a :	"Did your relatives or friends think it was strange or complain about it?"	"Yes"
Thought interference	Root:	"Over the past year, have you ever felt that your thoughts were directly interfered with or controlled by some outside force or person?"	"Yes"
	Targeted ^a :	"Did this come about in a way that many people would find hard to believe, for instance, through telepathy?"	"Yes"
Paranoia	Root:	"Over the past year, have there been times when you felt that people were against you?"	"Yes"
	Targeted ^a :	"Have there been times when you felt that people were deliberately acting to harm you or your interests?"	"Yes"
	Targeted ^a :	"Have there been times when you felt that a group of people were plotting to cause you serious harm or injury?"	"Yes"
Strange experience	Root:	"Over the past year, have there been times when you felt that something strange was going on?"	"Yes"
	Targeted ^a :	"Did you feel it was so strange that other people would find it very hard to believe?"	"Yes"
Hallucination	Root:	"Over the past year, have there been times when you heard or saw things that other people couldn't?"	"Yes"
	Targeted ^a :	"Did you at any time hear voices saying quite a few words or sentences when there was no-one around that might account for it?"	"Yes"

310 Notes: (a) Targeted questions were only asked in case of positive response to the root question.

311 Table 2. Ethnic differences in psychotic symptoms prevalence

	Mania		Thought interference		Paranoia		Strange Experience		Hallucination		Composite screening measure ^h	
	% (n)	Age and gender adjusted OR (95% CI) ^a	% (n)	Age and gender adjusted OR (95% CI) ^a	% (n)	Age and gender adjusted OR (95% CI) ^a	% (n)	Age and gender adjusted OR (95% CI) ^a	% (n)	Age and gender adjusted OR (95% CI) ^a	% (n)	Age and gender adjusted OR (95% CI) ^a
Ethnic group												
White British ^b	0.7 (26)	1.00 (reference)	1.5 (43)	1.00 (reference)	0.9 (37)	1.00 (reference)	3.3 (107)	1.00 (reference)	1.1 (39)	1.00 (reference)	5.4 (186)	1.00 (reference)
White Irish ^c	1.5 (11)	2.13 (0.94-1.85)	1.7 (12)	1.15 (0.42-3.13)	1.3 (10)	1.38 (0.50-3.81)	4.5 (40)	1.38 (0.81-2.36)	1.6 (13)	1.50 (0.61-3.72)	7.4 (65)	1.39 (0.92-2.12)
Caribbean ^d	1.9 (32)	2.47 (1.32-4.63)	2.6 (55)	1.75 (1.05-2.92)	3.4 (70)	3.43 (2.01-5.86)	7.2 (138)	2.14 (1.53-2.99)	2.7 (47)	2.39 (1.37-4.15)	12.7 (236)	2.38 (1.84-3.07)
Indian ^e	1.1 (24)	1.50 (0.76-2.95)	1.6 (29)	1.03 (0.54-1.94)	1.5 (40)	1.39 (0.73-2.64)	4.3 (74)	1.21 (0.80-1.81)	0.6 (18)	0.55 (0.27-1.14)	7.0 (137)	1.21 (0.89-1.65)
Pakistani ^f	1.4 (21)	1.78 (0.85-3.72)	1.7 (25)	1.12 (0.58-2.16)	3.6 (55)	3.13 (1.77-5.55)	4.1 (61)	1.06 (0.70-1.59)	0.9 (14)	0.75 (0.33-1.71)	8.3 (126)	1.36 (1.01-1.84)
Bangladeshi ^g	0.4 (4)	0.46 (0.14-1.50)	1.1 (11)	0.73 (0.32-1.69)	2.0 (25)	1.69 (0.87-3.25)	1.8 (21)	0.44 (0.25-0.78)	0.5 (8)	0.46 (0.19-1.12)	4.5 (52)	0.70 (0.48-1.03)
Total	0.8 (118)		1.5 (175)		1.0 (237)		3.4 (441)		1.1 (139)		5.6 (802)	

312 Notes: (a) Odds ratio with 95% confidence interval. (b) Reference group, N=3,444 (c) N=848. (d) N=1,877. (e) N=2,538. (f) N=1,908. (g) N=1,240. (h) Positive screening on any PSQ symptom.

Table 3. Fit statistics and model results for multiple-group confirmatory factor analyses

Model fit and difference test statistics						
Model	Invariance assumption	Overall fit	RMSEA	CFI	Difference test	
					comparison	Fit statistics
1.1	Configural invariance ^{a, b}	$\chi^2=15.914$, df=10, p=0.10	0.017	0.995		
1.2	Strong factorial invariance ^{a, c}	$\chi^2=25.116$, df=18, p=0.12	0.014	0.994	1.2 with 1.1	$\chi^2=10.165$, df=8, p=0.25
1.3	Strict factorial invariance ^{a, d}	$\chi^2=57.758$, df=34, p=0.007	0.018	0.980	1.3 with 1.2	$\chi^2=37.521$, df=16, p=0.002
1.4	Partial strict factorial invariance (θ_{PSY3} freely estimated across groups) ^{a, e}	$\chi^2=40.411$, df=30, p=0.097	0.013	0.991	1.4 with 1.2	$\chi^2=16.170$, df=12, p=0.18

Invariant across groups:		
Factor loadings	Estimate (SE)	95% CI
Thought interference	1.00 (0.00)	N/A
Paranoid symptoms	0.92 (0.17)	0.59 to 1.25
Strange experience	1.29 (0.24)	0.82 to 1.76
Hallucination	0.84 (0.13)	0.59 to 1.09
Response thresholds	Estimate (SE)	95% CI
Thought interference	4.16 (0.49)	3.20 to 5.12
Paranoid symptoms	4.12 (0.63)	2.89 to 5.35
Strange experience	4.16 (0.57)	3.04 to 5.28
Hallucination	3.88 (0.43)	3.04 to 4.72
Residual variances	Estimate (SE)	95% CI
Thought interference	1.00 (0.00)	N/A
Strange experience	1.00 (0.00)	N/A
Hallucination	1.00 (0.00)	N/A

Non-invariant across groups: residual variance for paranoid symptoms		
Group	Estimate (SE)	95% CI
White British	1.00 (0.00)	N/A
White Irish	1.51 (0.56)	0.41 to 2.61
Caribbean	1.48 (0.34)	0.81 to 2.15
Indian	1.39 (0.34)	0.72 to 2.06
Pakistani	2.64 (0.66)	1.35 to 3.93

Group comparisons							
Group	Based on composite PSQ response ^f			Based on latent factor means			
				Assuming full factorial invariance across groups (model 1.3)		Assuming different residual variances for the paranoia item (model 1.4)	
	Crude OR	(95% CI)	p-value	η (SE)	p-value	η (SE)	p-value
White British	1.00			0.00 (ref)		0.00 (ref)	
White Irish	1.31	(0.83 to 2.07)	0.24	0.23 (-0.99 to 1.45)	0.71	0.12 (-1.06 to 1.30)	0.84
Caribbean	2.54	(1.95 to 3.30)	<0.001	1.06 (0.14 to 1.98)	0.023	0.93 (0.05 to 1.81)	0.040
Indian	1.29	(0.94 to 1.79)	0.27	0.60 (-0.36 to 1.56)	0.22	0.50 (-0.42 to 1.42)	0.29
Pakistani	1.60	(1.18 to 2.17)	0.003	0.86 (-0.05 to 1.76)	0.065	0.24 (-0.74 to 1.22)	0.64

Notes: (a) Factor loading for marker item thought interference fixed to 1 across groups. (b) Factor loadings and response thresholds allowed to vary, residual variances fixed to 1 across groups. (c) Factor loadings and response thresholds constrained to equality, residual variances fixed to 1 for British Whites but freely estimated in other groups. (d) Factor loadings and response thresholds constrained to equality, residual variances fixed to 1 across groups. (e) As model 1.3, but allowing the residual variances for paranoid symptoms to be freely estimated across groups. (f) To ensure that group differences in manifest and latent response patterns were based on the same selection of PSQ items, we excluded the mania item from the composite screening measure. (g) N=10,618.

Figure 1. Example of a continuous latent response model with three manifest variables and a single latent variable, tested simultaneously in two groups

Notes: Groups are denoted by subscripts [a] and [b]. Manifest variables are represented by $[x_1, x_2, x_3]$ in squares. Latent continuous response variables are represented by $[x_1^*, x_2^*, x_3^*]$ in circles. Latent trait variables are represented by $[\eta]$ in circles. The relationship of latent trait $[\eta_a]$ with latent response variable $[x_{1a}^*]$ is denoted by factor loading $[\lambda_{1a}]$. The level of latent response $[x_{1a}^*]$ at which positive response to manifest variable $[x_{1a}]$ is likely to occur is denoted by response threshold $[\tau_{1a}]$. Residual variance in observed measure $[x_{1a}]$ is denoted by $[\theta_{1a}]$.

References

- Adebimpe, V. R., Klein, H. E. & Fried, J.** (1981). Hallucinations and Delusions in Black Psychiatric Patients. *Journal of the National Medical Association* **73**, 517-520.
- Arnold, L. M., Keck, P. E., Jr., Collins, J., Wilson, R., Fleck, D. E., Corey, K. B., Amicone, J., Adebimpe, V. R. & Strakowski, S. M.** (2004). Ethnicity and first-rank symptoms in patients with psychosis. *Schizophrenia Research* **67**, 207-12.
- Barrio, C., Yamada, A. M., Atuel, H., Hough, R. L., Yee, S., Berthot, B. & Russo, P. A.** (2003). A tri-ethnic examination of symptom expression on the positive and negative syndrome scale in schizophrenia spectrum disorders. *Schizophrenia Research* **60**, 259-69.
- Bauer, S. M., Schanda, H., Karakula, H., Olajossy-Hilkesberger, L., Rudaleviciene, P., Okribelashvili, N., Chaudhry, H. R., Idemudia, S. E., Gscheider, S., Ritter, K. & Stompe, T.** (2011). Culture and the prevalence of hallucinations in schizophrenia. *Comprehensive Psychiatry* **52**, 319-25.
- Bebbington, P. & Nayani, T.** (1995). The Psychosis Screening Questionnaire. *International Journal of Methods in Psychiatric Research* **5**, 11-19.
- Bentler, P. M.** (1990). Comparative fit indexes in structural models. *Psychological Bulletin* **107**, 238-46.
- Borsboom, D. & Cramer, A. O.** (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology* **9**, 91-121.
- Byrne, B. M.** (1991). The Maslach Burnout Inventory: Validating Factorial Structure and Invariance Across Intermediate, Secondary, and University Educators. *Multivariate Behavioral Research* **26**, 583-605.
- Carmines, E. G. & McIver, J. P.** (1981). Analysing models with unobserved variables. In *Social measurement: Current issues* (ed. G. W. Bohrnstedt and E. F. Borgatta). Sage: Beverly Hills, CA.
- Chang, N., Newman, J., D'Antonio, E., McKelvey, J. & Serper, M.** (2011). Ethnicity and symptom expression in patients with acute schizophrenia. *Psychiatry Research* **185**, 453-455.
- Chu, C. C., Sallach, H. S., Zakeria, S. A. & Klein, H. E.** (1985). Differences in psychopathology between black and white schizophrenics. *International Journal of Social Psychiatry* **31**, 252-7.
- Cohen, C. & Marino, L.** (2013). Racial and Ethnic Differences in the Prevalence of Psychotic Symptoms in the General Population. *Psychiatric Services* **64**, 1103-1109.
- Deshon, R. P.** (2004). Measures are not invariant across groups with error variance homogeneity. *Psychology Science* **46**, 137-49.
- Erens, B., Primatesta, P. & Prior, G.** (2001). Health(1999) Survey for England: The Health of Minority Ethnic Groups.
- Fabrega, H., Jr., Mulsant, B. M., Rifai, A. H., Sweet, R. A., Pasternak, R., Ulrich, R. & Zubenko, G. S.** (1994). Ethnicity and psychopathology in an aging hospital-based population. A comparison of African-American and Anglo-European patients. *Journal of Nervous and Mental Disease* **182**, 136-44.
- Flora, D. B. & Curran, P. J.** (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* **9**, 466-91.
- Hu, L. t. & Bentler, P. M.** (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* **6**, 1-55.
- Janssen, I., Hanssen, M., Bak, M., Bijl, R. V., de Graaf, R., Vollebergh, W., McKenzie, K. & van Os, J.** (2003). Discrimination and delusional ideation. *The British Journal of Psychiatry* **182**, 71-6.
- Johns, L. C., Nazroo, J. Y., Bebbington, P. & Kuipers, E.** (2002). Occurrence of hallucinatory experiences in a community sample and ethnic variations. *The British Journal of Psychiatry* **180**, 174-178.
- Jöreskog, K. G.** (1971). Simultaneous factor analysis in several populations. *Psychometrika* **36**, 409-426.
- Kendler, K. S.** (1982). Demography of paranoid psychosis (delusional disorder): a review and comparison with schizophrenia and affective illness. *Archives of General Psychiatry* **39**, 890-902.
- King, M., Nazroo, J., Weich, S., McKenzie, K., Bhui, K., Karlsen, S., Stansfeld, S., Tyrer, P., Blanchard, M., Lloyd, K., McManus, S., Sproston, K. & Erens, B.** (2005). Psychotic symptoms in the general

population of England--a comparison of ethnic groups (The EMPIRIC study). *Social Psychiatry and Psychiatric Epidemiology* **40**, 375-81.

Kleinman, A. (1987). Anthropology and psychiatry. The role of culture in cross-cultural research on illness. *The British Journal of Psychiatry* **151**, 447-454.

Littlewood, R. & Lipsedge, M. (1981). Some social and phenomenological characteristics of psychotic immigrants. *Psychological Medicine* **11**, 289-302.

Lubke, G. H. & Dolan, C. V. (2003). Can Unequal Residual Variances Across Groups Mask Differences in Residual Means in the Common Factor Model? *Structural Equation Modeling: A Multidisciplinary Journal* **10**, 175-192.

Maslowski, J., Jansen van Rensburg, D. & Mthoko, N. (1998). A polydiagnostic approach to the differences in the symptoms of schizophrenia in different cultural and ethnic populations. *Acta Psychiatrica Scandinavica* **98**, 41-6.

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy* **86**, 95-104.

Millsap, R. E. & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research* **39**, 479-515.

Modood, T., Berthoud, R., Lakey, J. & et al. (1997). *Ethnic minorities in Britain: diversity and disadvantage*. Policy Studies Institute: London.

Morgan, C., Charalambides, M., Hutchinson, G. & Murray, R. M. (2010). Migration, Ethnicity, and Psychosis: Toward a Sociodevelopmental Model. *Schizophrenia Bulletin* **36**, 655-664.

Muthén, B. & Muthén, L. Mplus discussion forum.

Muthén, L. K. & Muthén, B. O. (1998-2010). *Mplus User's Guide*. Muthén & Muthén: Los Angeles, CA.

Nazroo, J. (1997). *Ethnicity and mental health: findings from a national community survey*. Policy Studies Institute: London.

Nuevo, R., Chatterji, S., Verdes, E., Naidoo, N., Arango, C. & Ayuso-Mateos, J. L. (2012). The Continuum of Psychotic Symptoms in the General Population: A Cross-national Study. *Schizophrenia Bulletin* **38**, 475-485.

Radhakrishnan, J., Mathew, K., Richard, J. & Verghese, A. (1983). Schneider's first rank symptoms--prevalence, diagnostic use and prognostic implications. *The British Journal of Psychiatry* **142**, 557-559.

Sproston, K. & Nazroo, J. (2002). *Ethnic Minority Psychiatric Illness Rates in the Community (EMPIRIC)*. TSO: London.

StataCorp (2011). *Stata Statistical Software: release 12*. StataCorp, LP: College Station, TX.

Suhail, K. & Cochrane, R. (2002). Effect of culture and environment on the phenomenology of delusions and hallucinations. *International Journal of Social Psychiatry* **48**, 126-38.

Thomas, P., Mathur, P., Gottesman, I. I., Nagpal, R., Nimgaonkar, V. L. & Deshpande, S. N. (2007). Correlates of hallucinations in schizophrenia: A cross-cultural evaluation. *Schizophrenia Research* **92**, 41-49.

UK Data Archive *Ethnic Minority Psychiatric Illness Rates in the Community (EMPIRIC): user guide for UK data archive*.

Vandenberg, R. J. & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods* **3**, 4-70.

Vanheusden, K., Mulder, C. L., van der Ende, J., Selten, J. P., van Lenthe, F. J., Verhulst, F. C. & Mackenbach, J. P. (2008). Associations between ethnicity and self-reported hallucinations in a population sample of young adults in The Netherlands. *Psychological Medicine* **38**, 1095-102.

Veling, W., Selten, J.-P., Mackenbach, J. P. & Hoek, H. W. (2007). Symptoms at first contact for psychotic disorder: Comparison between native Dutch and ethnic minorities. *Schizophrenia Research* **95**, 30-38.

Viswanath, B. & Chaturvedi, S. (2012). Cultural Aspects of Major Mental Disorders: A Critical Review from an Indian Perspective. *Indian Journal of Psychological Medicine* **34**, 306-312.

Weisman, A. G., Lopez, S. R., Ventura, J., Nuechterlein, K. H., Goldstein, M. J. & Hwang, S. (2000). A comparison of psychiatric symptoms between Anglo-Americans and Mexican-Americans with schizophrenia. *Schizophrenia Bulletin* **26**, 817-24.

- Whaley, A. L.** (1998). Cross-Cultural Perspective on Paranoia: A Focus on the Black American Experience. *Psychiatric Quarterly* **69**, 325-343.
- World Health Organisation** Key terms and definitions in mental health: Psychosis.
- World Health Organisation** (2010). International Statistical Classification of Diseases and Related Health Problems 10th revision.
- Wu, A. D., Li, Z. & Zumbo, B.** (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation* **12**.
- Yamada, A. M., Barrio, C., Morrison, S. W., Sewell, D. & Jeste, D. V.** (2006). Cross-ethnic evaluation of psychotic symptom content in hospitalized middle-aged and older adults. *General Hospital Psychiatry* **28**, 161-8.