
Article:
Statham, V., Emerson, L.-M. and Rowse, G. orcid.org/0000-0003-3292-4008 (2019) A systematic review of self-report measures of paranoia. Psychological Assessment, 31 (2). pp. 139-158. ISSN 1040-3590

https://doi.org/10.1037/pas0000645

Abstract

Paranoia can be conceptualised as consisting of a hierarchy of cognitions, ranging from commonly experienced thoughts about less severe perceived threats, up to less common, persecutory thoughts about extreme threats, which are associated with distressing psychosis. This review systematically appraises self-report paranoia questionnaires validated for use among the general population; the type of paranoia assessed, measurement or psychometric properties, and subsequent validation with clinical samples are all considered. A systematic literature search was performed using PubMed, Web of Science, and PsycInfo databases. Study methodologies and measurement properties were evaluated according to COnsenus-based Standards for the selection of health-based Measurement Instruments (Mokkink et al., 2012). Twenty-six studies, describing the validation of nine paranoia-related questionnaires, were identified. Questionnaires were reviewed in relation to the hierarchy of paranoia; with two questionnaires assessing 'low-level' paranoia, four assessing persecutory thoughts, and the remainder assessing paranoia across this continua. Questionnaires assessing the full hierarchy of paranoid thoughts, alongside associated dimensions such as pre-occupation, conviction, and distress, offer the most comprehensive assessment of paranoia in both non-clinical and clinical populations. Of the measures which do this, the Green et al. (2008) Paranoid Thoughts Scale had the strongest evidence for its measurement properties and is therefore recommended as the most reliable and valid self-report assessment of paranoia currently available. However, this review illustrated that generally paranoia questionnaires lack high quality evidence for their measurement properties. Implications of these findings for clinical practice and research are discussed.

Keywords: paranoia; self-report; questionnaires; validation; Psychometric; properties; measures; assessment

Public significance statement

This systematic review identified nine self-report questionnaires that have been developed to assess paranoia and were designed for use with the general population. An analysis of studies that used these questionnaires suggested that the Green et al. (2008) Paranoid Thoughts Scale has the best evidence for the reliability and validity of its test scores.

Introduction

Paranoia has traditionally been conceptualised as a symptom of psychosis-related diagnoses such as 'Schizophrenia' (American Psychiatric Association, 2013). However, increasing evidence shows that as with other clinically-relevant experiences (e.g. obsessive-intrusive thoughts; Berry & Laskey, 2012; voice hearing; Beavan, Read, & Cartwright, 2011), paranoid thoughts are also experienced by those without mental health difficulties. Freeman (2006) reviewed studies assessing different types of paranoid thoughts, in general population samples, over different time periods, and found varying estimates of the prevalence of paranoid cognitions, ranging from approximately 2% to 42%. Data from a large, nationally-representative UK sample has also found paranoid thinking to be associated with a variety of difficulties, including poorer physical and mental health (e.g. anxiety, worry, insomnia, suicidal ideation), reduced social functioning, lack of social support, and increased use of alcohol and cannabis (Freeman et al., 2011). While there may be debate as to whether these associated difficulties contribute towards the development of paranoia, or are a consequence of it, they highlight the potential gains of the study of this phenomenon, both to individuals affected by paranoid experiences, and to wider society.

What constitutes a 'paranoid' experience is often not well defined within the literature. Bentall et al. (2009) suggest that paranoia occurs as a result of a combination of cognitive biases (e.g. threat anticipation, jumping to conclusions, difficulty understanding others' mental states), and accompanying "emotion-related processes, such as anxiety, depression, and self-esteem" (p. 244). Paranoia can also be conceptualised as a more stable

personality trait (often described as "suspiciousness"), that can vary between individuals and within theories of schizotypy is associated with an increased vulnerability to develop psychotic symptoms (Johns & van Os, 2001). Accordingly, examples of extreme paranoid personality traits are associated with diagnoses such as 'Paranoid Personality Disorder (American Psychiatric Association, 2013). While paranoia may be associated with certain "suspicious" personality presentations, this review focuses in more detail upon the specific types of threat-based cognitions, thoughts and beliefs that could be described as 'paranoid'. Understanding the ideational experience of paranoia in isolation, rather than as part of a more stable personality structure, accounts for the fact that paranoid thoughts can fluctuate from moment to moment (Ben-Zeev, Ellington, Swendsen, & Granholm, 2010; Nittel et al., 2018) and can decrease in response to psychological interventions (Freeman & Garety, 2014).

What makes a cognition 'paranoid' is often not well defined, and thoughts can be described as 'paranoid' and 'persecutory' synonymously, when the meaning of these descriptors may differ (McKay, Langdon, & Coltheart, 2006). Academic research often focuses upon 'persecutory' thoughts, which are defined as explicit concerns about threats of current/ongoing harm to oneself, enacted by an intentional perpetrator (Freeman & Garety, 2000). However, paranoia can be conceptualised more broadly as including thoughts relating to an exaggerated and "persistent misconception of one-self as the target of another's thoughts or action" (Fenigstein & Vanable, 1992, p. 130), often described as 'ideas of reference'.

Persecutory thoughts and paranoid ideas of reference are distinguished somewhat from concerns about threats to society or wider social groups (e.g. broader conspiracy theories) by the focus of threat to oneself. However, there is likely to be overlap between these constructs. Ideas of reference more broadly could also be part of non-paranoid psychological difficulties, such as the self-focused attention seen among socially anxious

individuals (Clark & Wells, 1995), or those with "grandiose delusions" (e.g. relating to inflated sense of worth, or a special identity, Knowles, McCarthy-Jones, & Rowse, 2011). This perhaps explains why paranoia and ideas of reference emerged as distinct facets of positive schizotypy in Cicero and Kern's (2010) factor analytic study. What distinguishes paranoid ideas of reference from other self-referential thoughts may be how these thoughts are appraised, and whether they are associated with assumptions of ill will, hostility, or suspicious intent (Fenigstein & Vanable, 1992). Accordingly, questionnaires that only assess ideas of reference include items that may have positive appraisals and thus would not necessarily assess paranoia (e.g. thinking that people are waving at you - The Referential Thinking Scale, Lenzenweger, Bennet, & Lilenfield, 1996), as well as those more related to paranoia, which are likely to imply hostility,. In support of the importance of self-referential thoughts within the construct of paranoia, Stefanis et al. (2004) also demonstrated that ideas of reference load on to a paranoia factor, along with social anxiety and suspiciousness.

Freeman et al.'s (2005) hierarchy of paranoia provides a framework to organise these different paranoid cognitions, and includes thoughts that are less explicitly persecutory (e.g. ideas of reference), within a broad conceptualisation of paranoia. Freeman et al. (2005) order their identified paranoid cognitions in a hierarchy according to the severity of perceived threat, beginning with social evaluative concerns at the bottom (defined as interpersonal worries, such as fears of rejection/vulnerability/the world being dangerous), followed by ideas of reference, and finally persecutory thoughts relating to mild (e.g. people trying to cause irritation), moderate (e.g. people going out of their way to get at you), then severe (e.g. people trying to cause you significant harm) threats, at the top of the hierarchy. There are other theoretical models of paranoid cognition, which generally seek to explain persecutory beliefs in isolation and focus upon the origin and maintenance of these experiences (e.g. Bentall et al., 2009; Freeman, Garety, Kuipers, Fowler, & Bebbington, 2002). Furthermore,

as previously discussed, there are schizotypy theories which describe 'suspicious' personality traits (see Grant, Green, & Mason, 2018 for review) that may be associated with a greater incidence of paranoid cognitions. Thus, the paranoia hierarchy is the only widely cited model known to the authors that provides a structured account of the types of thought content can be said to be part of paranoid experience, and it is therefore used to structure this review.

Subsequent studies examining paranoid thoughts from Freeman et al.'s (2005) hierarchy have found that those from the lower hierarchy emerge as a distinct factor from persecutory thoughts (from the upper hierarchy) within factor analyses (Green et al., 2008; Ibáñez-Casas et al., 2015). Ideas of reference and social evaluative concerns are proposed as being the building blocks for the development of more explicit persecutory thoughts, and thus assessing both types of cognition alongside each other is argued to provide a more comprehensive understanding of paranoid ideation (Freeman et al., 2005). Relatedly, many studies find strong associations between paranoia and self-focused or self-conscious cognitive styles (Combs & Penn, 2004; Freeman et al., 2012; Smári, Stefánsson, & Thorgilsson, 1994).

Green et al. (2008) and Ibáñez-Casas et al. (2015) found that ideas of reference in social situations (social reference thoughts) were the most commonly endorsed paranoid thoughts among the general population, whereas persecutory ideas were the most commonly endorsed paranoid thoughts among clinical participants. Nevertheless, both types of thoughts were much more prevalent among individuals with persecutory delusions (PDs), suggesting that the entire hierarchy has clinical relevance to paranoia (Green et al., 2008).

There may be also factors other than paranoid thought content that influence whether these cognitions are clinically-relevant experiences. For example, paranoid thoughts that are more frequent, distressing, and appraised with more conviction and preoccupation are more common among clinical populations (Green et al., 2008; Ibáñez-Casas et al., 2015). Indeed, Peters, Joseph, Day, and Garety (2004) argue that the distress, conviction, and preoccupation

associated with persecutory cognitions determine how 'delusion-like' they are. Alternatively, Trower and Chadwick (1995) distinguish 'poor me' paranoia, where persecution is perceived as unjust or undeserved, and 'bad me' paranoia, where persecution is perceived as a deserved consequence of an individual's actions. Research suggests that 'poor me' paranoia is more common among those with psychosis-related diagnoses (Melo & Bentall, 2013; Melo, Concoran, Shryane, & Bentall, 2009).

Much of our understanding of paranoia among both clinical and non-clinical populations has been obtained using self-report questionnaires. Within these questionnaires paranoia is defined and assessed differently, which is may have influenced endorsement rates, and contributed towards the varying prevalence estimates for delusions and paranoid cognitions in the general population (Freeman, 2006). As persecutory thoughts are more common among clinical samples, and ideas of reference are more common among non-clinical samples as opposed to those experiencing psychosis (Green et al., 2008; Ibáñez-Casas et al., 2015), prevalence estimates are likely to be influenced by both the type of paranoid thought content from the hierarchy of perceived threat (Freeman et al., 2005) that is being assessed, and the population to which the questionnaire is administered.

Aside from within large symptom inventories (e.g. Minnesota Multiphasic Personality Inventory-2 Restructured Form; Ben-Porath & Tellegen, 2011), there are no paranoia specific self-report measures developed primarily with clinical samples. Rather, diagnostic interview tools tend to be preferred (e.g. Composite Diagnostic Interview; Kessler, Andrews, Mroczek, Ustun, & Wittchen, 1998). This preference stems from the historical use of diagnosis to classify distressing psychotic experiences and also arguments (summarised by Bell, Fiszdon, Richardson, Lysaker,& Bryson, 2007) that those experiencing psychosis may struggle to self-report accurately due to holding unusual beliefs, experiencing cognitive deficits, or desiring to minimise their experiences (e.g. due to stigma, as a defensive coping strategy). The

evidence for these arguments is mixed, and varies based upon the construct assessed (Baumstark et al., 2013; Kim et al., 2010; Selton, Wiersma, & van den Bosch, 2000). For PDs, Lincoln, Ziegler, Lüllmann, Müller, and Rief (2010) found a strong relationship between self-reported and observed-rated experiences, whereas Liraud, Droulout, Parrot, and Verdoux (2004) did not. However, a lack of association between self-reported and observer-rated paranoia does not necessarily indicate that an observers are more accurate than those self-reporting their experiences. Furthermore, self-report assessments have additional advantages such as their ability to be distributed widely, with fewer resources required, and potentially less impact of social desirability bias compared with a face-to-face interview.

The primary aim of this review is to critically evaluate existing self-report measures that were developed to assess paranoia with general population samples. However, it is also acknowledged that the inclusion of individuals with psychosis in the development and subsequent validation of these measures is important, to evidence the construct validity of scales and examine whether items are clinically-relevant. Indeed, not evidencing the clinical relevance of items has been a criticism when assessing other constructs in the general population, such as obsessive-intrusive thoughts (Berry & Laskey, 2012). Furthermore, questionnaires validated clinically and non-clinically have greater potential utility. Psychotic-like experiences that occur without significant distress or impairment increase the later risk of symptoms that may warrant a clinical diagnosis (Hanssen, Bak, Bijl, Vollebergh, & van Os, 2005; Welham et al., 2009). Thus, assessing paranoia across clinical and non-clinical populations could highlight variables that increase the likelihood of paranoia-related distress. Questionnaires validated for use with individuals experiencing distressing paranoia could also be used clinically within assessment or for outcome measurement within interventions.

In this review, we therefore aim to provide a critical appraisal of the measurement properties of self-report measures of paranoia that were developed using non-clinical

participants, or a mixed clinical and non-clinical group. Additionally, studies validating questionnaires with clinical populations, that were originally developed with non-clinical or mixed samples were included. The inclusion of these measures therefore encompasses the full continua of experience through non-clinical to clinical As well as evaluating the measurement properties of these questionnaires, we aim to use Freeman et al.'s (2005) proposed hierarchy of perceived threat, as a framework to categorise the construct of paranoia that is assessed by each measure. Thus, the reviewed questionnaires relate to conceptualisations of 'low level' paranoia, from the lower hierarchy (e.g. Paranoia Scale, Fenigstein & Vanable, 1992), 'persecutory beliefs', from the upper hierarchy (e.g. Persecutory Ideation Questionnaire, McKay et al., 2006) and paranoia constructs that span the entirety of the hierarchy (e.g. Green et al. Paranoid Thoughts Scale, Green et al., 2008).

Method

As this paper describes a literature review, no ethical approval was required for the research.

Search Strategy

The methods undertaken in this review were informed by guidelines in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement 2009 (Moher, Liberati, Tetzlaff, & Altman, 2009). A systematic search using PubMed, Web of Science, and PsycInfo databases was performed on January 4, 2017 (see Appendix A, supplementary materials). Synonyms of terms for the construct of interest (e.g. paranoia, suspiciousness, persecutory), population for questionnaire development (e.g. general population, non-clinical, community), instrument type (e.g. questionnaire, scale, inventory), and questionnaire properties (e.g. psychometric, reliability, validity), were used to search the titles, abstracts, and keywords of publications. Papers containing keywords for comorbid difficulties associated with paranoia (e.g. dementia, Parkinson's) were excluded.

Initial searching identified 2432 papers. Firstly duplicate papers were removed from the search results (n = 707), followed by articles which after abstract and title screening did not meet the inclusion criteria (n = 1667). The full text of remaining papers (n = 58) was screened, followed by an ancestry search of studies included after this stage. Database and ancestry searching was used to find papers pertaining to both the original development and subsequent psychometric validation of the identified questionnaires. A citation search was also performed for studies documenting the initial development of each measure. The screening and data extraction process was completed by the primary author.

Inclusion and Exclusion Criteria

The following inclusion criteria were applied: 1. Studies must describe a self-report questionnaire rather than observer/interview-based assessments; 2. Studies must describe the initial development of a questionnaire, or indicate within the abstract that the aim is to validate the measurement properties of the questionnaire; 3. Studies must assess measurement properties outlined by Terwee, de Vet, Prinsen, and Mokkink (2011), or complete item-response theory (IRT) analyses (Kean & Reilly, 2014), or latent class analyses (Dayton & Macready, 2006); 4. Included questionnaires must have a scale or subscale for the assessment of paranoia or persecutory delusions. Furthermore, all questionnaire subscales must assess constructs relating to paranoia or delusions. Questionnaires measuring a range of delusions must have a specific persecutory delusion subscale, and present psychometric data relating to this subscale specifically. Thus, in-keeping with the review's specific focus upon the assessment of paranoid ideation, paranoia subscales were not included if they were part of larger questionnaires either measuring non-delusional elements of psychosis (e.g. voice hearing, negative symptoms), other mental health difficulties (e.g. depression, personality disorder), or personality traits; 5. In line with Freeman and Garety (2000), the paranoia assessed must relate to fears of present/ongoing harm to the self (rather than to society, or

social groups); 6. Questionnaires must have been developed using at least a small proportion of participants who were recruited from the general population; 7. Studies evaluating the measurement properties of existing questionnaires (i.e. not the original development papers), may include samples from any population (e.g. clinical/non-clinical/mixed); 8. Studies must describe questionnaires developed to assess paranoia among adults (aged 18 +). However, articles describing questionnaires originally developed with adult populations, then applied to younger samples (aged 14 +), were included; 9. Articles must be published in peer-review journals; 10. Full articles must be available in English.

Articles were excluded if: 1. The questionnaires solely measured cognitive biases involved in paranoia, or reactions to/appraisals of paranoid experiences; 2. Questionnaires assessed paranoia solely in relation to another condition or difficulty - thus not assessing paranoia distinctly, but its overlap with other constructs. For example, dementia, Parkinson's disease, substance misuse, learning disability, paranoid personality disorder, or depression.

Quality Appraisal

The COnsenus-based Standards for the selection of health-based Measurement Instruments (COSMIN) protocol for systematic reviews of self-report questionnaires (Terwee et al., 2011) was followed to appraise the measurement properties of questionnaires. COSMIN definitions of measurement properties were therefore adopted, as were COSMIN standards for how to appraise these properties. The COSMIN appraisal tool was developed by systematically reviewing existing criteria for good measurement properties, following which a multi-disciplinary panel of experts reached a consensus upon which properties to include within the tool, and how their quality would be judged (Mokkink, Terwee, Patrick, et al., 2010). Thus, the COSMIN protocol was deemed a comprehensive, systematically developed framework, that was grounded in the knowledge of experts. Mokkink, Terwee, Gibbons et al. (2010) evaluated the inter-rater reliability of COSMIN appraisal ratings and found 80%

agreement between raters on at least two thirds of items. Adjustments to the tool and manual were made to address areas where reliability was weaker (Mokkink et al., 2012). The COSMIN tool has now been applied within numerous systematic reviews aiming to appraise the quality of questionnaires (e.g. Sutton et al., 2016; Wigham & McConachie, 2014).

To establish the quality of papers included in this review the quality of the methodologies used to assess measurement properties was firstly appraised. The COSMIN tool (Mokkink et al., 2012) appraises methodologies which assess different forms of reliability, namely internal consistency, test-retest reliability, and measurement error; as well as different types of validity, namely content/face validity, criterion validity, and construct validity (includes structural validity, testing of hypotheses about related/unrelated constructs, and cross-cultural validity). The responsiveness of measures and also IRT methodologies can also be evaluated. The appraisal items used to assess each measurement property are provided in the supplementary materials (Appendix B). However, items included how missing items were handled, study samples sizes, and whether the unidimensionality of scales was evidenced (e.g. for internal consistency). For each applicable appraisal item, studies were rated 'poor', 'fair', 'good', or 'excellent'. Following the recommended 'worst score counts' procedure, the lowest item rating was taken to represent the overall methodological quality of analyses establishing that measurement property (Terwee et al., 2012).

COSMIN definitions of measurement properties were followed. For example, while some studies claimed to evidence criterion validity by comparing clinical and non-clinical groups on their paranoia scores, COSMIN defines these analyses as assessing construct validity. If the methodology used to establish a measurement property was cited within a different paper, where possible this was obtained and consulted for the required information. The methodology for content validity was rated if a questionnaire was being validated for the first time, or with a new population (e.g. a new culture or clinical population).

Once the methodological quality of psychometric analyses had been appraised, the second stage was to appraise whether the psychometric findings themselves met the recommended standards (e.g. internal consistency: Cronbach's alpha ≥ .7). An accompaniment to the COSMIN methodological checklist was used, covering the same aspects of reliability, validity, and responsiveness (Terwee et al., 2011). Each measurement property was assessed positively, negatively, or indeterminately. The standards required for each measurement property, are provided in supplementary materials, Appendix C.

An overall rating for the strength of each measurement property, for each questionnaire, was created by combining the methodological quality appraisal score for a measurement property with ratings for the quality of the psychometric property itself. Evidence was rated as either positive (+) or negative (-), and the strength of evidence in either direction was rated according to the categories shown in Table 1.

Table 1
Ratings for the strength of evidence for each measurement property

Level	Rating	Criteria
Strong	**+++ or ---**	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	**++ or --**	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	**+ or -**	One study of fair methodological quality
Conflicting	**+/-**	Conflicting findings
Unknown	**?**	Only studies of poor methodological quality
Indeterminate	**I**	All included studies reported indeterminate findings

Note. + = positive evidence**,** and - = negative evidence. Indeterminate category created by the author. Adapted from COSMIN website: COSMIN.nl

A second, independent researcher (postgraduate trainee clinical psychologist) conducted the quality appraisal procedure for studies (n = 7) relating to three randomly selected paranoia questionnaires. Quality appraisal was similarly conducted by combining the appraisal of the study methodology with an appraisal of the psychometric findings reported to obtain an overall rating. Inter-rater reliability for the ratings of overall strength of evidence for measurement properties was good (Kvalseth, 1989), with a Cohen's Kappa = .80.

Disagreements were resolved through discussion and consultation with COSMIN recommendations (Terwee et al., 2012). Initial ratings were then adjusted if necessary.

Results

Twenty-six papers were identified which described the measurement properties of nine different paranoia-related questionnaires (Table 2): the Paranoia Scale (PS; Fenigstein & Vanable, 1992), Paranoia/Suspiciousness Questionnaire (PSQ; Rawlings & Freeman, 1996), Persecutory Ideation Questionnaire (PIQ; McKay et al., 2006), Persecution and Deservedness Scale (PaDS; Melo et al., 2009), Peters et al. Delusions Inventory (PDI; Peters, Joseph, & Garety, 1999), State Social Paranoia Scale (SSPS; Freeman et al., 2007), Paranoia Checklist (PC; Freeman et al., 2005), State Paranoia Checklist (SPC; Schlier, Moritz, & Lincoln, 2016), and the Green et al. Paranoid Thoughts Scale (GPTS; Green et al., 2008).

The remaining papers retrieved reported adapted versions of these measures or further validated their measurement properties: PS (Barreto Carvalho et al., 2014; Combs, Penn, & Fenigstein, 2002; Smári et al., 1994), PSQ (Huppert, Smith, & Apfeldrof, 2002), PIQ (Jones, Fernyhough, de-Wit, & Meins, 2008; Van Dongen, Buck, Koole, & Van Marle, 2011), PDI (Cella, Sisti, Rocchi, & Preti, 2011; Jones & Fernyhough, 2007; Jung et al., 2008; Lincoln et al., 2010; López-Ilundain, Pérez-Nievas, Otero, & Mata, 2006; Peters et al., 2004; Prochwicz & Gawęda, 2015; Rocchi et al., 2008; Verdoux et al., 1998), PC (Lincoln et al., 2010; Moritz, Van Quaquebeke, & Lincoln, 2012), and GPTS (Ibáñez-Casas et al., 2015). Lincoln et al. (2010) presented psychometric evaluations of both the PDI and the PC and findings were considered separately for each measure.

For all nine questionnaires, the lead author reviewed the content of items and the construct of paranoia that the authors of the measure claimed to assess. This allowed the examination of how the themes of questionnaire items related to Freeman et al.'s (2005) paranoia hierarchy and questionnaires were categorised based upon this model.

Table 2

Paranoia questionnaires identified and evidence reported for their measurement properties

Author & location	Year	Construct of paranoia/PDs	# items	Sample	Paranoia subscales	Measurement properties
Paranoia Scale						
Fenigstein & Vanable [1] United States of America (USA)	1992	'Normal', 'non-pathological' paranoia. Suspiciousness/assumptions of hostility reminiscent of clinical paranoia, occurring independent of psychiatric problems	20	Four different student samples, n ranged from 119 to 180		IC. All samples $\alpha \geq .81$ R. r = .70 (n = 180) HT. Associations with measures of trust ($rs \geq .30 \leq .32$), experience/inward expression of anger ($rs \geq .45 \leq .51$), outwardly expressed anger (r = .18*), belief in control of others (r = .34**) and need for personal control (r = .29**) SV. 1-factor structure explaining 25% of the variance (N = 581)
Smári et al. [2] Iceland	1994		20	N = 30 Patients with schizophrenia diagnoses		IC. α = .87 HT. Associations with a feeling of being watched (r = .27**) and scores on a clinician-rated measure of paranoia (r = .51**)
Combs et al. [3] USA	2002		20	n = 191 (non-Hispanic Whites) n = 102 (African-Americans) Students		IC. non-Hispanic Whites, α = .88, African-Americans, α = .79 HT. The two ethnic groups differed similarly on the PS and clinical measurements of paranoia: African-American students had significantly higher levels of paranoia (for all comparisons p < .005)
Barreto Carvalho et al. [4] Portugal	2014		20	N = 1218 Adolescent high school pupils aged 14 to 22	Mistrust thoughts (8-items), persecutory ideas (8-items), self-depreciation (3-items)	IC. α ≥ .72 for subscales SV. 3-factor structure explaining 46.6% of variance
Paranoia/Suspiciousness Questionnaire						
Rawlings & Freeman [5] Australia	1996	Paranoia/suspiciousness among the non-psychiatric population.	47	n = 264 (Sample 1) n = 297 (Sample 2) Students	Interpersonal suspiciousness/hostility (12-items), negative mood/withdrawal (7-items), anger/impulsiveness (9-items), mistrust/wariness (6-items), perceived hardship/resentment (7-items) Six-items had no subscale	IC. α = .87 (total scale, n = 297), α ranged between 64 to .89 for subscales (N = 561) SV. 5-factor structure (N = 561) R. r = .82 (n = 74)

Author & location	Year	Construct of paranoia/PDs	# items	Sample	Paranoia subscales	Measurement properties
Paranoia/Suspiciousness Questionnaire						
Rawlings & Freeman [5] Australia	1996	Paranoia/suspiciousness among the non-psychiatric population.	47	n = 264 (Sample 1) n = 297 (Sample 2) Students	Interpersonal suspiciousness/hostility (12-items), negative mood/withdrawal (7-items), anger/impulsiveness (9-items), mistrust/wariness (6-items), perceived hardship/resentment (7-items) Six-items had no subscale	IC. α = .87 (total scale, n = 297), α ranged between 64 to .89 for subscales (N = 561) SV. 5-factor structure (N = 561) R. r = .82 (n = 74)
Huppert et al. [6] USA	2002			n = 33 (patients with schizophrenia-related diagnoses) n = 46 (patients with anxiety/depression)		IC. Total scale $\alpha \geq .85$ for both samples R. r = .67 (n = 23) HT. Positive, statistically significant ($p < .05$) correlations with scores on 9 different self-report measures of anxiety and depression: $rs \geq .32 \leq .73$
Persecutory Ideation Questionnaire						
McKay et al. [7] Australia	2006	'Persecutory' ideation	10	n = 98 (students) n = 25 (patients with experience of PDs)		IC. α = .87 (students) and .90 (patients) HT. Positively correlated with PSQ scores of students (r = .85***) and clinical participants (r = .85***) Correlation with observer-rated PDs among clinical participants (r = .61***). Insignificant correlation between PSQ scores and observed-rated PDs when PIQ scores were partialed out (r = -.14), versus significant correlations between observed-rated PDs and PIQ scores with PSQ scores partialed out (r = .51*)
Jones et al. [8] United Kingdom (UK)	2008		Reduce from 10 to 7-items	n = 183 (PIQ e-questionnaire) n = 188 (paper-version of PIQ) Students		IC. $\alpha \geq .84$ for PIQ-7 and PIQ-10 (paper and online versions) SV. 1-factor structure, excluding three items from original measure, demonstrated with both samples
Van Dongen et al [9] Holland	2011		10	n = 269 (community sample) n = 88 (individuals with schizophrenia-related diagnoses)		IC. α = .78 (community sample) and .89 (clinical sample) R. ICC = .82 (n = 38, community participants) HT. Positively correlated with self-reported positive psychotic symptoms (r = .51***), but removing persecutory items from the comparison measure hardly affected this correlation (minimal divergence), r = .51***. Significantly higher PIQ scores among clinical

						participants (U = 256.00**)
Author & location	**Year**	**Construct of paranoia/PDs**	**# items**	**Sample**	**Paranoia subscales**	**Measurement properties**
Persecution and Deservedness Scale						
Melo et al. [10] UK/Portugal	2009	Persecutory beliefs and the perceived 'deservedness' of persecution.	10	n = 318 (British students) n = 290 (Portuguese students) n = 45 (patients with PDs)	Persecution beliefs and deservedness beliefs relating to the same 10-items	Analyses using combined British/Portuguese sample: IC. α = .84 (Persecution). For deservedness calculated an ICC = .38. SV. 1-factor structure explaining 42% of the variance (Persecution subscale). 1-factor structure (deservedness subscale) HT. Persecution scores correlated strongly with PS scores (r_s = .78***) and self-reported depression (r_s = .57***). Deservedness scores correlated moderately with PS scores (r_s = .28***) and self-reported depression (r_s = .35***). Significantly higher persecution scores for patients as opposed to students ($p < .001$). CCV. 'Substantially identical' factor structures for British and Portuguese samples independently
Peters et al. Delusions						
Peters et al. [11] UK	1999	PDs in the general population. Attenuated versions of delusions	40	N = 272. (students and researcher acquaintances)	5-item subscale designed to assess PDs. However, factor analysis found three paranoia-related subscales: persecution (5-items), suspiciousness (3-items), and paranoid ideation (4-items) Items assessed on dimensions of conviction, pre-occupation and distress	SV. 11-factor structure explaining 59% of the variance
Verdoux et al. [12] France	1998		21	N = 444 (GP surgery attendees)	One PD-related subscale: "suspiciousness and persecutory ideas" (4-items)	SV. 7-factor structure explaining 55.3% of the variance
Peters et al. [13] UK	2004		21	N = 444 (university staff, students and research acquaintances)	Two items selected from each of the three PD-related, factor analytically identified subscales by Peter's et al. (1999)	SV. Select the two highest loading items from each factor identified by Peters et al. (1999) to create a shortened questionnaire

Autor & location	Year	Construct of paranoia/PDs	# items	Sample	Paranoia subscales	Measurement properties
Peters et al. Delusions Inventory						
Jung et al. [14] Korea	2004		40	N = 310 (community sample)	Initially identify "persecutory ideas" and "jealousy and suspiciousness" subscales - do not state number of items The authors later conclude that a unidimensional scale is more appropriate	SV. 10-factor structure explaining 57% of the variance. However, they argue that the dominant factor suggests a unidimensional structure (un-rotated explains 26% of variance)
Jones & Fernyhough [16] UK	2007		21	N = 493 (students)	Dispute the existence of previously established paranoia-related subscales	IC. Verdoux et al.'s (1998) suspiciousness and persecutory ideas subscale (α = .50) López-Ilundain et al.'s (2006) paranoid subscale (α = .26) SV. Lack of "valid multifactorial structure"
López-Illudain et al. [15] Spain	2006		21	N = 356 (community sample)	Factor analysis identified a "paranoid" subscale (2-items)	SV. 7-factor structure explaining 53.7% of the variance
Rocchi et al. [17] Italy	2008		21	n = 89 (outpatients with psychosis-related diagnoses n = 210 (community sample)	Refer to a "paranoia dimension" of the PDI (4-items)	For combined clinical/non-clinical sample: largest class found in latent class analysis (n = 140; 41.1%) related to a high probability of endorsing PDI items from the paranoia dimension
Lincoln et al. [18] Germany	2010		40	N = 80 (patients with psychosis-related diagnoses)	Peters et al. (1999) original 5-item PD scale	HT. Positively correlated with observer-rated PDs (r = .34***)
Cella et al. [19] UK & Italy	2011		21	n = 400 (British) n = 400 (Italian) Community samples		For combined British/Italian sample: latent class analysis identified a class (n = 330; 41.3%) associated with endorsement of two items with paranoid themes

Author & location	Year	Construct of paranoia/PDs	# items	Sample	Paranoia subscales	Measurement properties
Peters et al. Delusions Inventory						
Prochwicz & Gawęda [20] Poland	2015		40	N = 421 (community sample)	Initially identified subscales for 'suspiciousness', 'ideation of persecution and body distortion', and 'ideation of persecution' - number of items not reported The authors later argue for a unidimensional scale	SV. 14-factor structure explaining 58.68% of variance. However, scree plot suggests a unifactorial structure
State Social Paranoia Scale						
Freeman et al. [21] UK	2007	Assesses the expectation of harm from an intentional perpetrator in a recent situation	10	n = 100 (community sample) n = 64 (students) n = 21 (those at high risk of developing psychosis)		IC. $\alpha \geq .84$ for all samples R. ICC = .74 (n = 42) HT. Positively correlated with interviewer-rated paranoia (r = .73***), GPTS scores (r = .41***), visual analogue paranoia (r = .59***) and character hostility ratings (r = .63***) in the community sample, and PS scores in the student (r = .31*) and clinical (r = .44*) samples. Negatively correlated with perceptions of VR characters as positive (r = -.27***) or neutral (r = -.44***)
Paranoia Checklist						
Freeman et al. [22] UK	2005	Assesses paranoid thoughts of a "more clinical nature" than the PS	18	N = 1202 (students)	Items rated on dimensions of frequency, conviction and distress	IC. $\alpha \geq .90$ for all rating scales HT: Positively correlated with PS frequency (r = .71***), conviction (r = .62***), and distress (r = .58***) scores
Lincoln et al. [18] Germany	2010		18	N = 80 (patients with psychosis-related diagnoses)	Items rated on dimensions of frequency, conviction and distress	HT. Observer-rated PDs positively correlated with PS frequency (r = .43**), conviction (r = .39**), and distress (r = .38**) scores
Moritz et al. [23] Germany	2012		18	N = 1899 (community sample)	'Unspecified suspiciousness' (11-items) and 'psychotic paranoia' (5-items) 2 items had no subscale.	SV. 2-factor structure explaining 64% of the variance

Author & location		Construct of paranoia/PDs	# items	Sample	Paranoia subscales	Measurement properties
State Paranoia Checklist						
Schlier et al. [24] Germany	2006	State-adapted version of the PC assessing paranoia "in the moment", rather than as a trait	13, 5, and 3-item version	n = 1893 (community sample 1) n = 1966 (community sample 2)		Sample 1: SV. 1-factor structure for all versions RSP. Change effect size for 13-item, d = .17, 5-item, d = .19, and 3-item SPCs, d = .27 HT: All versions of the PC were correlated with trait measures of paranoia (rs ≥ .47 ≤ .55) and measures of social anxiety (rs ≥ .42 ≤ .46). Within a regression, PS frequency and distress scores were significantly predicted by anxiety, anger, depression and shame, but not significantly predicted by joy. Sample 2: IC: $\alpha \geq .74$ for all versions
Green et al. Paranoid Thoughts Scale						
Green et al. [25] UK	2008	Assesses a hierarchy of paranoid thoughts; from social reference thoughts to persecutory ideas.	32	n = 353 (university staff or students) n = 50 (individuals with PDs)	Persecution (16-items) and social reference (16-items) Items rated on dimensions of preoccupation, conviction, and distress	FA. On pool of 95 items. 2-factor structure explaining 49.7% of the variance (non-clinical sample). 16-tems per factor retained IC. α ≥ .90 for both samples on both subscales R. ICC ≥ .80 for all subscales (n = 164, non-clinical) HT. For both samples all GPTS scales were positively correlated with other measures of paranoia (rs ≥ .35 ≤ .86) anxiety (rs ≥ .34 ≤ .49) and depression (rs ≥ .42 ≤ .60) Significantly higher scores for clinical participants ($p < .001$) RSP. GPTS change scores correlated with change scores on interview-based paranoia measure (n = 30, clinical sample)
Ibáñez-Casas et al. [26] Spain	2015		32	n =151 (community sample) n = 40 (patients with delusions)	Persecution (16-items) and social reference (16-items)	IC: $\alpha \geq .90$ for both samples on all subscales SV: 2-factor structure explaining 61.7% of the variance (non-clinical sample) HT. Positively correlated with PDI (Smaller correlation with anxiety and depression measures. Higher scores for clinical group**: cut off of 92 gives 97.35% specificity and 65% sensitivity.

Note. *p < .05, **p < .01, ***p < .001, IC = Internal Consistency, R = Reliability, CV = Content Validity, SV = Structural Validity, HT = Hypothesis Testing, CCV = Cross-Cultural Validity, RSP = Responsiveness. ICC = intraclass correlation coefficient. For FA explained variance is included in the table where reported.

The PS (Fenigstein & Vanable, 1992) and PSQ (Rawlings & Freeman, 1996) measure commonly occurring paranoia among the general population, rather than so-called 'pathological' paranoia. These measures were deemed to best assess the lower levels of Freeman et al.'s (2005) paranoia hierarchy. Conversely, the PIQ (McKay et al., 2006), PaDS (Melo et al., 2009), PDI (Peters, Joseph, & Garety, 1999), and SSPS (Freeman et al., 2007) assess persecutory ideas, from the top levels of the paranoia hierarchy. The PaDS also assesses the perceived deservedness of persecution and the PDI assesses delusion-like qualities of persecutory ideas (conviction, pre-occupation, and distress). The PC (Freeman et al., 2005), SPC (Schlier, Moritz, & Lincoln, 2016), and the GPTS (Green et al., 2008), assess paranoia across the hierarchy, including ideas of reference and persecutory ideation.

Quality Analysis

Methodological quality ratings for each paper are shown in Table 3, along with ratings illustrating the overall strength of evidence for the measurement properties of each questionnaire. A full breakdown of the methodological ratings can be requested from the author.

Many studies did not describe how missing data were handled. As this can be a source of bias (Mokkink et al., 2012), such study methodologies were not rated better than 'fair'. Furthermore, the limited piloting of questionnaires meant that content validity and cross-cultural validity methodologies were rated 'poor', and no good psychometric evidence for these properties was reviewed. Methodologies for assessing structural validity and testing construct validity hypotheses were relative strengths for many studies, and accordingly these properties received stronger ratings. No studies assessed measurement error, criterion validity, or used IRT. Only two studies assessed the responsiveness of a questionnaire to measure change over time. Finally, no questionnaires included embedded validity indicators to assess the accuracy of the self-reported experiences.

Table 3

Quality ratings for study methodologies and ratings for overall evidence for the measurement properties

	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity	Responsiveness
Paranoia Scale						Icelandic, Portuguese & African-American samples	
Evidence for measurement property	++	-	+	--	++	**?**	
Methodological quality of studies							
Fenigstein & Vanable (1992)	Fair	Fair	Good	Fair	Poor		
Smári et al. (1994)	Poor		Poor		Fair	Poor*	
Combs et al. (2002)	Poor		Poor		Fair	Poor	
Barreto Carvalho et al. (2014)	Fair		Poor	Fair		Poor*	
PSQ							
Evidence for measurement property	-	+	**I**	**I**	**?**		
Methodological quality of studies							
Rawlings & Freeman (1996)	Fair	Fair	Fair	Fair			
Huppert et al. (2002)	Poor	Poor	Poor		Poor		
PIQ						Dutch sample	
Evidence for measurement property	+	+	**?**	**I**	+	**?**	
Methodological quality of studies							
McKay et al. (2006)	Poor		Poor		Fair		
Jones et al. (2008)	Fair			Fair			
Van Dongen et al. (2011)	Poor	Fair	Poor		Poor	Poor	

(continued)

	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity	Responsiveness
PaDS						Portuguese sample	
Evidence for measurement property	**P: ++** **D: ?**		**?**	**--**	**P: +** **D: -**	**?**	
Methodological quality of studies							
Melo et al. (2009)	P: Good D: Poor		Poor	Good	Fair	Poor	
PDI						French, Spanish, Korean, Italian, German, and Polish samples	
Evidence for measurement property	**-**		**?**	**+/-**	**-**	**?**	
Methodological quality of studies							
Peters et al. (1999)			Poor	Good			
Verdoux et al. (1998)			Poor	Fair		Poor	
Peters et al. (2004)			Poor				
Jung et al. (2008)			Poor	Fair		Poor	
López-Illundain et al. (2006)			Poor	Fair		Poor	
Jones & Fernyhough (2007)	Fair			Fair			
Rocchi et al. (2008)			CNR			Poor*	
Lincoln et al. (2010)			CNR		Fair	Poor*	
Cella et al. (2011)			CNR			Poor*	
Prochwicz & Gawęda (2015)			Poor	Fair		Poor	

(continued)

	Internal consistency	Reliability	Content validity	Structural validity	Hypothesis testing	Cross-cultural validity	Responsiveness
SSPS							
Evidence for measurement property	**?**	**+**	**?**		**+**		
Methodological quality of studies							
Freeman et al. (2007)	Poor	Fair	Poor		Fair		
PC						German sample	
Evidence for measurement property	**?**		**?**	**+**	**+/-**	**I**	
Methodological quality of studies							
Freeman et al. (2005)	Poor		Poor		Fair		
Lincoln et al. (2010)			CNR		Fair	Poor*	
Moritz et al. (2012)			CNR	Fair		Fair*	
SPC						German sample	
Evidence for measurement property	**+**		**CNR**	**I**	**+**	**I**	**?**
Methodological quality of studies							
Schlier et al. (2016)	Fair		CNR	Fair	Fair	Fair*	Poor
GPTS						Spanish sample	
Evidence for measurement property	**?**	**++**	**++**	**?**	**++**	**?**	**?**
Methodological quality of studies							
Green et al. (2008)	Poor	Good	Poor	Poor	Fair		Poor
Ibáñez-Casas et al. (2015)	Poor		Excellent	Poor	Fair	Poor	

Note. +++ or --- (strong positive or negative evidence), ++ or – (moderate positive or negative evidence), ‘+ or – (limited positive or negative evidence), +/- (conflicting findings), ? (only studies of poor methodological quality), or I (quality not possible to determine). P = persecution subscale. D = deservedness subscale. Although all papers were written in English, for some properties information needed to rate methodological quality was contained in another non-English language paper. In such cases either no items in an appraisal section could be rated (CNR) or ratings were based on a subset of items (*). Cross-cultural validity was rated for studies using measures in a different language or culture. Blank cells indicate where a measurement property was not examined within a paper.

Measures Assessing the Lower Paranoia Hierarchy

The PS (Fenigstein & Vanable, 1992) and the PSQ (Rawlings & Freeman, 1996) were designed to assess 'normal', 'non-pathological' paranoia, and the items in the questionnaire best reflect social evaluative concerns and ideas of reference within the lower levels of Freeman et al.'s (2005) hierarchy. Items from both measures also go beyond the hierarchy, assessing constructs related to paranoia, such as self-depreciation (PS; Barreto Carvalho et al. 2014) and anger/impulsiveness (PSQ; Rawlings & Freeman, 1996). Persecutory ideas from further up the paranoia hierarchy are not assessed; described as "obviously psychotic" (Fenigstein & Vanable, 1992, p. 131) and less relevant to the 'normal' population.

Rawlings and Freeman (1996) identified a 5-factor structure for the PSQ without stating the explained variance. The factor structure of the PS was also unclear, with Fenigstein and Vanable (1992) retaining a 1-factor structure, whereas with Portuguese adolescents, Barreto Carvalho et al. (2014) retained a 3-factor structure. These conflicting results could reflect methodological problems with the initial factor analysis of the measure, or limited validity of scores across age or these cultures. Evidence for the cross-cultural validity of the PS in Portuguese was poor, as studies did not conduct factor analyses to replicate the structural validity of scores on the measure and used samples dissimilar to the original development sample (Combs et al., 2002; Smári et al., 1994). The PS had evidence of good internal consistency of its test scores and some mixed findings with regards to construct validity (Barreto Carvalho et al., 2014; Combs et al., 2002; Fenigstein & Vanable, 1992; Smári et al., 1994). However, its test-retest reliability correlations were not adequate (Fenigstein & Vanable, 1992). Test-retest reliability was evidenced for scores on the PSQ. However, no other measurement properties were rated positively, which was often due to methodological limitations (Rawlings & Freeman, 1996).

The PS (Smári et al., 1994) and PSQ (Huppert et al., 2002) were validated with clinical participants, and the studies reported positive psychometric findings relating to internal consistency, reliability, and hypothesis testing. However, methodological problems meant that the quality of evidence for these areas was often rated poorly. Furthermore, there was little consideration of how appropriate these specifically 'non-clinical' assessments were for a clinical population, as arguably neither questionnaire could accurately assess the range of potential paranoid thoughts experienced by those diagnosed with schizophrenia.

Measures Assessing the Upper Paranoia Hierarchy

The PIQ (McKay et al., 2006), PaDS (Melo et al., 2009), PDI (Peters et al., 1999), and SSPS (Freeman et al., 2007) were designed with scales to measure persecutory beliefs, reflecting the upper levels of the paranoia hierarchy (Freeman et al., 2005). Freeman and Garety's (2000) definition of 'persecutory' was utilised in the development of items for the PIQ, PaDS, and SSPS. Alternatively, the PDI used a definition developed by experts (Peters et al., 1999). However, some PaDS items appear to only 'imply' persecutory ideas (Melo et al., 2009); it being questionable whether items such as "There are people that think of me as a bad person" specifically assess a perception of being at risk of harm. The PDI rates delusions on dimensions of conviction, pre-occupation and distress, whereas the PaDS also measures how deserved persecution is perceived to be (Trower & Chadwick, 1995). While the 'persecution' scale of the PaDS had some acceptable measurement properties, the properties of the 'deservedness' subscale are less evidenced, due to large amounts of missing data (Melo et al., 2009).

The PDI has items to assess PDs, alongside questions assessing other types of delusions (e.g. grandiose; Peters et al., 1999). Statements were worded to represent 'attenuated' versions of delusions, appropriate for general population samples. Although the 40-item PDI was designed with four PD items, Peters et al. (1999) identified three

components through factor analysis which relate to ‘paranoia’, covering a broader construct than just persecution (e.g. suspiciousness). However, Peters et al. argued that they had not aimed to “measure a limited number of well-defined subscales… but rather to sample as wide a variety of delusions as possible” (p. 562).

Six further studies reported PDI subscales relating to paranoia (Jung et al., 2008; López-Ilundain et al., 2006; Peters et al., 1999; Peters et al., 2004; Prochowitz & Gawęda, 2015; Verdoux et al., 1998), with a lack of consistency in the type and number of subscales identified. Furthermore, Jones and Fernyhough (2007) demonstrated the inadequate internal consistency of scores on previously identified paranoia subscales of the PDI, and reported a better fitting unidimensional factor structure, measuring general delusion-proneness. Similarly, while Jung et al. (2008) and Prochowitz and Gawęda (2015) initially extracted factors relating to persecution, they argued that the first underlying factor for the measure is highly dominant and suggested that a unidimensional factor structure is preferable. Finally, although latent class analyses using the PDI identified a ‘paranoid’ class of participants, the ‘paranoid’ items endorsed by participants were not consistent across samples (Cella et al., 2011; Rocchi et al., 2008).

Both the SSPS (Freeman et al., 2007) and 10-item PIQ (McKay et al., 2006; Van Dongen et al., 2011) have evidence of construct validity and test-retest reliability for their test scores. However, they were designed as unidimensional scales, without any assessment of structural validity (Freeman et al., 2007; McKay et al., 2006). Jones et al. (2008) did show that scores on a 7-item PIQ had good internal consistency and better fitted a unidimensional structure than the 10-item measure.

The SSPS assesses state persecutory ideation in the moment (as opposed to persecutory ideas over weeks/months) and was designed for studies where paranoia is assessed in a virtual reality (VR) environment (Freeman et al., 2007). However, there has

been no assessment of how responsive the scale is to momentary changes in paranoia, which is particularly important for a state measure.

Content validity and cross-cultural validity ratings were poor for all measures assessing the upper paranoia hierarchy, due to methodological limitations. For the PDI, factor structures were variable and the cross-cultural validity of scores from various European samples could not always be assessed as papers with the data needed to appraise these analyses were not available in English (German version; Lincoln, Keller, & Rief, 2009; Italian version; Preti, Marongiu, Petretto, Miotto, & Masala, 2002).

Most measurement properties for the persecutory measures were established with non-clinical populations. However, the PIQ was also validated with clinical participants (McKay et al., 2006). Construct validity hypotheses for PIQ, PaDS, and PDI were also supported, showing significant differences in scores between clinical and non-clinical samples (McKay et al., 2006; Melo et al., 2009), and correlations with observer-rated PDs (Lincoln et al., 2010; McKay et al., 2006).

Measures Assessing Paranoia Spanning the Full Hierarchy

Rather than focusing upon the lower or upper paranoia hierarchy, the PC (Freeman et al., 2005) and GPTS (Green et al., 2008) assess a range of paranoid thoughts at all levels. Freeman et al. (2005) did not establish an a priori construct for their measure, but based upon their findings argued that the PC assesses the hierarchy of paranoid thought, from social evaluative concerns up to persecutory beliefs. Green et al. (2008) later used this hierarchy to structure the GPTS item generation.

The PC assesses paranoia on dimensions of conviction and distress (Freeman et al., 2005), and the GPTS on dimensions of preoccupation, conviction, and distress (Green et al., 2008). Factor analyses showed that both measures have a 2-factor structure (‘persecution’ & ‘social reference’; Green et al., 2008; Ibáñez-Casas et al., 2015; ‘normal suspicions’ &

'pathological delusions'; Moritz et al., 2012). For the GPTS (Green et al., 2008), factors mapped on to the lower and higher ends of the paranoia hierarchy. However, methodologies were rated poorly for structural validity, internal consistency, and cross-cultural validity, due to sample size limitations (Green et al., 2008; Ibáñez-Casas et al., 2015). For the PC, some items from the 'pathological' factor did not reflect extreme persecutory beliefs from the paranoia hierarchy, and were instead described as 'clinically relevant' because they are bizarre and reflect 'first-rank' symptoms (Moritz et al., 2012; e.g. I detected coded messages about me in the press/TV/Radio).

The GPTS was designed for use with clinical and non-clinical participants, and validation studies involving both groups provided some moderate evidence for its measurement properties (e.g. reliability, hypothesis testing; Green et al., 2008; Ibáñez-Casas et al., 2015). The PC was subsequently applied with a clinical sample, where Lincoln et al. (2010) found a correlation between scores and observer-rated PDs. Lincoln et al. (2010) and Moritz et al. (2012) reported that the German version of the PC has good measurement properties. However, the cited papers were not available in English (Lincoln et al., 2009). Furthermore, although Freeman et al. (2005) reported good internal consistency for scores on the English PC, the unidimensionality of the scale is not evidenced, reducing the methodological quality.

The PC has also been developed in to a state measure of paranoia (SPC; Schlier et al., 2016); the 18 items were rephrased to ask how much they apply 'at the moment'. Schlier et al. (2016) generated 13-item, 5-item, and 3-item SPC scales, and demonstrated that the shorter scales (5-item and 3-item) were more responsive to momentary changes in paranoia. However, the data used were obtained from other studies with methodological limitations. Furthermore, COSMIN guidance cautions that higher effect sizes do not always necessarily indicate good responsiveness. The authors use effect sizes to demonstrate responsivity

without stating what the expected effect size for the interventions studied would be, making it difficult to judge their appropriateness. Further comparison of change scores with other measures would clarify that the SPC versions are appropriately responsive. All SPCs were undimensional scales (although no explained variance was reported) with good internal consistency. The 13 and 5-item measures were argued to encompass all levels of the paranoia hierarchy, with the 3-item version having reduced content validity, but still capturing key elements of persecutory thinking (Schier et al., 2016).

Discussion

This review aimed to critically evaluate existing self-report measures of paranoia, based upon the constructs of paranoia that they assess and their measurement properties. While the review identified measures developed in non-clinical populations, their applicability to clinical samples was also considered.

Nine questionnaires were identified, assessing paranoid beliefs relating to either the lower or upper levels of the paranoia hierarchy, or encompassing the full hierarchy (Freeman et al., 2005). A comprehensive conceptualisation of paranoia should include thoughts relating to varying degrees of threat and consider associated appraisals and distress. The PC (Freeman et al., 2005) and GPTS (Green et al., 2008) were the two measures fulfilling these criteria, capturing social reference paranoid thoughts commonly experienced across the population, as well as persecutory beliefs common among clinical samples, and endorsed by some of the general population. Between these measures, when combining the quality of the methodologies of analyses and the psychometric statistics reported, the GPTS has the most evidence for good measurement properties among clinical and non-clinical populations (Green et al., 2008; Ibáñez-Casas et al., 2015). It also is argued to have the most clearly defined construct underlying its items. This review therefore concludes that on the basis of current evidence, the GPTS (Green et al., 2008) offers the most valid and informative

assessment of paranoia. However, some psychometric findings (e.g. internal consistency, structural validity) require replication with a larger sample.

The PS (Fenigstein & Vanable, 1992) and PSQ (Rawlings & Freeman, 1996) were designed to measure 'subclinical' paranoia (analogous to the social evaluative concerns described by Freeman et al., 2005), as opposed to persecutory beliefs from the upper hierarchy. However, more recent research challenges the assumption that persecutory beliefs are always associated with psychosis, showing that they are also endorsed by some non-clinical participants (Green et al., 2008; McKay et al., 2006). By excluding supposedly 'extreme' paranoid thoughts, the PS and PSQ are unable to capture the range of paranoid experiences among a non-clinical sample, and are even less applicable for those with psychosis, who have more persecutory beliefs.

The PDI (Peters et al., 1999), PaDS (Melo et al., 2009), PIQ (McKay et al., 2006), and SSPS (Freeman et al., 2007) measure persecutory ideas evident in the upper paranoia hierarchy (Freeman et al., 2005). Researchers may assess persecutory beliefs in isolation, due to their clinical relevance. However, ideas of reference, which are not assessed by these questionnaires, may also be clinically-relevant if they cause distress and impairment. Some of the persecutory questionnaires do assess appraisals of beliefs, such as perceived deservedness (PaDS; Melo et al., 2009) and conviction, pre-occupation, and distress (PDI; Peters et al., 1999). The measurement properties of the PaDS deservedness scale (Melo et al., 2009), however, require further validation. Furthermore, the evidence reviewed suggested that the PDI should be used to assess general delusion proneness, rather than PDs specifically. Although only papers reporting paranoia-related subscales were included in this review, the use of the PDI to assess general delusion-proneness is also supported by other factor-analytic studies (Fonseca-Pedrero, Paino, Santarén-Rosell, Lemos-Giráldez, & Muñiz, 2012; Kim et al., 2013).

The PIQ (McKay et al., 2006) does not assess appraisals of persecutory ideas, but has more evidence for acceptable measurement properties with clinical and non-clinical populations. However, further factor analyses are required to establish whether a 10-item or 7-item measure is preferable. Given the increasing popularity of VR studies the SSPS (Freeman et al., 2007) is also a useful tool, but requires further evaluation of its responsiveness.

When measuring persecutory beliefs from the top of the hierarchy (Freeman et al., 2005), prevalence rates are likely to be lower in the general population (e.g. PIQ; McKay et al., 2006), whereas scores obtained using the PS (Feningstein & Vanable, 1992) and PSQ (Rawlings & Freeman, 1996) may be higher. However, total scores from the latter measures do not indicate the prevalence of paranoia specifically, as they include the assessment of associated experiences (e.g. anger/impulsivity). Measures such as the GPTS (Green et al., 2008) and PC (Freeman et al., 2005) therefore offer the best estimates of paranoia prevalence, capturing the full range of potential paranoid thoughts.

The limitations of the reviewed questionnaires have implications for studies that have used these measures. For example, by excluding the measurement of persecutory beliefs, studies using the PS and PSQ in clinical samples (e.g. Smári et al., 1994; Craig, Hatton, Craig, & Bentall, 2004) are unlikely to have measured a construct of paranoia appropriate for this population. Similarly, studies identifying PDI subscales that measure specific types of delusions, such as PDs (e.g. Jung et al., 2008), are using the measure in a way not intended by its original authors (Peters et al., 1999). Studies using the PDI to report the prevalence of PDs (e.g. Verdoux et al., 1998) may therefore not have assessed these experiences appropriately. Finally, studies using the SSPS in VR settings (e.g. Freeman et al., 2015) have only assessed persecutory thoughts, therefore not capturing potentially more commonly-occurring thoughts from the lower paranoia hierarchy (Freeman et al., 2005). The SPC (Schlier et al., 2016) is a

state measure assessing a broader range of paranoid experiences, but requires further psychometric validation (e.g. reliability, structural validity). As every measure reviewed lacked high quality evidence for particular measurement properties, this should be considered a limitation of all studies using self-report paranoia questionnaires.

Clinical Implications

When using paranoia questionnaires in practice, clinicians should consider that measures assessing the full paranoia hierarchy (GPTS; Green et al., 2008; PC; Freeman et al., 2005) will assess a greater range of service users' experiences. Thoughts from the upper section of the hierarchy may be experienced frequently, and thoughts from the lower hierarchy still have potential to cause distress. Relatedly, measures assessing distress (GPTS; Green et al., 2008; PC; Freeman et al., 2005) can highlight more troubling paranoid experiences and evaluate distress reduction during therapy, which may be a better outcome than reductions in thought frequency. The GPTS (Green et al., 2008) and PC (Freeman et al., 2005) also assess appraisals of paranoid thoughts and could be used to assess the outcomes of interventions which aim to target these (e.g. metacognitive therapy; Moritz & Woodward, 2007). Thus far, the GPTS (Green et al., 2008) has been used in randomised controlled trials to assess the impact of various psychological interventions upon paranoia (e.g. Freeman et al., 2017; Garety et al., 2017). However, the validity of findings from these studies could be enhanced if there was better evidence for particular psychometric properties of the measure, such as responsiveness to change.

The psychometric evidence for the reviewed measures suggests that when using self-report paranoia questionnaires in practice clinicians should be mindful of limitations in their validity. Of the measures assessing a range of paranoid thoughts, along with appraisals and distress, the GPTS (Green et al., 2008) has the most robust psychometric evidence obtained using clinical participants and is therefore the most recommended. The scope of this

questionnaire are makes it appropriate to assess paranoia among those with psychosis, and those at risk of developing it, who may have fewer persecutory thoughts and less distress.

If the GPTS is used clinically to track change over time then collecting further data about clinical and non-clinical norms on the scale could help to establish levels of clinically significant change, as has been done with measures such as the CORE-OM (Barkham et al., 2001). As there is an increasing emphasis upon the distinctiveness of different psychotic experiences (Bentall et al., 2014), there is likely to be value in further validating paranoia-specific questionnaire for use in interventions that specifically target these experiences.

Clinicians may wish to use measures other than the GPTS for specific purposes. If persecutory ideas specifically are an individual's primary difficulty, the PIQ (McKay et al., 2006) could be used, and is the persecutory measure most validated with clinical samples. Clinicians might also wish to assess the perceived deservedness of persecution, and could therefore use the PaDS (Melo et al., 2009). However, they should be aware of the limitations of using these measures, highlighted in this review.

Limitations

The questionnaires favoured in this review were based upon (GPTS; Green et al., 2008), or resulted in (PC; Freeman et al. 2005), the development of Freeman et al.'s (2005) paranoia hierarchy. The hierarchy is one conceptualisation of paranoia, and adopting an alternative definition may have influenced the conclusions of the review. However, Freeman et al.'s hierarchy is currently the most comprehensive model of paranoid cognition, with other research often failing to distinguish paranoid and persecutory beliefs (McKay et al., 2006). If there was a richer discussion within the academic literature about how thought content can be defined as 'paranoid', this would have perhaps enhanced the appraisal of the paranoia constructs within questionnaires. The Freeman et al. (2005) model itself could also benefit from a more detailed definition of the levels of the hierarchy, such as a more thorough

discussion of what constitutes a 'social evaluative concern' and how this relates to paranoia. Indeed, the lower parts of the hierarchy may be considered inappropriate in the assessment of paranoia due to the overlap with other difficulties (e.g. anxiety). However, the prevalence of social reference thoughts among paranoid samples (Green et al., 2008) and the strong association between persecutory thoughts and self-consciousness (Combs & Penn, 2004; Freeman et al., 2012; Smári et al., 1994), indicates the close relationship between the lower and upper paranoia hierarchy. Moreover, the authors' focus upon paranoia in the general population indicated a need for a broad and inclusive paranoia model.

The exclusion from the review of personality and psychotic symptom measures with paranoia subscales is a limitation, as researchers may wish to use these subscales in isolation to assess paranoia. However, it is argued that questionnaires which are focussed purely on paranoia and delusions are likely to have a more tightly-defined construct of paranoid ideation, whereas broader measures may lack this. Scales specifically designed to assess paranoia will also provide more psychometric data relevant to the assessment of paranoia, whereas broader measures may report properties of scales that include non-paranoid items. Within this review some measures did take items from other questionnaires assessing a range of constructs (e.g. the PSQ included items from large inventories of psychiatric symptoms and schizotypal personality). The inclusion of questionnaires that integrated items from schizotypy scales may also present a direct conflict with the aim of the review; to focus upon paranoid cognitions, as opposed to personality structures. However, this relates to a wider discussion about whether one can distinguish an item assessing a suspicious personality trait from an item assessing a paranoid thought.

Non-English papers were not accessed, limiting the ability to thoroughly evaluate the cross-cultural validity of some questionnaires (e.g. the Korean PaDS; Ko & Kim, 2016; and the Iranian GPTS; Abdolmohammadi, Mohammadzadeh, Ahmadi, & Ghadiri Sourman,

2016). Furthermore, unpublished papers were also not included in the review, which may have led to some publication bias.

The search strategy used within this review resulted in papers only being included if their study aims clearly referred to the validation of measurement properties of a questionnaire. This strategy is recommended for systematic reviews (Terwee et al., 2011), due to challenges identifying wider studies systematically, and to exclude studies without specific hypotheses about reliability or validity. However, it also meant that some psychometric data may have been missed if it was not part of the central aims of the study. Thus, despite some of the identified measures being widely cited (e.g. the original PS paper, Feningstein & Vanable, 1992, is cited over 500 times), often a very small proportion of these papers were included within the review. The lack of eligible studies could be a reflection of the small proportion of studies that provide subsequent validation of the measures, which may indicate the need for researchers to more routinely assess the measurement properties of paranoia questionnaires that they use.

The COSMIN protocol for systematic reviews was followed for the initial database search procedure (Terwee et al., 2011). However, the COSMIN protocol also indicates a second subsequent search, including the names of instruments found in the initial search, along with terms for measurement properties and the target population. While this second search was not completed, a citation search was instead performed and deemed satisfactory in achieving the same outcome. All published papers that cited the original development articles for questionnaires were thus included in this review.

Most studies included in the review were appraised poorly on particular COSMIN items (e.g. not reporting missing data, not piloting items), meaning that properties were rated ‘fair’ or ‘poor’, even if other criteria were met at a ‘good’ or ‘excellent’ level. This masked some of the variation between studies methodological quality. Those reviewing other self-

report measures have described COSMIN criteria as overly strict (Burton, Abbott, Modini, & Touyz, 2016). However, evaluating measurement properties in accordance with gold standard recommendations facilitated a higher quality systematic review and the methodological weaknesses identified illustrate areas of improvement for future research. Furthermore, while COSMIN provided a useful evaluative framework, the critique of the measures within this paper extends beyond this by reflecting upon the implications of psychometric findings, even if the methodologies used did not always meet the highly stringent COSMIN criteria.

Future Research

This review has highlighted a need for further validation of the existing paranoia measures. Future studies should do this using rigorous methodologies recommended by COSMIN (Mokkink et al., 2012), to ensure that the reviewed studies' limitations are not repeated. Studies could also employ IRT analyses to assess questionnaire properties. IRT could be used to examine whether, in line with Freeman et al.'s (2005) hierarchy of paranoia, clinical and non-clinical participants respond differently on items assessing different types of paranoid cognition (e.g. persecutory, social reference). With regards to the development of new paranoia questionnaires, authors should also pilot items with experts with professional and lived experience of paranoia, to ensure that the content reflects realistic paranoid experiences. Given arguments that it may be difficult for those experiencing paranoia to self-report these experiences (Bell et al., 2007), when developing paranoia questionnaires in the future it may be beneficial to consider including embedded validity indicators.

Obtaining more evidence for clinical and non-clinical norms on some of the most psychometrically valid paranoia measures such as the GPTS (Green et al., 2008) would increase their clinical applicability. Evaluating the measurement error of this tool would also enable estimates of reliable clinical change to be developed. There have been increasing efforts to design and evaluate interventions designed specifically to target paranoia or PDs

(e.g. Freeman et al., 2016). Increasing the quality of psychometric evidence for paranoia-specific measures could allow the tools to be used in research evaluating such interventions, and arguably lead to a more reliable and valid assessment of changes in paranoia.

To further investigate the value of assessing paranoid thoughts relating to varying degrees of threat, future research could assess thoughts from lower down the paranoia hierarchy in clinical populations. Studies could examine the distress associated with these thoughts and compare them with persecutory beliefs, higher in the hierarchy. Furthermore, building upon findings using observer-rated tools that non-distressing paranoid beliefs are predictive of later paranoia-related distress and psychosis (e.g. Hanssen et al., 2005; Welham et al., 2009), self-report questionnaires could be used longitudinally to examine the role of frequency, content, and appraisals made about paranoid thoughts in this process. For example, persecutory thoughts that are appraised as preoccupying and convincing may be associated with an increased risk of later psychosis.

As indicated within the limitations section of this review, future literature reviews could be conducted to include the paranoia subscales within broader measures of schizotypy, personality, or psychopathology. It would be of particular interest to compare how paranoia is conceptualised and assessed between these scales. For example, are items to assess a paranoid cognition actually distinct from items that measure a paranoid personality trait? Other reviews could also extend their scope to include measures that assess threats of harm to wider society (e.g. conspiracy theories). While this review focussed upon paranoid ideation, paranoid imagery is also prevalent among those with PDs (Schulze, Freeman, Green, & Kuipers, 2013), and considering how best to assess this is also of interest. Furthermore, although cognitive models conceptualise psychotic experiences as intrusions into awareness (Morrison, 2001), no measures have assessed the process characteristics of paranoid thoughts, and whether they do arise intrusively.

References

Abdolmohammadi, K., Mohammadzadeh, A., Ahmadi, E., & Ghadiri Sourman, F. (2016). Validation of the Paranoid Thoughts Scale in the Iranian population. Qom University of Medical Sciences Journal, 10, 47-53. doi:10.1080/13546800903378211

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., . . . McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE–OM: Toward practice-based evidence in the psychological therapies. Journal of Consulting and Clinical Psychology, 69, 184-196. doi: 10.1037/0022-006x.69.2.184

Barreto Carvalho, C., Pereira, V. M. D. P., da Motta, C., Pinto-Gouveia, J., Caldeira, S. N., Peixoto, E. B., & Fenigstein, A. (2014). Paranoia in the general population: A revised version of the General Paranoia Scale for adolescents. European Scientific Journal, 10, 128-141. doi:10.1111/cp.12065

Baumstarck, K., Boyer, L., Boucekine, M., Aghababian, V., Parola, N., Lançon, C., & Auquier, P. (2013). Self-reported quality of life measure is reliable and valid in adult patients suffering from schizophrenia with executive impairment. Schizophrenia Research, 147, 58–67. doi:10.1016/j.schres.2013.03.008

Beavan, V., Read, J., & Cartwright, C. (2011). The prevalence of voice-hearers in the general population: A literature review. Journal of Mental Health, 20, 281-292. doi:10.3109/09638237.2011.562262

Bell, M., Fiszdon, J., Richardson, R., Lysaker, P., & Bryson, G. (2007). Are self-reports valid for schizophrenia patients with poor insight? Relationship of unawareness of illness to psychological self-report instruments. Psychiatry Research, 151, 37–46. doi:10.1016/j.psychres.2006.04.012

Ben-Porath, Y. S., & Tellegen, A. (2011). MMPI-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form): Manual for administration, scoring, and interpretation. Minneapolis, MN: University of Minnesota Press

Bentall, R. P., de Sousa, P., Varese, F., Wickham, S., Sitko, K., Haarmans, M., & Read, J. (2014). From adversity to psychosis: Pathways and mechanisms from specific adversities to specific symptoms. Social Psychiatry and Psychiatric Epidemiology, 49, 1011-1022. doi:10.1007/s00127-014-0914-0

Bentall, R. P., Rowse, G., Shryane, N., Kinderman, P., Howard, R., Blackwood, N., … Corcoran, R. (2009). The Cognitive and Affective Structure of Paranoid Delusions. Archives of General Psychiatry, 66, 236-247. doi:10.1001/archgenpsychiatry.2009.1

Ben-Zeev, D., Ellington, K., Swendsen, J., & Granholm, E. (2010). Examining a cognitive model of persecutory ideation in the daily life of people with schizophrenia: A computerized experience sampling study. Schizophrenia Bulletin, 37, 1248–1256. doi:10.1093/schbul/sbq041

Berry, L. M., & Laskey, B. (2012). A review of obsessive intrusive thoughts in the general population. Journal of Obsessive-Compulsive and Related Disorders, 1, 125-132. doi:10.1016/j.jocrd.2012.02.002

Burton, A. L., Abbott, M. J., Modini, M., & Touyz, S. (2016). Psychometric evaluation of self-report measures of binge eating symptoms and related psychopathology: A systematic review of the literature. International Journal of Eating Disorders, 49, 123-140. doi:10.1002/eat.22453

Cella, M., Sisti, D., Rocchi, M. B., & Preti, A. (2011). Delusional profiles among young adults: A latent class analysis of delusion proneness. Psychiatry Research, 185, 97-101. doi:10.1016/j.psychres.2010.04.054

Cicero, D. C., & Kerns, J. G. (2010). Multidimensional factor structure of positive schizotypy. Journal of Personality Disorders, 24, 327–343. doi:10.1521/pedi.2010.24.3.327

Clark, D. M. & Wells, A. (1995). A cognitive model of social phobia. In R. Heimberg, M. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), Social phobia: Diagnosis, assessment and treatment (pp. 69–93). New York, NY: Guilford Press.

Combs, D. R., & Penn, D. L. (2004). The role of subclinical paranoia on social perception and behavior. Schizophrenia Research, 69, 93–104. doi:10.1016/s0920-9964(03)00051-3

Combs, D. R., Penn, D. L., & Fenigstein, A. (2002). Ethnic differences in subclinical paranoia: An expansion of norms of the Paranoia Scale. Cultural Diversity and Ethnic Minority Psychology, 8, 248. doi:10.1037//1099-9809.8.3.248

Craig, J. S., Hatton, C., Craig, F. B., & Bentall, R. P. (2004). Persecutory beliefs, attributions and theory of mind: Comparison of patients with paranoid delusions, Asperger's syndrome and healthy controls. Schizophrenia Research, 69, 29-33. doi:10.1016/s0920-9964(03)00154-3

Dayton, C. M., & Macready, G. B. (2006). Latent class analysis in psychometrics. In C. R. Rao & S. Sinharay (Eds.), Handbook of statistics 26 (pp. 421-446). Oxford, UK: Elsevier.

Fenigstein, A., & Vanable, P. A. (1992). Paranoia and self-consciousness. Journal of Personality and Social Psychology, 62, 129. doi:10.1037//0022-3514.62.1.129

Freeman, D. (2006). Delusions in the nonclinical population. Current Psychiatry Reports, 8, 191-204. doi:10.1007/s11920-006-0023-1

Freeman, D., Dunn, G., Fowler, D., Bebbington, P., Kuipers, E., Emsley, R., … Garety, P. (2012). Current paranoid thinking in patients with delusions: The presence of

cognitive-affective biases. Schizophrenia Bulletin, 39, 1281–1287. doi:10.1093/schbul/sbs145

Freeman, D., Emsley, R., Dunn, G., Fowler, D., Bebbington, P., Kuipers, E., … Garety, P. A. (2015). The stress of the street for patients with persecutory delusions: A test of the symptomatic and psychological effects of going outside into a busy urban area. Schizophrenia Bulletin, 41, 971-979. doi:10.1348/014466500163400

Freeman, D., & Garety, P. A. (2000). Comments on the content of persecutory delusions: Does the definition need clarification? British Journal of Clinical Psychology, 39, 07-414. doi:10.1348/014466500163400

Freeman, D., & Garety, P. A. (2014). Advances in understanding and treating persecutory delusions: A review. Social Psychiatry and Psychiatric Epidemiology, 49, 1179–1189. doi:10.1007/s00127-014-0928-7

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., … & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. The British Journal of Psychiatry, 186, 427-435. doi:10.1192/bjp.186.5.427

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. British Journal of Clinical Psychology, 41, 331-347. doi:10.1348/014466502760387461

Freeman, D., McManus, S., Brugha, T., Meltzer, H., Jenkins, R., & Bebbington, P. (2011). Concomitants of paranoia in the general population. Psychological Medicine, 41, 923-936. doi:10.1017/S0033291710001546

Freeman, D., Pugh, K., Green, C., Valmaggia, L., Dunn, G., & Garety, P. (2007). A measure of state persecutory ideation for experimental studies. The Journal of Nervous and Mental Disease, 195, 781-784. doi:10.1093/med:psych/9780199206315.003.0002

Freeman, D., Sheaves, B., Goodwin, G.M., Yu, L-M., Nickless, A., Harrison, P.J., . . . Espie, C.A. (2017). The effects of improving sleep on mental health (OASIS): A randomised controlled trial with mediation analysis. Lancet Psychiatry, 4, 749-758. doi:10.1016/S2215-0366(17)30328-0

Freeman, D., Waite, F., Emsley, R., Dunn, G., Kingdon, D., Davies, L., & Fitzpatrick, R. (2016). The efficacy of a new translational treatment for persecutory delusions: Study protocol for a randomised controlled trial (The Feeling Safe Study). Trials, 17, 1-8. doi:10.1186/s13063-016-1245-0

Fonseca-Pedrero, E., Paino, M., Santarén-Rosell, M., Lemos-Giráldez, S., & Muñiz, J. (2012). Psychometric properties of the Peters et al. Delusions Inventory 21 in college students. Comprehensive Psychiatry, 53, 893-899. doi:10.1016/j.comppsych.2012.01.007

Garety, P. A., Ward, T., Kuipers, E., Waller, H., Hardy, A., Freeman, D., . . . Bebbington, P. (2017). SlowMo, a digital therapy targeting reasoning in paranoia, versus treatment as usual in the treatment of people who fear harm from others: Study protocol for a randomised controlled trial. Trials, 18, 510-543. doi: 10.1186/s13063-017-2242-7

Green, C. E. L., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., & Garety, P. A. (2008). Measuring ideas of persecution and social reference: The Green et al. Paranoid Thought Scales (GPTS). Psychological Medicine, 38, 101-111. doi:10.1017/s0033291707001638

Grant, P., Green, M. J., & Mason, O. J. (2018). Models of schizotypy: The importance of conceptual clarity. Schizophrenia Bulletin, sby012. doi:10.1093/schbul/sby012

Hanssen M., Bak, M., Bijl R., Vollebergh, W., van Os, J. (2005). The incidence and outcome of subclinical psychotic experiences in the general population. British Journal of Clinical Psychology, 44, 181–191. doi:10.1348/014466505x29611

Huppert, J. D., Smith, T. E., & Apfeldorf, W. J. (2002). Use of self-report measures of anxiety and depression in outpatients with schizophrenia: Reliability and validity. Journal of Psychopathology and Behavioral Assessment, 24, 275-283. doi:10.1023/A:1020783101092

Ibáñez-Casas, I., Femia-Marzo, P., Padilla, J. L., Green, C. E., De Portugal, E., & Cervilla, J. A. (2015). Spanish adaptation of the Green Paranoid Thought Scales. Psicothema 27, 74-81. doi:10.7334/psicothema2014.103

Johns, L. C., & van Os, J. (2001). The continuity of psychotic experiences in the general population. Clinical Psychology Review, 21, 1125–1141. doi:10.1016/s0272-7358(01)00103-9

Jones, S. R., & Fernyhough, C. (2007). Reliability of factorial structure of the Peters et al. Delusions Inventory (PDI-21). Personality and Individual Differences, 43, 647-656. doi:10.1016/j.paid.2007.01.018

Jones, S. R., Fernyhough, C., De-Wit, L., & Meins, E. (2008). A message in the medium? Assessing the reliability of psychopathology e-questionnaires. Personality and Individual Differences, 44, 349-359. doi:10.1016/j.paid.2007.08.014

Jung, H. Y., Chang, J. S., Yi, J. S., Hwang, S., Shin, H. K., Kim, J. H., ... & Kim, Y. S. (2008). Measuring psychosis proneness in a nonclinical Korean population: Is the Peters et al. Delusions Inventory useful for assessing high-risk individuals? Comprehensive Psychiatry, 49, 202-210. doi:10.1016/j.comppsych.2007.08.011

Kean, J., & Reilly, J. (2014). Item response theory. In F. M. Hammond, J. F. Malec, T. G. Nick & R. M. Buschbacher (Eds.), Handbook for clinical research: Design, statistics and implementation (pp. 195-198). New York, NY: Demos Medical Publishing.

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H. U. (1998). The World Health Organization Composite International Diagnostic Interview short- form

(CIDI-SF). International Journal of Methods in Psychiatric Research, 7, 171-185. doi:10.1002/mpr.47

Kim, Y., Chang, J. S., Hwang, S., Yi, J. S., Cho, I. H., & Jung, H. Y. (2013). Psychometric properties of Peters et al. Delusions Inventory-21 in adolescence. Psychiatry Research, 207, 189-194. doi:10.1016/j.psychres.2012.09.002

Kim, S. H., Jung, H. Y., Hwang, S. S., Chang, J. S., Kim, Y., Ahn, Y. M., & Kim, Y. S. (2010). The usefulness of a self-report questionnaire measuring auditory verbal hallucinations. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 34, 968–973. doi:10.1016/j.pnpbp.2010.05.005

Knowles, R., McCarthy-Jones, S., & Rowse, G. (2011). Grandiose delusions: A review and theoretical integration of cognitive and affective perspectives. Clinical Psychology Review, 31, 684–696. doi:10.1016/j.cpr.2011.02.009

Ko, H., & Kim, B. (2016). Reliability and validity of the Korean-Persecution and Deservedness Scale. Korean Journal of Clinical Psychology, 35, 61-79. doi:10.15842/kjcp.2016.35.1.004

Kvalseth, T. O. (1989). Note on Cohen's kappa. Psychological Reports, 65, 223-26. doi:10.2466/pr0.1989.65.1.223

Lenzenweger, M. F., Bennett, M. E., & Lilenfeld, L. R. (1997). The Referential Thinking Scale as a measure of schizotypy: Scale development and initial construct validation. Psychological Assessment, 9, 452–463. doi:10.1037/1040-3590.9.4.452

Lincoln, T. M., Keller, E., & Rief, W. (2009). Die erfassung von wahn und halluzinationen in der normalbevölkerung: Deutsche adaptationen des Peters et al. Delusions Inventory (PDI) und der Launay Slade Hallucination Scale (LSHS-R). Diagnostica, 55, 29-40. doi:10.1026/0012-1924.55.1.29

Lincoln, T. M., Ziegler, M., Lüllmann, E., Müller, M. J., & Rief, W. (2010). Can delusions be

self-assessed? Concordance between self-and observer-rated delusions in schizophrenia. Psychiatry Research, 178, 249-254. doi:10.1016/j.psychres.2009.04.019

Liraud, F., Droulout, T., Parrot, M., & Verdoux, H. (2004). Agreement between self-rated and clinically assessed symptoms in subjects with psychosis. The Journal of Nervous and Mental Disease, 192, 352–356. doi:10.1097/01.nmd.000126702.30745.1d

López-Ilundain, J. M., Pérez-Nievas, E., Otero, M., & Mata, I. (2006). Peter's Delusions Inventory in Spanish general population: Internal reliability, factor structure and association with demographic variables (dimensionality of delusional ideation). Actas Espanolas de Psiquiatria, 34, 94-104. Retrieved from https://www.researchgate.net/publication/7225768_Peter%27s_delusions_inventory_in_Spanish_general_population_Internal_reliability_factor_structure_and_association_with_demographic_variables_dimensionality_of_delusional_ideation

McKay, R., Langdon, R., & Coltheart, M. (2006). The Persecutory Ideation Questionnaire. The Journal of Nervous and Mental Disease, 194, 628-631. doi:10.1093/med:psych/9780199206315.003.0002

Melo, S. S., & Bentall, R. P. (2013). 'Poor me' versus 'bad me' paranoia: The association between self-beliefs and the instability of persecutory ideation. Psychology and Psychotherapy: Theory, Research and Practice, 86, 146-163. doi:10.1111/j.2044-8341.2011.02051.x

Melo, S., Corcoran, R., Shryane, N., & Bentall, R. P. (2009). The Persecution and Deservedness Scale. Psychology and Psychotherapy: Theory, Research and Practice, 82, 247-260. doi:10.1037/t20074-000

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The Prisma statement. Annals of Internal

Medicine, 151, 264-269. doi:10.1093/ptj/89.9.873

Mokkink, L., Terwee, B., Patrick, C., Alonso, D., Stratford, L., Knol, J., … Vet, W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. Quality of Life Research, 19, 539-549. doi: 10.1007/s11136-010-9606-8

Mokkink, L., Terwee, C., Gibbons, E., Stratford, P., Alonso, J., Patrick, D., . . . De Vet, H. (2010). Inter-rater reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) Checklist. Quality Of Life Research, 19, 25. doi: 10.1186/1471-2288-10-82

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., … de Vet, H. C. W. (2012). The COSMIN checklist manual. Retrieved from http://www.cosmin.nl/cosmin_checklist.html

Moritz, S., Van Quaquebeke, N., & Lincoln, T. M. (2012). Jumping to conclusions is associated with paranoia but not general suspiciousness: A comparison of two versions of the probabilistic reasoning paradigm. Schizophrenia Research and Treatment, 2012, Article ID 384039. doi:10.1155/2012/384039

Moritz, S. & Woodward, T. S. (2007). Metacognitive training in schizophrenia: From basic research to knowledge translation and intervention. Current Opinion in Psychiatry, 20, 619-625. doi:10.1097/yco.0b013e3282f0b8ed

Morrison, A. P. (2001). The interpretation of intrusions in psychosis: An integrative cognitive approach to hallucinations and delusions. Behavioural and Cognitive Psychotherapy, 29, 257-276. doi:10.1017/s1352465801003010

Nittel, C. M., Lincoln, T. M., Lamster, F., Leube, D., Rief, W., Kircher, T., & Mehl, S. (2018). Expressive suppression is associated with state paranoia in psychosis: An

experience sampling study on the association between adaptive and maladaptive emotion regulation strategies and paranoia. British Journal of Clinical Psychology. doi:10.1111/bjc.12174

Peters, E. R., Joseph, S. A., Day, S., & Garety, P. A. (2004). Measuring delusional ideation: The 21-item Peters et al. Delusions Inventory (PDI). Schizophrenia Bulletin, 30, 1005-1022. doi:10.1093/oxfordjournals.schbul.a007116

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: Introducing the PDI (Peters et al. Delusions Inventory). Schizophrenia Bulletin, 25, 553-576. doi:10.1093/oxfordjournals.schbul.a033401

Preti, A., Marongiu, S., Petretto, D. R., Miotto, P., & Masala, C. (2002). Esperienze psichiche inusuali: Studio di validazione della versione italiana del Peters et al. Delusions Inventory. Rivista Sperimentale di Freniatria, 3-4, 81-100. Retrieved from www.rivistafreniatria.it

Prochwicz, K., & Gawęda, Ł. (2015). The Polish version of the Peters et al. Delusions Inventory: Factor analysis, reliability and the prevalence of delusion-like experiences in the Polish population. Psychiatria Polska, 49, 1203-1222. doi:10.12740/pp/42916

Rawlings, D., & Freeman, J. L. (1996). A questionnaire for the measurement of paranoia/suspiciousness. British Journal of Clinical Psychology, 35, 451-461. doi:10.1111/j.2044-8260.1996.tb01199.x

Rocchi, M. B., Sisti, D., Manca, S., Siddi, S., Mura, T., & Preti, A. (2008). Latent class analysis of delusion-proneness: Exploring the latent structure of the Peters et al. Delusions Inventory. The Journal of Nervous and Mental Disease, 196, 620-629. doi: 10.1097/NMD.0b013e31818132a3

Selten, J.-P., Wiersma, D., & van den Bosch, R. J. (2000). Clinical predictors of discrepancy

between self-ratings and examiner ratings for negative symptoms. Comprehensive Psychiatry, 41, 191–196. doi:10.1016/s0010-440x(00)90047-6

Schlier, B., Moritz, S., & Lincoln, T. M. (2016). Measuring fluctuations in paranoia: Validity and psychometric properties of brief state versions of the Paranoia Checklist. Psychiatry Research, 241, 323-332. doi:10.1016/j.psychres.2016.05.002

Schulze, K., Freeman, D., Green, C., & Kuipers, E. (2013). Intrusive mental imagery in patients with persecutory delusions. Behaviour Research and Therapy, 51, 7-14. doi:10.1016/j.brat.2012.10.002.

Smári, J., Stefánsson, S., & Thorgilsson, H. (1994). Paranoia, self-consciousness, and social cognition in schizophrenics. Cognitive Therapy and Research, 18, 387-399. doi:10.1007/BF02357512

Sutton, J. L., Gould, R. L., Daley, S., Coulson, M. C., Ward, E. V., Butler, A. M., … Howard, R. J. (2016). Psychometric properties of multicomponent tools designed to assess frailty in older adults: A systematic review. BMC Geriatrics, 16. doi:10.1186/s12877-016-0225-2

Stefanis, N. C., Smyrnis, N., Avramopoulos, D., Evdokimidis, I., Ntzoufras, I., & Stefanis, C. N. (2004). Factorial composition of self-rated schizotypal traits among young males undergoing military training. Schizophrenia Bulletin, 30, 335-350. doi:10.1093/oxfordjournals.schbul.a00708

Terwee, C. B., de Vet, H. C. W., Prinsen, C. A. C., & Mokkink, L. B. (2011). Protocol for systematic reviews on measurement properties. Retrieved from http://www.cosmin.nl/Systematic%20reviews%20of%20measurement%20properties.html

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. Quality of Life Research, 21, 651-657. doi:10.1007/s11136-011-9960-1

Trower, P., & Chadwick, P. (1995). Pathways to defense of the self: A theory of two types of paranoia. Clinical Psychology: Science and Practice, 2, 263-278. doi:10.1111/j.1468-2850.1995.tb00044.x

Van Dongen, J. D. M., Buck, N. M. L., Kool, A. M., & Van Marle, H. J. C. (2011). Psychometric evaluation of the Dutch Persecutory Ideation Questionnaire (PIQ) and its relation to aggression. Personality and Individual Differences, 51, 527-531. doi:10.1016/j.paid.2011.05.014

Verdoux, H., van Os, J., Maurice-Tison, S., Gay, B., Salamon, R., & Bourgeois, M. (1998). Is early adulthood a critical developmental stage for psychosis proneness? A survey of delusional ideation in normal subjects. Schizophrenia Research, 29, 247-254. doi:10.1016/s0920-9964(97)00095-9

Welham, J., Scott, J., Williams, G., Najman, J., Bor, W., O'Callaghan, M., & McGrath, J. (2009). Emotional and behavioural antecedents of young adults who screen positive for non-affective psychosis: A 21-year birth cohort study. Psychological Medicine, 39, 625–634. doi:10.1017/s0033291708003760

Wigham, S., & McConachie, H. (2014). Systematic review of the properties of tools used to measure outcomes in anxiety intervention studies for children with Autism Spectrum Disorders. PLoS ONE, 9, e85268. doi:10.1371/journal.pone.0085268