Published in final edited form as:

Psychiatry Res. 2014 May 15; 216(2): 263–268. doi:10.1016/j.psychres.2014.01.043.

Measuring insight through patient self-report: An in-depth analysis of the factor structure of the Birchwood Insight Scale

Sean D. Cleary^a, Sanaa Bhatti^b, Beth Broussard^c, Sarah L. Cristofaro^d, Claire Ramsay Wan^d, and Michael T. Compton^{c,e,*}

^aThe George Washington University School of Public Health and Health Services, Department of Epidemiology and Biostatistics, Washington, DC, USA

^bThe George Washington University School of Medicine and Health Sciences, Department of Psychiatry and Behavioral Sciences, Washington, DC, USA

^cLenox Hill Hospital, Department of Psychiatry, 111 E. 77th Street, New York, NY, USA

^dEmory University School of Medicine, Department of Psychiatry and Behavioral Sciences, Atlanta, GA, USA

^eHofstra North Shore—LIJ School of Medicine at Hofstra University, Department of Psychiatry, Hempstead, NY, USA

Abstract

Little research has focused on item analysis and factor structure of the most commonly used measures of insight. We examined the factorial structure of the Birchwood Insight Scale (BIS), a brief, easy-to-administer, self-report measure. We studied the BIS in 327 first-episode psychosis patients, including a test sample ($n=163$) and a validation sample ($n=164$). We then used data from 100 patients with chronic serious mental illnesses as a second, external validation sample. Exploratory factor analysis was conducted with the test subsample, and confirmatory factor analyses with the two validation samples. Confirmatory factor analyses (in both the first-episode psychosis validation sample and the chronic serious mental illness sample) indicated that a single-factor solution, with seven items loading on a single factor—with item 1 (“Some of your symptoms are made by your mind”) eliminated—was the best-fitting model. Seven of the eight original BIS items loading on a single factor fit the data well in these samples. Researchers using this efficient measure of patient-reported insight should assess the item distributions and factor structure of the BIS in their samples, and potentially consider eliminating item 1.

Keywords

Awareness of illness; Birchwood Insight Scale; First-episode psychosis; Insight; Psychosis; Schizophrenia; Serious mental illness

*Corresponding author at: Lenox Hill Hospital, Department of Psychiatry, 111 E. 77th Street, New York, NY 10075, USA. Tel.: +1 212 434 3215; fax: +1 212 434 3306. mcompton@nshs.edu (M.T. Compton).

1. Introduction

Insight, in the context of serious mental illnesses like schizophrenia and other psychotic disorders, pertains to one's recognition of having a mental illness and awareness that treatment could be beneficial. Rather than being a categorical phenomenon, insight has come to be viewed as continuous, multi-dimensional, and modifiable. Insight was defined by David as being composed of three distinct, yet somewhat overlapping dimensions: the recognition that one has a mental illness, adherence with treatment, and the ability to re-label unusual mental events (e.g., delusions and hallucinations) as pathological (David, 1990). Poor insight is common across the various serious mental illnesses, especially schizophrenia and bipolar disorder (Pini et al., 2001). Although for some this may be a stable feature, for others it improves with time and/or treatment (Wiffen et al., 2010). This is of importance as greater insight has been associated with lower symptomatology, better treatment adherence, improved social functioning, and better work performance (Lysaker et al., 2002; Mintz et al., 2003; Olfson et al., 2006; Melle and Barrett, 2012).

Several approaches exist for measuring insight in schizophrenia and other serious mental illnesses. Lincoln et al. (2007) suggested that a one-item measure might suffice for a screening of overall insight in this population. However, multiple-item scales allow one to learn more about specific aspects of insight and their clinical relevance (Baier, 2010). Such scales shown to be reliable and valid in this population include, but are not limited to, the Scale to Assess Unawareness of Mental Disorder (SUMD; Amador and Kronengold, 2004), which measures current and previous awareness of having a mental illness, the effects of medications, consequences, and the signs and symptoms of the illness; and the Beck Cognitive Insight Scale (BCIS; Beck et al., 2004), measuring patients' capacity for distancing themselves from and reevaluating anomalous beliefs and misinterpretations. The SUMD includes 74 items that form four subscales (current awareness, retrospective awareness, current attribution, and retrospective attribution); although it has satisfactory psychometric properties, it may be too long for studies that are not focusing specifically on insight but yet would benefit from a numerical insight score. The BCIS has two subscales (nine items tapping self-reflectiveness and six on self-certainty); its reported psychometric properties include Cronbach's alpha values of >0.70 for both subscales, moderate stability, and modest to moderate correlations with the SUMD.

The *Birchwood Insight Scale* (BIS; Birchwood et al., 1994) measures the three aforementioned dimensions of insight (ability to re-label symptoms, awareness of mental illness, and recognition of a need for treatment). The BIS, increasingly used in psychiatric research—especially studies involving persons with schizophrenia and other psychotic disorders—has a number of advantages, not the least of which are its brevity and ease of administration. In light of the instrument's substantial use despite limited information on psychometric properties and factor structure, we thoroughly explored item responses and the factor structure of the BIS. Specifically, we studied the BIS in a relatively large sample of first-episode psychosis patients — a large enough sample to allow us to randomly select a test sample as well as a validation sample. We then used data from a very different sample, comprised of patients with chronic serious mental illnesses, as a second, external validation sample. Our goal was to determine the most appropriate factor structure (and thus subscales)

of the BIS across these samples, to inform our own work and that of other researchers using this measure of patient self-reported insight.

2. Methods

2.1. Settings and samples

A total of 327 first-episode psychosis patients were recruited between August 2003 and December 2012 from five inpatient psychiatric units for two consecutive studies focusing on predictors of duration of untreated psychosis (Compton et al., 2008, 2009a, 2009b, 2011a) and the impact of premorbid cannabis use on the early course of psychotic disorders (Compton and Ramsay, 2009; Compton et al., 2009c). From the primary site (in Atlanta, Georgia), participants were recruited from a large, urban, university-affiliated, public-sector hospital ($n=249$), a suburban county psychiatric crisis center ($n=47$), and a local state psychiatric hospital ($n=18$). The two inpatient settings of the secondary site (Washington, DC), which only recently joined the study, included a private, university-affiliated hospital ($n=6$), and a community hospital ($n=7$). Criteria for inclusion were that participants were hospitalized for a first-episode, nonaffective psychotic disorder, aged 18–40 years, and English speaking. Participants were excluded if they had known or suspected mental retardation, a Mini-Mental State Examination (Folstein et al., 1975; Cockrell and Folstein, 1988) score of <24 , a significant medical condition that compromised ability to participate, a prior hospitalization >3 months before the index hospitalization, prior antipsychotic treatment for >3 months, or an inability to provide informed consent.

Sociodemographic characteristics of the first-episode psychosis sample (divided into a test sample and a validation sample, as described below) are shown in Table 1. The mean age at onset of psychosis was 21.5 ± 5.1 years. Diagnoses derived from the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID; First et al., 1998) were as follows: 196 (59.9%) had schizophrenia (139 with paranoid type, 34 with undifferentiated type, 19 with disorganized type, 2 with residual type, and 2 with catatonic type), 46 (14.1%) had schizophreniform disorder, 42 (12.8%) had psychotic disorder not otherwise specified, 32 (9.8%) had schizoaffective disorder (23 with depressive type and 9 with bipolar type), 6 (1.8%) had brief psychotic disorder, and 5 (1.5%) had delusional disorder.

With regard to the second sample of patients, who had chronic, serious mental illnesses, a total of 100 participants were recruited between April 2011 and April 2013 from four inpatient settings (Compton et al., 2011b). Recruitment sites included a state psychiatric hospital in Savannah, Georgia ($n=70$), and three crisis stabilization units in the southeastern region of Georgia ($n=30$). Criteria for inclusion required that participants were aged 18–65 years and English-speaking, were diagnosed with a psychotic or mood disorder, were discharged from the hospital or one of the three crisis stabilization units after a stay of ≥ 2 nights, and had at least one prior stay of ≥ 2 nights in the 6 months before the index admission. Participants were excluded if they had known or suspected mental retardation, a significant medical condition that compromised ability to participate, or an inability to provide informed consent.

Sociodemographic characteristics of the chronic, serious mental illness sample are shown in Table 1. Just over half of the sample was diagnosed with a mood disorder (54, 54.0%), and the remaining portion was diagnosed with a psychotic disorder (46, 46.0%).

Participants comprising both samples were administered the eight-item BIS as part of the research assessments of the larger studies. Due to the relatively low literacy among some study participants, and to standardize administration, research assessors read the items to all participants and recorded their verbal responses. As such, we made very minor modifications to the wording of items to make them suitable for researcher-rather than self-administration (e.g., “I am mentally well” was changed to “You are mentally well,” and “My stay in the hospital is necessary” was changed to “Your stay in the hospital is necessary.”)

2.2. Data analyses

We chose to use factor analysis as our analytic method primarily because (1) our goal was to characterize the factor structure (and specifically the proposed three-factor structure) of the BIS, (2) the BIS has been hypothesized to be multidimensional despite little research, and (3) the scale uses ordinal response options. While it is possible to examine multidimensional models using item-response theory (IRT) models, they typically require a much larger sample size (e.g., 200–500) than was available for accurate parameter estimates (Tsutakawa and Johnson, 1990). In addition, assumptions made when using IRT include a unidimensional latent structure of the underlying construct and dichotomous responses. Because neither of these assumptions could be made based on the available theoretical model of the BIS, we elected to use exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) methods, with a different sample for cross-validation to gain a clear understanding of the scale’s underlying factor structure. To maximize the use of the available data, we randomly divided the first-episode psychosis sample into test ($n=163$) and validation ($n=164$) subsamples. The distributions of the BIS items in these and the chronic serious mental illness sample were examined. EFA was conducted with the test subsample. Due to the ordinal nature of the response options, we used a maximum likelihood estimator with robust standard errors (MLR). Models with one, two, and three factors, composed of the original eight items and then seven items, as described below, were compared for fit. Models with two or more factors were rotated, using an oblique geomin rotation, to facilitate interpretation. CFA was conducted on the first-episode psychosis validation sample, and then the chronic serious mental illness sample for external validation. Three CFA models were specified and compared for fit: an eight-item single factor model and a seven-item single-factor model based on the results from the EFA, and a three-factor model representing the original specification of the BIS. Standard fit statistics were assessed, including the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), as well as Bentler comparative fit index (CFI) and non-normed fit index (NNFI). Correlations between factors were examined. Cronbach’s alpha, a measure of internal consistency reliability, was assessed for resulting factors/subscales. All analyses were conducted using SAS 9.3 (SAS Institute, 2002–2010) and Mplus 7.0 (Muthén and Muthén, 1998–2012) software.

3. Results

3.1. Distributions of BIS item responses across samples

The distributions of responses to the BIS items are displayed by sample in Table 2. Cross-tabulation of each item by the randomly selected sample subsets indicated that there were no significant differences between the first-episode psychosis test and validation samples for any item, and gender distributions were also equivalent, both suggesting that the random subsampling was valid. Significant differences between the first-episode psychosis sample and the chronic serious mental illness sample were observed for most BIS items, likely driven by the prominent demographic and clinical differences between the two groups.

3.2. Exploratory factor analysis in the first-episode psychosis test subsample

EFA was conducted using MLR in the first-episode psychosis test subsample ($n=163$). Table 3 displays the factor loadings and residuals for the one-factor (Model 1), two-factor (Model 2), and three-factor (Model 3) solutions. Model 3 (likelihood difference $\chi^2 = 50.19$, d.f. = 14, $p < 0.0001$) and Model 2 (difference $\chi^2 = 54.21$, d.f. = 11, $p < 0.0001$) fit the data better than Model 1, as indicated by the likelihood difference test results, but Model 3 did not fit the data better than Model 2 (difference $\chi^2 = 4.02$, d.f. = 3, $p > 0.05$).

In evaluating the structures of each of the models, items that loaded on more than one factor were eliminated. For Model 3, items 5 and 7 were dropped, leaving the third factor in Model 3 with only a single item (item 8). All other items in Model 3 loaded significantly on factor 1 (items 2, 4, and 6) and factor 2 (items 1 and 3); however, the residuals indicated that the underlying factors accounted for less than 5% of the variance for item 1 and only approximately 50% for items 5–7. For Model 2, five items (items 1, 3, 4, 7, and 8) did not load significantly on either factor and were eliminated. The resulting factor structure included only two items (items 6 and 7) for factor 1 and one item (item 2) for factor 2, a very unstable model, and one with poor face validity. Model 1 was the strongest with respect to all items loading significantly on the single factor, with item 1 being the weakest. Cronbach's alpha was acceptable (0.78).

Because item 1 loaded minimally relative to all other items, we reran the EFA without item 1. The same three models were evaluated, with results displayed in Table 4. Evaluation of the fit based on the likelihood difference revealed that Model 1 (the one-factor solution) fits as well as Model 2 and Model 3. Examination of the structure of Model 3 indicated loss of three items (items 2, 3, and 5) that loaded on more than one factor, with two of the three factors then having only one item. Model 2 had four items loading on both factors, resulting in a single factor with three items. Model 1 was the most acceptable, with all seven items loading significantly on the single factor, and a Cronbach's alpha of 0.80. Fit statistics for the EFAs involving both eight items and seven items (the latter with item 1 eliminated) are shown in Tables 3 and 4, respectively.

3.3. Confirmatory factor analysis in the first-episode psychosis validation sample

CFA was conducted using MLR with the first-episode psychosis validation sample ($n=164$). Based on the empirical results from the EFA as well as the theoretical subscale structure of

the BIS, three models were specified and the fit compared. Model 1 had eight items loading on a single factor and Model 2 had seven items loading on a single factor (with item 1 eliminated). Model 3 was specified as the original subscale structure, with the eight items loading on *re-labeling symptoms* (items 1 and 8), *awareness of illness* (items 2 and 7), and *need for treatment* (items 3, 4, 5, and 6). Table 5 displays the factor structure for each model, along with fit statistics. Model 2, seven items loading on a single factor, had the lowest AIC and BIC. Likelihood difference tests indicated significant improvement in fit for Model 2 compared to both Model 1 and Model 3. Although all of the factor loadings in Model 3 were significant, the correlations between the factors were all >0.55 , suggesting a unidimensional model rather than a three-factor model. In Model 1, item 1 was the lowest relative loading, although it is statistically significant, but Model 2 fits the data better.

3.4. Confirmatory factor analysis in the chronic serious mental illness external validation sample

We used data from an independent source, a sample of participants with chronic serious mental illness ($n=99$, one missing value) who also completed the BIS, to validate the results observed in the first-episode psychosis dataset. A CFA was conducted, as described above, comparing the three specified models. Results are displayed in Table 5. Despite the differences in the distribution of responses to the eight items (noted previously in Table 2), the CFA using the chronic serious mental illness sample supported the results from the first-episode psychosis validation sample. Specifically, Model 2 (seven items loading on a single factor, with item 1 eliminated) had the lowest AIC and BIC, indicating the best-fitting model. In Model 1 (eight items loading on a single factor), item 1 did not load significantly on this single-factor model. In Model 3 (the original subscale structure with three subscales), both item 1 and item 2 did not load significantly on factor 1, indicating that there are only two factors. Correlations between Model 3 factors were >0.55 , further supporting the conclusion that a unidimensional model best represents the data.

4. Discussion

Our results indicate that seven of the eight BIS items loading on a single factor fit the data well in all three samples. Findings suggest that researchers assess the item distributions and factor structure of the BIS in their samples, and potentially consider eliminating item 1. It is interesting to speculate about the poor performance of item 1. We postulate that, unlike the other items, which are relatively concrete, the use of “made by your mind” in item 1 is variably interpreted by respondents. Some participants with hallucinations, for example, come to think of their voices as “made by their mind” meaning to them that the voices require therapy and/or medications (which clinicians would deem as good insight). Others, however, experience their voices also as “made by their mind,” but, as such, not the manifestation of an illness (which clinicians would equate with poor insight). A similar dichotomy in how “made by your mind” is interpreted could occur in the context of depression, anxiety, and other psychiatric symptoms. Another potential reason for the poor performance of item 1 is that it refers to “some of your symptoms” (rather than just “your symptoms” or “all of your symptoms”), which could influence responses in a manner different from the other seven items. Finally, it is possible that the interpretation of “made by

your mind” could vary across cultural groups. Although the exact reason for the poor performance of item 1 can only be determined empirically (e.g., by asking participants how they understand this item), any of these potential problems would render the item unclear, resulting in variable interpretations and thus measurement error and poor validity of the item.

Insight is a multi-dimensional and complex construct that reflects a person’s understanding of his or her mental illness, its impact on functioning, and the need for treatment. Clinician-rated measures are of value in assessing patients’ expressed attitudes and behaviors, but self-report measures also contribute important information on patients’ opinions and internal experiences. The general accuracy of self-report from individuals with schizophrenia has been questioned, precisely due to poor insight and motivational deficits (Bell et al., 2007). Yet, clinician-rated instruments are subject to inter-rater biases and the possibility that raters’ assessments of insight are influenced by patients’ cognitive functioning and communication skills (Marks et al., 2000). Clinician-rated and self-report measures of insight in schizophrenia and in first-episode psychosis in particular have had a moderate to high level of agreement in the few studies that have used both (Sanz et al., 1998; Marks et al., 2000; Young et al., 2003; Jovanovski et al., 2007). However, this correlation was strongly influenced by the order of administration in two small studies (Young et al., 2003; Jovanovski et al., 2007) and self-report and clinician-rated scales have yielded somewhat different correlations with data on other psychiatric symptoms (Sanz et al., 1998). Self-report measures are quicker to administer and more easily incorporated into research and clinical settings. However, it is critical to test their psychometric properties in different populations. Although multidimensionality of the insight construct—and insight scales including the BIS—has been proposed, our results indicate that the items of the BIS reflect a single dimension. To accurately measure a more multidimensional insight construct would require that more items are included.

We acknowledge two key methodological limitations. First, a primary concern is one of selection bias. To the extent that our first-episode psychosis sample and chronic serious mental illness sample are both representative, the results should be generalizable. However, the first-episode psychosis sample was comprised largely of urban, socially disadvantaged, low-income, African American males, and all were hospitalized. While the chronic serious mental illness sample might be more generalizable, it included participants with either psychotic or affective disorders who were being discharged from the hospital, and the sample was drawn from a relatively rural region. Until these same analyses are conducted in dissimilar samples, we can only assume generalizability to similar samples. Second, as noted, we made very minor modifications to the wording of several items to allow us to read the items to participants (e.g., “your symptoms”) rather than having them self-administer the items (e.g., “my symptoms”). Thus, our analysis relied on data collected in a researcher-administered format (to avoid problems associated with low literacy) rather than a truly self-administered format. Although it is conceivable that this could influence responses, it is unlikely to have affected the factor structure.

Despite these limitations, a prominent strength of this analysis is that we were able to conduct both an EFA and then follow-up CFAs in two samples for validation; few studies

follow this stepwise approach in evaluating an instrument. The data in this report offer valuable information on the factor structure of the BIS from samples of first-episode psychosis patients and individuals with chronic serious mental illnesses in two different settings in the southeastern United States. Future studies should investigate if this structure is consistent in other populations and determine what other features, such as administration method, might influence findings.

Acknowledgments

The research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award nos. K23 MH067589 and R01 MH081011, as well as a grant from the Bristol-Myers Squibb Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Bristol-Myers Squibb Foundation.

References

- Amador, XF., Kronengold, H. Understanding and assessing insight. In: Amador, XF., David, AS., editors. *Insight and Psychosis: Awareness of Illness in Schizophrenia and Related Disorders*. 2. Oxford University Press; London: 2004. p. 197-229.
- Baier M. Insight in schizophrenia: a review. *Current Psychiatry Reports*. 2010; 12:356–361. [PubMed: 20526897]
- Beck AT, Baruch E, Balter JM, Steer RA, Warman DM. A new instrument for measuring insight: the Beck Cognitive Insight Scale. *Schizophrenia Research*. 2004; 68:319–329. [PubMed: 15099613]
- Bell M, Fiszdon J, Richardson R, Lysaker P, Bryson G. Are self-reports valid for schizophrenia patients with poor insight? Relationship of unawareness of illness to psychological self-report instruments. *Psychiatry Research*. 2007; 151:37–46. [PubMed: 17343920]
- Birchwood M, Smith J, Drury V, Healy J, Macmillan F, Slade M. A self-report Insight Scale for psychosis: reliability, validity and sensitivity to change. *Acta Psychiatrica Scandinavica*. 1994; 89:62–67. [PubMed: 7908156]
- Cockrell JR, Folstein MF. Mini-Mental State Examination (MMSE). *Psychopharmacology Bulletin*. 1988; 24:689–692. [PubMed: 3249771]
- Compton MT, Chien VH, Leiner AS, Goulding SM, Weiss PS. Mode of onset of psychosis and family involvement in help-seeking as determinants of duration of untreated psychosis. *Social Psychiatry and Psychiatric Epidemiology*. 2008; 43:975–982. [PubMed: 18604616]
- Compton MT, Gordon TL, Goulding SM, Esterberg ML, Carter T, Leiner AS, Weiss PS, Druss BG, Walker EF, Kaslow NJ. Patient-level predictors and clinical correlates of duration of untreated psychosis among hospitalized first-episode patients. *Journal of Clinical Psychiatry*. 2011a; 72:225–232. [PubMed: 21382306]
- Compton MT, Goulding SM, Gordon TL, Weiss PS, Kaslow NJ. Family-level predictors and correlates of the duration of untreated psychosis in African American first-episode patients. *Schizophrenia Research*. 2009a; 115:338–345. [PubMed: 19833482]
- Compton MT, Hankerson-Dyson D, Broussard B, Druss BG, Haynes N, Strode P, Grimes C, Li C, DiPolito JA, Thomas GV. Opening Doors to Recovery: a novel community navigation service for people with serious mental illnesses. *Psychiatric Services*. 2011b; 62:1270–1272. [PubMed: 22211204]
- Compton MT, Kelley ME, Ramsay CE, Pringle M, Goulding SM, Esterberg ML, Stewart T, Walker EF. Association of pre-onset cannabis, alcohol, and tobacco use with the age at onset of prodrome and age at onset of psychosis in first-episode patients. *American Journal of Psychiatry*. 2009c; 166:1251–1257. [PubMed: 19797432]
- Compton MT, Ramsay CE. The impact of pre-onset cannabis use on age at onset of prodromal and psychotic symptoms. *Primary Psychiatry*. 2009; 16:35–43.

- Compton MT, Ramsay CE, Shim RS, Goulding SM, Gordon TL, Weiss PS, Druss BG. Health services determinants of the duration of untreated psychosis among African-American first-episode patients. *Psychiatric Services*. 2009b; 60:1489–1494. [PubMed: 19880467]
- David AS. Insight and psychosis. *British Journal of Psychiatry*. 1990; 156:798–808. [PubMed: 2207510]
- First, MB., Spitzer, RL., Gibbon, M., Williams, JBW. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID). Biometrics Research. New York State Psychiatric Institute; New York, NY: 1998.
- Folstein MF, Folstein SE, McHugh PR. “Mini-mental state.” A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- Jovanovski D, Zakzanis KK, Atia M, Campbell Z, Young DA. A comparison between researcher-rated and self-report method of insight assessment in chronic schizophrenia revisited: a replication study using the SUMD and SAIQ. *Journal of Nervous and Mental Disease*. 2007; 195:165–169. [PubMed: 17299305]
- Lincoln TM, Lullmann E, Rief W. Correlates and long-term consequences of poor insight in patients with schizophrenia: a systematic review. *Schizophrenia Bulletin*. 2007; 33:1324–1342. [PubMed: 17289653]
- Lysaker PH, Bryson GJ, Bell MD. Insight and work performance in schizophrenia. *Journal of Nervous and Mental Disease*. 2002; 190:142–146. [PubMed: 11923648]
- Mark KA, Fatenau PS, Lysaker PH, Bond GR. Self-Appraisal of Illness Questionnaire (SAIQ): relationship to researcher-rated insight and neuropsychological function in schizophrenia. *Schizophrenia Research*. 2000; 45:203–211. [PubMed: 11042438]
- Melle I, Barrett EA. Insight and suicidal behavior in first-episode schizophrenia. *Expert Review of Neurotherapeutics*. 2012; 12:353–359. [PubMed: 22364334]
- Mintz AR, Dobson KS, Romney DM. Insight in schizophrenia: a meta-analysis. *Schizophrenia Research*. 2003; 61:75–88. [PubMed: 12648738]
- Muthén, LK., Muthén, BO. *Mplus User’s Guide*. 6. Muthén & Muthén; Los Angeles: 1998–2012.
- Olfson M, Marcus SC, Wilk J, West JC. Awareness of illness and nonadherence to antipsychotic medications among persons with schizophrenia. *Psychiatric Services*. 2006; 57:205–211. [PubMed: 16452697]
- Pini S, Cassano GB, Dell’Osso L, Amador XF. Insight into illness in schizophrenia, schizoaffective disorder, and mood disorders with psychotic features. *American Journal of Psychiatry*. 2001; 158:122–125. [PubMed: 11136644]
- Sanz M, Constable G, Lopez-Ibor I, Kemp R, David AS. A comparative study of insight scales and their relationship to psychopathological and clinical variables. *Psychological Medicine*. 1998; 28:437–446. [PubMed: 9572100]
- SAS Institute Inc. SAS 9.3. SAS Institute, Inc; Cary, North Carolina: 2002–2010.
- Tsutakawa RK, Johnson JC. The effect of uncertainty of item parameter estimation on ability estimates. *Psychometrika*. 1990; 55:371–390.
- Wiffen BD, Rabinowitz J, Lex A, David AS. Correlates, change and ‘state or trait’ properties of insight in schizophrenia. *Schizophrenia Research*. 2010; 122:94–103. [PubMed: 20382507]
- Young DA, Campbell Z, Zakzanis KK, Weinstein E. A comparison between an interview and a self-report method of insight assessment in chronic schizophrenia. *Schizophrenia Research*. 2003; 63:103–109. [PubMed: 12892864]

Table 1

Sociodemographic characteristics of the three study samples.

	First-episode psychosis test sample (<i>n</i> =165) ^a	First-episode psychosis validation sample (<i>n</i> =168) ^a	Chronic serious mental illness sample (<i>n</i> =100) ^a
Age at hospitalization, years	24.0±5.1 (range: 18–39)	23.5±4.3 (range: 18–38)	37.3±13.0 (range: 18–65)
Educational attainment, years	11.8±2.1 (range: 6–19)	11.8±2.3 (range: 5–19)	10.8±2.1 (range: 4–18)
Gender, male	118 (71.5%)	130 (77.4%)	53 (53.0%)
Race			
African-American	143 (86.7%)	147 (87.5%)	46 (46.0%)
Caucasian	13 (7.9%)	12 (7.1%)	52 (52.0%)
Other	9 (5.5%)	9 (5.4%)	2 (2.0%)
Relationship status ^b			
Single and never married	145 (87.9%)	149 (89.2%)	50 (50.0%)
Married or living with a partner	7 (4.2%)	8 (4.8%)	9 (9.0%)
Separated, divorced, or widowed	13 (7.8%)	10 (6.0%)	41 (41.0%)
Who the patient lived with before hospitalization ^c			
Family members	104 (63.4%)	113 (67.7%)	42 (42.0%)
Alone	14 (8.5%)	12 (7.2%)	25 (25.0%)
Boyfriend, girlfriend, spouse, or partner	11 (6.6%)	10 (6.0%)	10 (10.0%)
Friends or roommate	8 (4.8%)	9 (5.4%)	10 (10.0%)
Homeless	14 (8.5%)	13 (7.8%)	8 (8.0%)
Other	13 (7.9%)	8 (4.8%)	5 (5.0%)
Work status before hospitalization, unemployed	105 (63.6%)	113 (67.3%)	85 (85.0%)

^aSubsequent analyses relied on *n*=163, *n*=164, and *n*=99 across the three samples, respectively, due to sporadic missing data.^bAvailable *n* for first-episode psychosis validation sample is 167.^cAvailable *n* for first-episode psychosis test sample is 164, and for first-episode validation sample is 165.

Table 2

Distributions of Birchwood Insight Scale item responses across samples.

	First-episode psychosis		Chronic serious mental illness
	Test sample (%) n=163	Validation sample (%) n=164	External validation sample (%) n=99
1. Some of your symptoms are made by your mind.			
Disagree	22.1	24.4	39.4
Unsure	16.6	15.2	11.1
Agree	61.4	60.4	49.5
2. You are mentally well.			
Agree	65.6	68.9	45.5
Unsure	17.2	14.0	9.1
Disagree	17.2	17.1	45.5
3. You do not need medication.			
Agree	45.4	40.9	19.2
Unsure	17.2	18.3	5.1
Disagree	37.4	40.9	75.8
4. Your stay in the hospital is necessary.			
Disagree	49.1	46.3	35.4
Unsure	12.9	9.1	3.0
Agree	38.0	44.5	61.6
5. The doctor is right in prescribing medication for you.			
Disagree	20.2	20.1	8.1
Unsure	19.0	15.2	3.0
Agree	60.7	64.6	88.9
6. You do not need to be seen by a doctor or psychiatrist.			
Agree	33.7	37.8	13.1
Unsure	13.5	14.6	3.0
Disagree	52.8	47.6	83.8
7. If someone said you have a nervous or mental illness, they would be right.			
Disagree	49.1	43.9	29.3
Unsure	17.2	15.2	4.0
Agree	33.7	40.9	66.7
8. None of the unusual things you are experiencing are due to an illness.			
Agree	37.4	41.5	28.3
Unsure	26.4	23.8	13.1
Disagree	36.2	34.8	58.6

Table 3
Exploratory factor analysis of the Birchwood Insight Scale (first-episode psychosis test subsample, $n=163$)—8 items.

Item	Model 1			Model 2			Model 3			
	Factor 1 Loading (SE)	Residual		Factor 1 Loading (SE)	Factor 2 Loading (SE)	Residual	Factor 1 Loading (SE)	Factor 2 Loading (SE)	Factor 3 Loading (SE)	Residual
1	0.25 (−0.11)*	0.93		0.33 (0.23)	−0.10 (0.33)	0.92	0.01 (0.01)	0.21 (0.10)*	−0.02 (0.04)	0.96
2	0.76 (0.07)*	0.42		0.53 (0.28)	0.33 (0.16)*	0.42	1.38 (0.33)*	−0.56 (0.41)	0.00 (0.00)	0.00
3	0.79 (0.06)*	0.38		0.55 (0.29)	0.34 (0.20)	0.38	0.01 (0.01)	0.94 (0.07)*	0.14 (0.19)	0.00
4	0.60 (0.09)*	0.64		0.80 (1.19)	−0.16 (1.25)	0.48	0.56 (0.20)*	0.12 (0.23)	−0.09 (0.12)	0.62
5	0.71 (0.08)*	0.49		0.64 (0.31)*	0.12 (0.29)	0.49	0.36 (0.16)*	0.40 (0.16)*	0.01 (0.02)	0.48
6	0.76 (0.07)*	0.43		0.79 (0.21)*	0.00 (0.08)	0.37	0.69 (0.22)*	0.06 (0.23)	−0.03 (0.11)	0.48
7	0.77 (0.06)*	0.41		0.43 (0.90)	0.47 (0.84)	0.38	0.56 (0.09)*	0.00 (0.02)	0.27 (0.10)*	0.46
8	0.54 (0.09)*	0.71		0.00 (0.02)	0.78 (0.85)	0.40	0.00 (0.01)	0.01 (0.03)	0.98 (0.01)*	0.00
Free parameters		24				31				37
Log likelihood		−1140.01				−1132.98				−1115.23
Akaike information criteria		2328.02				2327.95				2304.45
Bayesian information criteria		2402.27				2423.86				2418.92
Pearson χ^2		2463.55				1940.56				2001.16
Likelihood ratio χ^2		598.43				544.22				548.24
d.f.		6524				6513				6510

* Significant at the 5% level.

Table 4Exploratory factor analysis of the Birchwood Insight Scale (first-episode psychosis test subsample, $n=163$)—7 items (item 1 eliminated).

Item	Model 1			Model 2			Model 3			
	Factor 1 Loading (SE)	Residual		Factor 1 Loading (SE)	Factor 2 Loading (SE)	Residual	Factor 1 Loading (SE)	Factor 2 Loading (SE)	Factor 3 Loading (SE)	Residual
2	0.77 (0.06) *	0.41		0.56 (0.10) *	0.33 (0.10) *	0.42	0.13 (0.05) *	0.47 (0.15) *	0.37 (0.15) *	0.41
3	0.79 (0.06) *	0.37		0.55 (0.11) *	0.37 (0.12) *	0.39	0.95 (0.03) *	−0.00 (0.04)	0.11 (0.04) *	0.00
4	0.60 (0.09) *	0.65		0.65 (0.10) *	0.02 (0.09)	0.57	0.24 (0.12)	0.50 (0.12) *	−0.01 (0.10)	0.58
5	0.72 (0.08) *	0.49		0.64 (0.12) *	0.15 (0.13)	0.49	0.46 (0.13) *	0.33 (0.14) *	0.07 (0.12)	0.48
6	0.76 (0.07) *	0.43		0.81 (0.10) *	0.05 (0.09)	0.31	0.13 (0.09)	0.81 (0.14) *	0.04 (0.08)	0.20
7	0.76 (0.06) *	0.42		0.48 (0.14) *	0.42 (0.15) *	0.42	0.17 (0.14)	0.39 (0.17) *	0.41 (0.23)	0.42
8	0.55 (0.09) *	0.70		−0.11 (0.03) *	0.99 (0.06) *	0.09	0.08 (0.08)	−0.06 (0.06)	0.92 (0.22) *	0.14
Free parameters		21				27				32
Log likelihood		−990.81				−984.27				−979.51
Akaike information criteria		2023.62				2022.54				2023.02
Bayesian information criteria		2088.59				2106.07				2122.02
Pearson χ^2		1438.67				1543.49				1407.87
Likelihood ratio χ^2		450.92				448.29				435.55
d.f.		2158				2153				2148

* Significant at the 5% level.

Table 5

Confirmatory factor analyses of the birchwood insight scale in the two validation samples.

Item	8-Item Model		7-Item Model		Original Scale Structure		
	Factor 1	R ²	Factor 1	R ²	Re-labeling symptoms	Need for treatment	Awareness of illness
	Loading (SE)		Loading (SE)		Loading (SE)	Loading (SE)	Loading (SE)
<i>First-Episode Psychosis Validation Subsample (n=164)</i>							
1	0.36 (0.10)	0.13	–	–	0.36 (0.10)	–	–
2	0.69 (0.10)	0.48	0.70 (0.09)	0.49	–	–	0.82 (0.09)
3	0.77 (0.06)	0.59	0.77 (0.06)	0.60	–	0.77 (0.07)	–
4	0.69 (0.07)	0.48	0.68 (0.08)	0.46	–	0.69 (0.08)	–
5	0.69 (0.08)	0.48	0.69 (0.08)	0.48	–	0.70 (0.08)	–
6	0.75 (0.07)	0.56	0.76 (0.07)	0.57	–	0.77 (0.06)	–
7	0.69 (0.07)	0.48	0.67 (0.07)	0.45	–	–	0.82 (0.07)
8	0.65 (0.07)	0.42	0.64 (0.07)	0.41	0.71 (0.12)	–	–
Free parameters	24			21			27
Log likelihood	–1126.48			–979.88			–1121.39
Akaike information criteria	2300.96			2001.76			2296.77
Bayesian information criteria	2375.35			2066.86			2380.47
Pearson χ^2	1984.01			1582.16			2249.52
Likelihood ratio χ^2	528.19			461.81			554.19
d.f.	6519			2161			6520
<i>Chronic serious mental illness validation sample (n=99)</i>							
1	0.06 (0.08)*	0.00*	–	–	0.05 (0.08)*	–	–
2	0.79 (0.09)	0.63	0.79 (0.09)	0.63	–	–	0.86 (0.09)
3	0.63 (0.12)	0.39	0.63 (0.12)	0.40	–	0.69 (0.11)	–
4	0.73 (0.10)	0.54	0.74 (0.10)	0.54	–	0.82 (0.09)	–
5	0.69 (0.15)	0.47	0.69 (0.15)	0.47	–	0.77 (0.14)	–
6	0.82 (0.09)	0.68	0.82 (0.09)	0.67	–	0.78 (0.12)	–
7	0.80 (0.10)	0.64	0.80 (0.10)	0.64	–	–	0.88 (0.10)
8	0.44 (0.12)	0.20	0.44 (0.12)	0.20	0.78 (0.25)	–	–
							0.60*

Item	8-Item Model		7-Item Model		Original Scale Structure		
	Factor 1	R^2	Factor 1	R^2	Re-labeling symptoms	Need for treatment	Awareness of illness
					Loading (SE)	Loading (SE)	Loading (SE)
Free parameters		24		21			27
Log likelihood		-534.54		-439.69			-531.13
Akaike Information criteria		1117.08		921.38			1116.26
Bayesian information criteria		1179.36		975.88			1186.32
Pearson χ^2		967.66		443.44			1008.16
Likelihood ratio χ^2		234.79		160.97			229.21
d.f.		6530		2160			6527

* Not significant at the 5% level.