Next-generation negative symptom assessment for clinical trials: Validation of the Brief Negative Symptom Scale

Gregory P. Strauss a,⁎, William R. Keller a, Robert W. Buchanan a,b, James M. Gold a, Bernard A. Fischer a,b, Robert P. McMahon a, Lauren T. Catalano a, Adam J. Culbreth a, William T. Carpenter a,b, Brian Kirkpatrick c

a University of Maryland School of Medicine, Department of Psychiatry and Maryland Psychiatric Research Center, Baltimore, MD, USA
b VA Maryland Health Care System, Baltimore, MD, USA
c Texas A&M College of Medicine, Scott & White Healthcare, Temple, TX, USA

A R T I C L E  I N F O

Article history:
Received 9 August 2012
Received in revised form 8 October 2012
Accepted 10 October 2012
Available online 3 November 2012

Keywords:
Anhedonia
Alogia
Asociality
Blunted Affect
Avolition

A B S T R A C T

The current study examined the psychometric properties of the Brief Negative Symptom Scale (BNSS), a next-generation rating instrument developed in response to the NIMH sponsored consensus development conference on negative symptoms. Participants included 100 individuals with a DSM-IV diagnosis of schizophrenia or schizoaffective disorder who completed a clinical interview designed to assess negative, positive, disorganized, and general psychiatric symptoms, as well as functional outcome. A battery of anhedonia questionnaires and neuropsychological tests were also administered. Results indicated that the BNSS has excellent internal consistency and temporal stability, as well as good convergent and discriminant validity in its relationships with other symptom rating scales, functional outcome, self-reported anhedonia, and neuropsychological test scores. Given its brevity (13-items, 15-minute interview) and good psychometric characteristics, the BNSS can be considered a promising new instrument for use in clinical trials.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Negative symptoms are a significant barrier to successful functional outcome and recovery in individuals with schizophrenia (Strauss et al., 2010, 2012). They also represent a primary unmet need in schizophrenia therapeutics, as no drug has received Food and Drug Administration (FDA) approval for an indication of negative symptoms. Although the importance of studying negative symptoms may be clear, ideas regarding which aspects of psychopathology should be considered part of the negative symptom construct have changed over the years. Symptom rating scales developed in the 1980s regarded such clinical features as poverty of content of speech, inappropriate affect, and attention, to be negative symptoms (e.g., Andreasen, 1982); however, factor analytic studies show that these symptoms are more closely tied to other aspects of pathology (e.g., disorganization) than negative symptoms (Buchanan and Carpenter, 1994). In 2005, NIMH hosted a consensus development conference on negative symptoms in an effort to address some of these issues. Several important conclusions resulted from this meeting. Among these were that: 1) there are at least 5 core domains within the negative symptom construct, including: restricted affect, alogia, avolition, anhedonia, and asociality; and 2) there was a need for the development of new negative symptom rating scales designed to measure these 5 domains, while excluding content thought to be unrepresentative of the negative symptom construct (Kirkpatrick et al., 2006).

In response to the consensus conference and the NIMH MATRICS initiative on negative symptoms, we previously reported the development of a new negative symptom rating instrument, the Brief Negative Symptom Scale (BNSS: Kirkpatrick et al., 2011). Guiding principles behind BNSS development included: (1) that it be concise so that these could be considered separately.

Our initial study of the BNSS involved conducting 20 video-taped interviews of people with schizophrenia, which were then rated by 7 individuals from 3 institutions who had varying academic backgrounds. Psychometric analyses indicated that the BNSS had strong inter-rater reliability (0.94), as well as internal-consistency, test–retest reliability, and convergent, discriminant, and predictive validity. Principal components analysis supported the construct...
validity of the BNSS, indicating a 2 factor solution reflecting an Emotional Expressivity domain and a Motivation/Pleasure domain. Thus, our initial study provided preliminary evidence that the BNSS has good psychometric properties.

In the current study, we aimed to extend our initial investigation by examining the psychometric properties of the BNSS in a larger sample of outpatients who received a more extensive battery of measures used to assess convergent and discriminant validity.

2. Methods

2.1. Participants

Participants included 100 individuals meeting DSM-IV criteria for schizophrenia (n=88) or schizoaffective disorder (n=12). Participants were recruited through the Maryland Psychiatric Research Center, Outpatient Research Program, and evaluated during a period of clinical stability, as indicated by no changes in medication type or dosage for a period of 4 or more weeks prior to the evaluation. Participants were diagnosed using a best estimate diagnostic approach that utilized information from the Structured Clinical Interview for DSM-IV (First et al, 1997), direct assessment, family informants, and past medical records. Exclusion criteria included substance abuse or dependence in the past 6 months and history of head injury or neurological disorder.

Participants were on average 42.2 (11.1) years old, had 12.7 (2.1) years of personal education, and 13.4 (2.8) years of parental education. Seventy-four percent of the participants were male, 62% were Caucasian, 32% were African-American, 1% Asian-American, 1% American Indian, and 4% were mixed race. Seventy-nine percent of the participants were prescribed second generation antipsychotic medications, 16% first-generation antipsychotic medications, 6% of participants were prescribed both second generation and first antipsychotic medications, and 1 participant was unmedicated (but clinically stable) at the time of evaluation. All participants provided informed consent for a protocol approved by the University of Maryland Institutional Review Board.

2.2. Procedures

A battery of psychiatric rating instruments were administered to all participants, including: 1) Brief Negative Symptom Scale (BNSS: Kirkpatrick et al., 2011), 2) Scale for the Assessment of Negative Symptoms (SANS: Andreasen, 1982; Buchanan et al., 2007), 3) Brief Psychiatric Rating Scale (BPRS: Overall and Gorham, 1962), 4) Schedule for the Deficit Syndrome (SDS: Kirkpatrick et al., 1989), and 5) Level of Function Scale (LOF: Hawk et al., 1975). Four clinical raters, who completed ratings for all scales, were trained to the following reliability standards prior to conducting the assessments: inter-rater agreement >0.80; 8/10 Kappa agreement on SDS. Raters had a bachelor’s degree or higher and at least one year of clinical experience. The BNSS was completed on a subset of participants again by the same rater at a later point to assess stability. A longer interval was selected to provide an estimate of measurement error that takes into account stability of symptom presentation, while minimizing potential for carry-over effects that bias temporal consistency estimates when retest intervals are short.

Rater training consisted of an in-depth review of the BNSS manual and workbook, as well as procedures for rating the BPRS, SANS, and LOF. The clinical raters watched a series of video-recorded interviews made by the first author, which covered content necessary for rating the BNSS, SANS, BPRS, and LOF, and then rated the participant in the video on those measures. Ratings were then discussed as a group, and raters were instructed in interviewing technique and observed in conducting interviews. The interviewers subsequently received ongoing supervision and participated in quarterly gold-standard interview meetings to maintain quality assurance.

Neuropsychological testers administered the MATRICS Cognitive Consensus Battery (MCCB: Nuechterlein and Green, 2006) to assess current cognitive functioning and the Wechsler Test of Adult Reading (WTAR: Wechsler, 2001) to estimate predicted premorbid intellectual functioning (i.e., the level of cognitive function an individual would have been expected to achieve without the disease). The Revised Chapman Physical (Chapman and Chapman, 1978) and Social Anhedonia (Eckblad et al., 1982) Scale questionnaires were used to assess self-reported anhedonia.

3. Results

3.1. Descriptive statistics and distribution of scores

Descriptive statistics for BNSS items and subscales are presented in Table 1 for Time 1 and Time 2 assessments. At both time points, the Alogia and Lack of Normal Distress subscales had skew >1.0, as did individual BNSS items for Intensity of Future Pleasure, Quantity of Speech, and Spontaneous Elaboration.

3.2. Internal consistency

Cronbach’s alpha, calculated to examine internal consistency, was 0.94, indicating that the items measure a single latent construct of negative symptoms. Item total correlations indicated that all BNSS items were significantly correlated with the BNSS total scale score (see Table 1). Furthermore, alpha if-item-deleted coefficients ranged from 0.93 to 0.95, suggesting no benefit from excluding any individual items.

3.3. Stability of measurement

Pearson correlations were calculated to estimate the stability of BNSS scores for 37 participants who were tested across two time points separated by 214 (88) days on average (range = 56 to 371 days). Patients were clinically stable at both time points. Results indicated good temporal stability for the BNSS total score (r = 0.93) and the 6 subscales. All individual items demonstrated good stability, with the exception of the Intensity of Future Pleasure item, which was substantially lower (see Table 1).

3.4. Discriminant validity

Discriminant validity was examined by evaluating the magnitude of correlations between the negative symptom scales (BNSS total score, SANS total score, and BPRS Negative Subscale) and BPRS Positive, Disorganized, and Total symptom scores (see Table 2). All 3 negative symptom scales had moderate relationships with the BPRS total, as would be expected. The BNSS, SANS, and BPRS Negative Subscale were not significantly correlated with BPRS Positive or Disorganized domains. The BPRS Depression item was not significantly correlated with the BNSS total score (r = 0.12, p = 0.82), anhedonia subscale (r = 0.15, p = 0.13), or individual anhedonia items (p’s > 0.12). The lack of significant correlations suggests that negative symptoms measured on the BNSS are not synonymous with depression, and that BNSS anhedonia items capture a form of affective disturbance that has little overlap with depression.

---

1 The 22-item version of the SANS developed in the CONSIST clinical trial (Cognitive and Negative Symptoms in Schizophrenia Trial) was used (Buchanan et al., 2007). Modifications implemented in the 22-item SANS include: 1) poverty of content of speech is not included in the subscale total score; 2) the avolition subscale is expanded by replacing impersistence in work or school with two items rating quality and level of role function, with appropriate alternate forms for inpatients and outpatients; 3) the asociality/anhedonia subscale is modified to directly rate anhedonia and asociality rather than rating recreational activities and relationships with friends and peers; and 4) attention items were not included.
90

sure of functional outcome.

1. Intensity of pleasure
2. Frequency of pleasure
3. Intensity of future pleasure

Subscale total

Distress subscale

4. Lack of normal distress

Asociality subscale

5. Asociality behavior
6. Asociality inner-experience

Subscale total

Avolition subscale

7. Avolition behavior
8. Avolition inner-experience

Subscale total

Blunted affect subscale

9. Facial expression
10. Vocal expression
11. Expressive gestures

Subscale total

Alogia subscale

12. Quantity of speech
13. Spontaneous elaboration

Subscale total

BNSS total score

The BNSS total score was significantly correlated with the Chapman Physical Anhedonia (PA) (r = 0.31, p < 0.01) and Social Anhedonia (SA) (r = 0.45, p < 0.001) Scales. The BNSS Intensity of Pleasure item significantly correlated with PA (r = 0.32, p < 0.01) and SA (r = 0.39, p < 0.001), as did the Frequency of Pleasure item (PA: r = 0.28, p < 0.01; PA: r = 0.41, p < 0.001). The Intensity of Future pleasure item was correlated with SA (r = 0.26, p < 0.02), but not PA (r = 0.15, p = 0.15).

The BNSS total score was significantly correlated with the MCCB total t-score, as well as the domain scores for Processing Speed, Attention/Vigilance, and Working Memory. Correlations between the BNSS total score and other MCCB domains were nonsignificant. The BPRS Negative subscale was not significantly correlated with any of the MCCB domains, and the SANS was correlated only with Processing Speed and Working Memory. The BNSS, SANS, and BPRS Negative factor were not significantly correlated with the WTAR (see Table 3).

### Table 3
Convergent Validity.

<table>
<thead>
<tr>
<th>Functional outcome</th>
<th>BNSS total</th>
<th>SANS total</th>
<th>BPRS Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOF total</td>
<td>−0.71***</td>
<td>−0.68***</td>
<td>−0.55***</td>
</tr>
</tbody>
</table>
| Premorbid and current cognition
| WTAR               | −0.02      | −0.03      | −0.04         |
| MCCB ProcSpd       | −0.32**    | −0.24*     | −0.05         |
| MCCB AttnVig       | −0.21†     | −0.11      | −0.17         |
| MCCB WM            | −0.36***   | −0.25*     | −0.18         |
| MCCB VerbLrn       | −0.18      | −0.15      | 0.04          |
| MCCB VisLrn        | −0.11      | −0.09      | −0.03         |
| MCCB ReaPS         | −0.14      | −0.10      | −0.07         |
| MCCB SocCog        | −0.05      | −0.01      | −0.02         |
| MCCB Overall       | −0.25†     | −0.17      | −0.09         |

Note: Values represent correlation coefficients (r-values) **p < 0.01; ***p < 0.001; WTAR = Wechsler Test of Adult Reading scaled score; MCCB ProcSpd = Processing Speed; MCCB AttnVig = Attention/Vigilance; MCCB WM = Working Memory; MCCB VerbLrn = Reasoning and Problem Solving; MCCB VisLrn = Visual Learning; MCCB ReaPS = Verbal Learning; MCCB SocCog = Social Cognition; MCCB Overall = Overall t-score.

### Table 2
Discriminant validity.

<table>
<thead>
<tr>
<th></th>
<th>BPRS Positive</th>
<th>BPRS disinorganization</th>
<th>BPRS total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNSS total</td>
<td>−0.06</td>
<td>0.04</td>
<td>0.32**</td>
</tr>
<tr>
<td>SANS total</td>
<td>−0.02</td>
<td>0.08</td>
<td>0.39**</td>
</tr>
<tr>
<td>BPRS Negative</td>
<td>−0.02</td>
<td>0.09</td>
<td>0.53***</td>
</tr>
</tbody>
</table>

Note: Values represent correlation coefficients (r-values) ***p < 0.01.
4. Discussion

4.1. Summary of psychometric findings

Reliability analyses indicated that the BNSS has excellent internal consistency and temporal stability. Item-total correlation analyses indicated that the intensity of future pleasure and distress items showed relatively low correlations with the total score. These items also demonstrated skewness $>1.0$, suggesting that relatively few patients were rated as having pathological scores. However, alpha-item deleted analyses suggested no benefit of excluding any individual BNSS items. Further studies should clarify this issue.

Discriminant validity was indicated by low correlations between the BNSS total score and the BPRS total score, as well as nonsignificant correlations with the BPRS Positive and Disorganized subscales. The BNSS total and anhedonia subscale total were not significantly correlated with BPRS depression, suggesting that these scores are not reflective of mood symptoms. Since extrapyramidal symptoms (EPS) were not assessed, future BNSS psychometric studies should include measures of EPS to estimate overlap with EPS induced bradykinesia. Given the high percentage of patients on a second generation antipsychotic in this sample, this may be less of an issue than in previous years.

Convergent validity was demonstrated by significant relationships between BNSS total scores and the SANS total score and BPRS Negative factor; these correlations were significant, but not so high as to suggest that the BNSS is redundant with these existing instruments. Subscale-level convergent validity was demonstrated by significant relationships with the four SANS subscales, which again were not high enough to suggest redundancy. Additionally, participants categorized as “deficit” on the SDS had more severe BNSS ratings than nondeficit patients. Future studies should also examine convergent validity of the BNSS in relation to other measures of negative symptoms, such as the NSA (Axelrod et al., 1993).

Convergent validity with social and vocational outcome was indicated by significant correlations with the LOF scale. The BNSS anhedonia subscale and individual items were also significantly correlated with the Chapman Physical and Social Anhedonia scales. We did not include a concurrent measure of anticipatory pleasure to examine the convergent validity of the BNSS intensity of future pleasure item, and this will be an important extension of the current study. With regard to convergent validity as measured by cognitive tests, one would expect current neuropsychological status to correlate with negative symptoms based upon the literature, especially the domains of processing speed working memory, and attention (see Harvey et al., 2006). However, measures of premorbid level of cognitive function, such as the WTAR, would not be expected to correlate with negative symptoms because they are “hold” tests which are relatively stable in the presence of cognitive decline/neurological insult and known to show minimal relationships with negative symptoms (Wexsler, 2001). Indeed, the BNSS, SANS, and BPRS negative factor were not significantly correlated with the WTAR. However, the BNSS total score was significantly correlated with the current neuropsychological status on the MCCB, including the MCCB total score, and domain scores for Processing Speed, Working Memory, and Attention/Vigilance. Correlations between the BNSS and these MCCB domains were somewhat higher than those observed with the SANS or the BPRS negative factor. Thus, the BNSS demonstrated good convergent validity with regard to its relationships with measures of current cognitive function.

4.2. Comparison of the BNSS and other negative symptom rating scales

The BNSS differs from existing measures (e.g., SANS, BPRS, PANSS, NSA, CAINS) (Overall and Gorham 1982; Andreasen, 1982; Kay et al., 1987; Axelrod et al., 1993; Blanchard et al., 2011) in several important ways. First, the BNSS is brief, consisting of only 13 items that can be rated in a 15-minute interview, or less if done in conjunction with other interview scales. In comparison, other instruments include more items (e.g., SANS ranges from 22 to 30 depending upon the version used) or longer interviews (e.g., CAINS, 45 min Blanchard et al., 2011). Second, the BNSS includes a manual with suggested questions, prompts, and a detailed discussion of procedures for rating individual items and performing the interview (BNSS manual, workbook, scoresheet available online with Kirkpatrick et al., 2011). Many scales do not include a manual (e.g., SANS, NSA), requiring investigators to create in-house procedures for ratings and interviewing, thereby making it difficult for different groups to administer similar interviews and establish inter-rater and cross-site reliability. Third, the BNSS manual is written at a reading level that is not advanced — it has clear, simple wording, especially in suggested probe questions. We have found that raters trained to use the BNSS find the manual, which is only 9 pages in length, easy to use and learn. As we had previously shown, we found that raters of all levels of experience are able to learn the BNSS and become reliable efficiently (see Kirkpatrick et al., 2011). Fourth, the BNSS covers the 5 Consensus Conference domains, and does not include items such as poverty of content of speech or inappropriate affect, which are not thought to be related to the negative symptom construct. Measures such as the SANS and NSA were not designed specifically to cover these content domains. Fifth, in two studies (Kirkpatrick et al., 2011; Strauss et al., in press), the BNSS has shown good separation of the 2 dimensions thought to underlie negative symptoms (i.e., Emotional Expressivity and Motivation-Pleasure); other instruments have produced less clean factor loadings or been less consistent in this regard (Blanchard and Cohen 2006; Horan et al., 2011).

The BNSS may also offer advantages at the level of individual items. For example, the BNSS includes separate items for internal-experience and outward behavior for the domains of avolition and asociaility. Other scales do not make this distinction between inner-experience and behavior (e.g., SANS, PANSS), or collapse these constructs into a single item (e.g., CAINS), making it impossible to assess discrepancies between inner-experience and behavior. We have demonstrated here some ways in which these discrepancies are important. However, there are many other possibilities for the importance of separating behavior and internal-experience. For example, the ability to assess these items separately may be beneficial for examining the efficacy of pharmacological or psychosocial interventions, which could conceivably affect a change in internal-experience prior to behavior. Such changes would be missed on an instrument that assesses only behavior or
collapses internal-experience and behavior into a single item. Similarly, having separate items for internal experience and outward behavior may be relevant for experimental psychopathology studies examining constructs such as reward learning and negative symptoms (e.g., Heerey and Gold, 2007; Barch and Dowd, 2010; Gold et al., 2008, 2012; Strauss et al., 2011a,b), which may be more strongly associated with the internal-experience construct. The BNSS also offers advances in the assessment of anhedonia. Several scales either do not assess anhedonia (e.g., BPRS, PANSS) or confute anhedonia and asociality (e.g., SANS). The 3 BNSS anhedonia items include coverage of multiple aspects of anhedonia, which have been identified in recent conceptualizations of the symptom (see Strauss and Gold, 2012), including items for retrospective pleasure, prospective pleasure, and frequency of pleasurable activities. The BNSS also includes a unique affective item measuring the lack of normal distress, which has been a core predictor of primary and enduring negative symptoms in the deficit schizophrenia subgroup (see Kirkpatrick et al., 2001; Kirkpatrick and Galderisi, 2008). The current study demonstrated that the BNSS Lack of Normal Distress item separated deficit and nondeficit schizophrenia patients categorized using the SDS, suggesting that it may have utility in identifying individuals with schizophrenia who present with primary negative symptoms.

Overall, given its conceptual and item-level advantages, brevity (13-items, 15-minute interview), and good psychometric characteristics, the BNSS can be considered a promising new instrument for use in clinical trials. Future studies are needed to determine whether the BNSS is sensitive to change in pharmacological trials.

Role of funding source
Research supported in part by the US National Institutes of Mental Health Grants: K23-MH092530 to Dr. Strauss; R01 MH080666 to Dr. Gold; PS0-MH082999 to Dr. Carpenter.

Contributors
Gregory Strauss, James Gold, Robert Buchanan, William Carpenter, and Brian Kirkpatrick designed the study. Statistical analysis and writing of the first draft of the manuscript was performed by Gregory Strauss and Adam Culbreth. Lauren Catalano and Adam Culbreth performed study-coordination and testing. All authors contributed to and have approved the final manuscript.

Conflict of interest
Dr. Carpenter served as a consultant for Eli Lilly, Lundbeck, Bristol Myers Squibb, Genentech, Shire Pharmaceuticals, Astra Zeneca, and Merck. Dr. Kirkpatrick has served as a consultant for Sunovion, Abbott, and Boehringer Ingelheim. Dr. Buchanan has served as consultant or on the advisory boards of Abbott, Astra Zeneca, Estrellas, Bristol-Meyer-Squibb, Cypress Bioscience, EnVivo, Janssen Pharmaceuticals, Inc., NuPathe, Inc., Pfizer, Roche, Schering-Plough, Solvay Pharmaceutical, Takeda, and Wyeth. Dr. Buchanan is also a DSMB member for Cephalon, Pfizer, and Otsuka. Dr. Gold has served as a consultant for Roche, Merck, Pfizer, Bristol Meyers Squibb, and Glaxo Smith Kline. Dr. McMahon has consulted for Agena. Drs. Buchanan and Fischer receive royalties from UpToDate Incorporated. Dr. Strauss, Dr. Koller, Mr. Culbreth, and Ms. Catalano have no financial relationships with commercial interests.

Acknowledgments
We would like to thank the subjects who participated in the study and staff at the Maryland Psychiatric Research Center who made the completion of the study possible. We are especially thankful to the staff at the Maryland Psychiatric Research Center, Jackie, Kiwanuka, Sharon August, Leeka Hubzin, and Jaime Brown who conducted subject recruitment and testing.

References