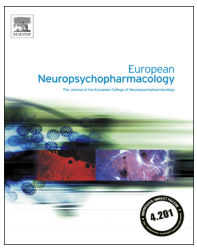
European Neuropsychopharmacology (2013) 23, 1051–1056

www.elsevier.com/locate/euroneuro



Psychometric analysis in support of shortening the Scale for the Assessment of Negative Symptoms

Stephen Z. Levine[a,*], Stefan Leucht[b]

[a]Department of Community Mental Health, University of Haifa, Haifa 31905, Israel
[b]Klinik und Poliklinik für Psychiatrie und Psychotherapie, Technische Universität München, München, Germany

Received 19 October 2012; received in revised form 13 November 2012; accepted 14 November 2012

KEYWORDS
Psychometrics;
Schizophrenia;
Symptom severity;
Predominant negative symptoms;
Scale for the Assessment of Negative Symptoms;
Clinical trial

Abstract
Despite recent emphasis on the measurement and treatment of negative symptoms, studies of the Scale for the Assessment of Negative Symptoms (SANS) identify different symptom clusters, offer mixed support for its psychometric properties and suggest that it is shortened. The current study objective is to examine the psychometric properties of the SANS and the feasibility of a short research version of the SANS. Data were re-analyzed from three clinical trials that compared placebo and amisulpride to 60 days. Participants had chronic schizophrenia and predominantly negative symptoms (n=487). Baseline data were examined with exploratory factor analysis and Item Response Theory (IRT) to identify a short SANS. The short and original SANS were compared: with confirmatory factor analysis at endpoint; and on symptom response with mixed modeling to compare. Results showed that at baseline the SANS consisted of three factors labeled Affective-flattening, Asociality and Alogia-inattentiveness. IRT suggested a short SANS with 11 items and 3 response options. Comparisons of the original and short SANS showed: the short version was a better fit to the data based on confirmatory factor analysis at endpoint; similar significant ($p<.001$) correlations between the baseline and subsequent scores; similar reliability; and similar significance ($p<.05$) on response based on mixed modeling. It is concluded that a short SANS is feasible to assess predominantly negative symptoms in chronic schizophrenia in research settings.

1. Introduction

Negative symptoms are integral to schizophrenia, and date to Kraepelin's early description of the destruction of the personality (Kraepelin, 1919). Negative symptoms are associated with deficits in cognitive, social and real-world functioning (Bowie et al., 2006; Harvey et al., 2006; Kirkpatrick et al., 2006). Meta-analysis reports that second-generation antipsychotic medications are effective in the treatment of the positive and not negative symptoms of schizophrenia (Leucht et al., 2009). There are currently intensive initiatives that aim to address the antipsychotic treatment of negative symptoms. For instance, new medications are being developed to treat negative symptoms (e.g., SPD489, asenapine). Also, a NIMH-MATRICS expert consensus group has been formed on negative

*Corresponding author. Tel.: +972 524 896 083; fax: +972 3 761 7374.
E-mail address: slevine@univ.haifa.ac.il (S.Z. Levine).

0924-977X/$ - see front matter © 2012 Elsevier B.V. and ECNP. All rights reserved.
http://dx.doi.org/10.1016/j.euroneuro.2012.11.008

symptoms with participants from academia, the Federal Drug Administration and industry (Alphs, 2006; Kirkpatrick et al., 2006; Kirkpatrick and Fischer, 2006; Marder et al., 2011). The consensus group highlighted methodological and assessment problems in clinical trials of negative symptoms (Kirkpatrick et al., 2006).

One of the most widely used (Kirkpatrick et al., 2006) measures to assess negative symptoms is the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983). It consists of five a priori symptom factors (i.e., symptom clusters) of affective flattening, alogia, avolition, anhedonia symptoms. A recent NIMH consensus statement on negative symptoms noted that the SANS is appropriate to use in clinical trials of negative symptoms but requires revision since empirically the symptom factors are inconsistently identified and it lacks validity (Kirkpatrick et al., 2006).

To examine the psychometric properties of the SANS, studies use factor analysis to identify factors (i.e., symptom clusters that covary). Cross-sectional studies of the SANS identify five (Peralta et al., 1995), four (Rabany et al., 2011), and three-different (Keefe et al., 1992; Mueser et al., 1994; Sayers et al., 1996) symptom factors. Longitudinal research into the SANS has replicated three factors (diminished expression, inattention-alogia and social amotivation) at a month after index hospitalization and again after 3–6 months (Kelley et al., 1999). Reliability (i.e., the accuracy of measurement; e.g., Cronbach's alpha) is another psychometric property studies report. Generally, reported reliabilities range from high (over .9) to modest (around .6) (Andreasen et al., 1995; Gorsuch, 1993; Keefe et al., 1992; Kline, 1993; Mueser et al., 1994; Nunnally and Bernstein, 1978; Peralta et al., 1995; Rabany et al., 2011; Sayers et al., 1996). Reliability of .6 means that 60% of the SANS reflects the "true score" (i.e., symptom severity), and the remaining 40% is attributable measurement error. High measurement error reduces accuracy when predictive methods are used (e.g., prediction of treatment response). Some items reduce reliability (e.g., "grooming and hygiene") (Mueser et al., 1994; Rabany et al., 2011), or contribute little to the symptom factors ("poor eye contact") (Rabany et al., 2011).

Existing SANS studies use classical psychometric theory that has been extended by Item Response Theory (IRT; as summarized in Table 1; Baker, 2001; Embretson and Reise, 2000; Reise and Henson, 2003). A series of initiatives (e.g., the NIH Patient-Reported Outcomes Measurement Information System; Cella et al., 2007) encourage the use of advanced psychometric item modeling in particular IRT to examine outcomes (Reeve et al., 2007; Teresi and Fleishman, 2007). In clinical trials of cancer IRT has been used to shorten patient reported outcomes (Garcia et al., 2007). In psychiatric epidemiology it has been used, for example, to reduce the length of the Center for Epidemiologic Studies-Depression scale (Cole et al., 2004). In clinical trials of schizophrenia, reanalysis of clinical trial data using IRT has shown that a PANSS reduced in length and with fewer rating options is feasible (Khan et al., 2011; Levine et al., 2011; Santor et al., 2007). IRT can identify overlapping items, and superfluous rating options. Thus it has the potential to reduce the number of items, and response options of measures without compromising the reliability and validity of a measure. Also, IRT identifies reliability at different severity levels for each symptom and symptom cluster. For instance, IRT may acknowledge the possibility

Table 1 The terminology of Item Response Theory.

Term	Definition	SANS relevance
Key terms		
Classical test theory	Traditional psychometric methods (e.g., Cronbach's α)	Assesses syndromes not symptoms
Polychotomous responses	Measures with more than two response options These are observed responses	E.g., "Absent", "Mild"
Item	A question is a measure	The SANS questions E.g., inappropriate affect
Theta θ	Latent construct assessed with a scale	Symptom severity is latent, whereas SANS ratings are observed
Assessment method: parameter estimates		
Discrimination parameter (α)	Index of the strength the relationship between item and the latent construct it aims to measure	Acknowledges the disparity between the hallucinations item assessment and latent hallucination severity
Threshold parameter (β)—difficulty	Parameter to index the severity of an item response along the latent (θ) continuum of item response categories	Acknowledges the disparity between the severity rating (e.g., "Severe") and latent symptom severity
Assessment method: pictured graphically		
Information curves	Indexes items or scales values over latent severity (θ) to assess reliability at different severity levels	E.g., SANS items or scales *may* be less reliable if latent symptom severity is high
Item characteristic curve	Probabilistic index of the relationship between ratings on each category for an item and level on the latent construct (θ)	E.g. examines the distinction between "severity" and "Extreme severity" according to latent severity