Schizophrenia Bulletin Open
doi:10.1093/schizbullopen/sgab013

What is the Best Latent Structure of Negative Symptoms in Schizophrenia? A Systematic Review

Bernardo Haguiara[1,], Gabriela Koga[1], Elton Diniz[1,2], Lais Fonseca[1,2], Cinthia H Higuchi[1,], Simão Kagan[1], Acioly Lacerda[1,2], Christoph U Correll[3–5], and Ary Gadelha[*,1,2]

[1]Laboratory of Integrative Neurosciences, Department of Psychiatry, Escola Paulista de Medicina/Universidade Federal de São Paulo, SP, Brazil; [2]Schizophrenia Program, Department of Psychiatry, Escola Paulista de Medicina/Universidade Federal de São Paulo, SP, Brazil; [3]Department of Psychiatry, The Zucker Hillside Hospital, Northwell Health, Glen Oaks, NY, USA; [4]Department of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [5]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany

*To whom correspondence should be addressed; 241 Major Maragliano Street - Vila Mariana 04017-030, São Paulo, SP, Brazil. tel: +551134662170; e-mail: aryararipe@gmail.com

Bernardo Haguiara and Gabriela Koga are co-first authors.

Background: **Determining the best latent structure of negative symptoms in schizophrenia could benefit assessment tools, neurobiological research, and targeted interventions. However, no review systematically evaluated studies that assessed and validated latent models of negative symptoms.** ***Objective:*** **To identify and evaluate existing latent structure models in the literature of negative symptoms and to determine the best model.** ***Method:*** **Systematic search of MEDLINE, EMBASE, and Scopus on July 19, 2020, for confirmatory factor analysis models of negative symptoms in patients with schizophrenia. The available evidence was assessed through 2 sets of criteria: (1) study design quality—based on negative symptoms assessment and modeling strategy and (2) psychometric quality and model fit—based on fit indices and factor definition quality.** ***Results:*** **In total, 22 studies (*n* = 17 086) from 9 countries were included. Studies differed greatly regarding symptom scales, setting, and sample size (range = 86–6889). Dimensional models included 2–6 factors (median = 4). Twelve studies evaluated competing models and adopted appropriate instruments to assess the latent structure of negative symptoms. The 5-factor and hierarchical models outperformed unitary, 2-factor, and 3-factor models on all direct comparisons, and most of the analyses derived from the Brief Negative Symptom Scale. Considering the quality criteria proposed, 5-factor and hierarchical models achieved excellent fit in just one study.** ***Conclusions:*** **Our review points out that the 5-factor and hierarchical models represent the best latent structure of negative symptoms, but the immaturity of the relevant current literature may affect the robustness of this conclusion. Future studies should address current limitations regarding psychometric properties and also address biological and clinical validity to refine available models.**

Key words: schizophrenia/negative symptoms/latent structure/factor analysis/systematic review

Introduction

Negative symptoms represent core features of schizophrenia since the first descriptions of the disorder[1,2] and refer to a diminution or absence of expected behaviors and inner experiences related to motivation, interest, or expression.[3–7] Although a major contributor to poor real-life functioning in people with schizophrenia,[3,8,9] no specific approved treatment exists for negative symptoms,[10,11] which is explained both by the insufficient knowledge about its neurobiology and by the challenging assessment of these symptoms.

To address such issues, the National Institute of Mental Health (NIMH) organized, in 2005, the NIMH-MATRICS (Measurement and Treatment Research to Improve Cognition in Schizophrenia) Consensus Statement on Negative Symptoms.[12] It acknowledged 5 constructs within the negative domain: blunted affect (decrease in the observed expression of emotion, ie, facial and vocal expression, and expressive gestures),[13,14] alogia (reduction in the quantity of speech and in its spontaneous elaboration)[4], anhedonia (diminished capacity to experience pleasant emotions)[15], asociality (reduction in social initiative due to decreased interest in forming close relationships with others)[16], and avolition (reduced initiation and persistence of goal-directed activity).[17]

Despite improvements that allowed the development of more sophisticated tools, obtaining accurate and reliable measurements of negative symptoms remains challenging due, in part, to the overlap between the symptoms that make up this dimension. A particular negative symptom may be represented by more than one item of a scale, eg, affective flattening could influence simultaneously social withdrawal and speech spontaneity, thus being accounted for multiple times. In this scenario, factor analysis methods are useful, as they reduce measurement errors by generating latent structure models that present independent unobservable variables (factors) to explain correlations between observed variables (here, the items of a scale).[18] Two possible approaches to factor analysis are exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

EFA is a data reduction technique that infers the presence of latent factors responsible for shared variance among a set of items. It can be useful to generate hypotheses about latent structures; however, EFAs do not specify an underlying structure for the observed variables but rather assume that each item could be related to each latent factor, not ensuring a true validation. EFAs using negative symptom scale items alone find support for 2 dimensions that reflect motivation and pleasure (MAP) and expression (EXP),[16,19–22] but CFA models find that 1- and 2-factor solutions usually offer a poor fit for the data.[23,24]

CFA allows testing a priori hypothesis and objective comparison with other theoretical models. This way, CFA provides more accurate conclusions about latent models, being more appropriate to evaluate the latent structure of negative symptoms. Recently, the evidence supporting the Consensus 5-factor model has been reviewed,[25] and results from CFA studies using mainly the Brief Negative Symptom Scale (BNSS) have consistently shown the robustness of the 5-factor and hierarchical models over the 2-factor or unitary models. Despite that, no study has systematically reviewed and evaluated all the literature covering CFA studies that investigated latent models for negative symptoms, even though nonsystematic reviews on the topic are available.[4,19,25–28]

Thus, we performed a systematic review of the literature searching for latent structure models of negative symptoms in schizophrenia and aimed to identify the best latent structure. To define the best model, we assessed the quality of each study, psychometric parameters, and direct comparisons between models obtained from the same sample.

Methods

Search Strategy

We conducted a systematic review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyzes (PRISMA)[29,30] guidelines. Searching MEDLINE, EMBASE, and Scopus from database inception until July 19, 2020, we used the following search terms: *(((((((“schizophrenia spectrum and other psychotic disorders”[MeSH Terms])) OR “schizophrenic disorders”) OR schizophrenia)) AND (((“negative symptoms”) OR (“negative symptom scale”)))) AND (((((((“confirmatory factor analysis”) OR “latent structure analysis”) OR “latent structure”) OR “latent model”) OR “validation study”)))))*. We also reviewed manually reference lists of included articles and relevant reviews.

Inclusion and Exclusion Criteria

We included original, peer-reviewed articles without language restriction and excluded posters, letters, editorials, and reviews. We included samples with individuals aged ≥16 years old meeting diagnostic criteria for schizophrenia according to *Diagnostic and Statistical Manual of Mental Disorders* (DSM)-III,[31] DSM-IV,[32] or DSM-V,[33] although the addition of non-affective psychotic disorders was not a reason for exclusion if most of the subjects (>50%) had a schizophrenia diagnosis. We selected studies that performed CFA to evaluate the latent structure of negative symptoms. We excluded studies with EFA and principal component analysis (PCA) since such approaches do not test theoretical models of latent factors.[18]

Data Collection

Two authors (G.K. and B.H.) independently screened all retrieved records by title and abstract and assessed selected studies in full for eligibility. Any conflicts were resolved by consensus or consulting another investigator (E.D.). Additionally, the following information was extracted independently (G.K. and B.H.) from each included study: metadata, sample characteristics, intraclass correlation coefficient (ICC), negative symptom rating scales, model origin, estimation method, competing models and discriminative criteria employed, the best models, fit indices, and shortest number of items with loadings higher than 0.5 per factor.

Quality Assessment

Quality assessment was based on study design, psychometrics, and model fit.

Study Design. The study design adequacy was evaluated according to the reliability of the assessment of negative symptoms; appropriateness of the adopted instrument to assess the latent structure of negative symptoms; and modeling strategy, which encompassed model origin, estimation method employed for the CFA, and use of competing model’s approach.

Small samples produce less stable factor analysis solutions.[34] However, there is no consensus regarding the minimum sample size for CFAs. This issue cannot be

determined by rules of thumb[35,36] and would only be adequately tested if done via sensitivity power analyses for each included study,[35,36] which is far from the scope of the present study. Therefore, sample size was not included in the quality assessment.

Assessment reliability of negative symptoms was evaluated through ICC. Based on previous studies,[37–39] we considered assessments with ICC ≤ 0.5 (or with ICC not informed/performed) as "hardly reliable," 0.5 < ICC < 0.8 as "reliable," and ICC ≥ 0.8 as "highly reliable."

Regarding instruments appropriateness, the Positive and Negative Syndrome Scale (PANSS) and Negative Symptom Assessment (NSA) lack the items needed to cover all constructs determined by the NIMH-MATRICS Consensus and, so, cannot properly assess the latent structure of negative symptoms.[23,40] On the other hand, the BNSS, the Clinical Assessment Interview for Negative Symptoms (CAINS), and the Scale for the Assessment of Negative Symptoms (SANS) possess the items to cover the Consensus' domains, making them more suitable to answer our research question.[23,40]

EFA or PCA should be followed by CFA using a different sample.[41–43] Therefore, studies that performed EFA/PCA to obtain the tested model and used the same sample to perform the CFA and studies that performed CFA based on theoretical models not preceded by EFA/PCA were considered "methodologically limited."

The selection of an estimator must be based on distributional patterns of the data and assumptions, which makes appropriate to report and justify its choice, while relying on statistical software's default settings is not advisable.[18] Thus, studies that provided no information about the estimation method employed were also considered "methodologically limited."

The last step of quality assessment of study design consisted of verifying whether the best model presented had been compared with other models and, if so, which were the competing models. The strongest test of a proposed model is to identify and test competing models that represent truly different, but highly plausible, hypothesized structural relationships.[36] Using the same sample to test competing theories on the latent structure of negative symptoms provides more robust evidence than testing a single isolated model. Thus, studies that compared different models obtained from the same sample were well regarded, whereas the test of a single model reduced the quality of evidence. Studies that tested models with the same number of factors obtained by different items of the same scale were penalized, as well as studies that compared a model exclusively with the null model, since it is widely accepted that negative symptoms are a separate factor in schizophrenia.[4,12,28]

Psychometrics and Model Fit. Psychometric quality and model fit were assessed considering the following parameters: descriptive fit indices—comparative fit index (CFI), non-normed fit index (NNFI), normed fit index (NFI), root mean square error of approximation (RMSEA), weighted root mean square residual (WRMR), or goodness-of-fit index (GFI) —and factor definition quality. We also evaluated information criteria—Akaike Information Criterion (AIC) and/or Bayesian Information Criterion (BIC) —when studies compared non-nested models, with the lowest values being used to determine optimal model fit.[44] Chi-square values were not considered to evaluate model fit due to its high sensitivity to sample size as well as the ratio χ^2/df.[45,46]

The following standards for appropriate fit indices were considered:[47–49] GFI > 0.90, CFI > 0.95, NFI > 0.95, NNFI > 0.95, RMSEA < 0.06, and WRMR < 1.00. Models presenting no index with adequate fit were considered to have "poor fit"; models with at least one adequate fit index were classified as "acceptable"; and models showing all indices with adequate fit were considered "excellent."

Factor definition quality was based on the number of items with loadings > 0.5[36] per factor. Factors with less than 3 or 4 items per factor require larger sample sizes and are more likely to provide unstable solutions.[18,50–54] On the other hand, there is support for the use of the few best indicators for the development of theoretically sophisticated models.[55] Thus, we considered factors defined by a single item as "poorly defined," factors defined by 2 indicators as "acceptable," and factors defined by at least 3 items as "well defined." Models obtained through an EFA in a different sample and that used the EFA's factor loadings to define factors, instead of CFA's, had their quality lowered.

Additional details of quality assessment are provided in the supplementary material.

Determining the Best Latent Structure

Following the quality assessment, we selected the studies that evaluated competing models derived from the SANS, BNSS, and CAINS. After that, we compiled the models tested in a single table to verify which models performed better according to each study's criteria. Finally, we evaluated the best models according to the predefined fit quality criteria.

Results

Our search yielded 1680 records found in the database search and 9 through other sources (figure 1). After removing duplicates, 1495 records titles and abstracts were screened, out of which 69 were selected for full-text reading, resulting in 22 included studies in our qualitative analysis.

We extracted data from 26 latent models of negative symptoms. Table 1 summarizes the included studies, with 5 published between 1992 and 1996 and 17 between 2013

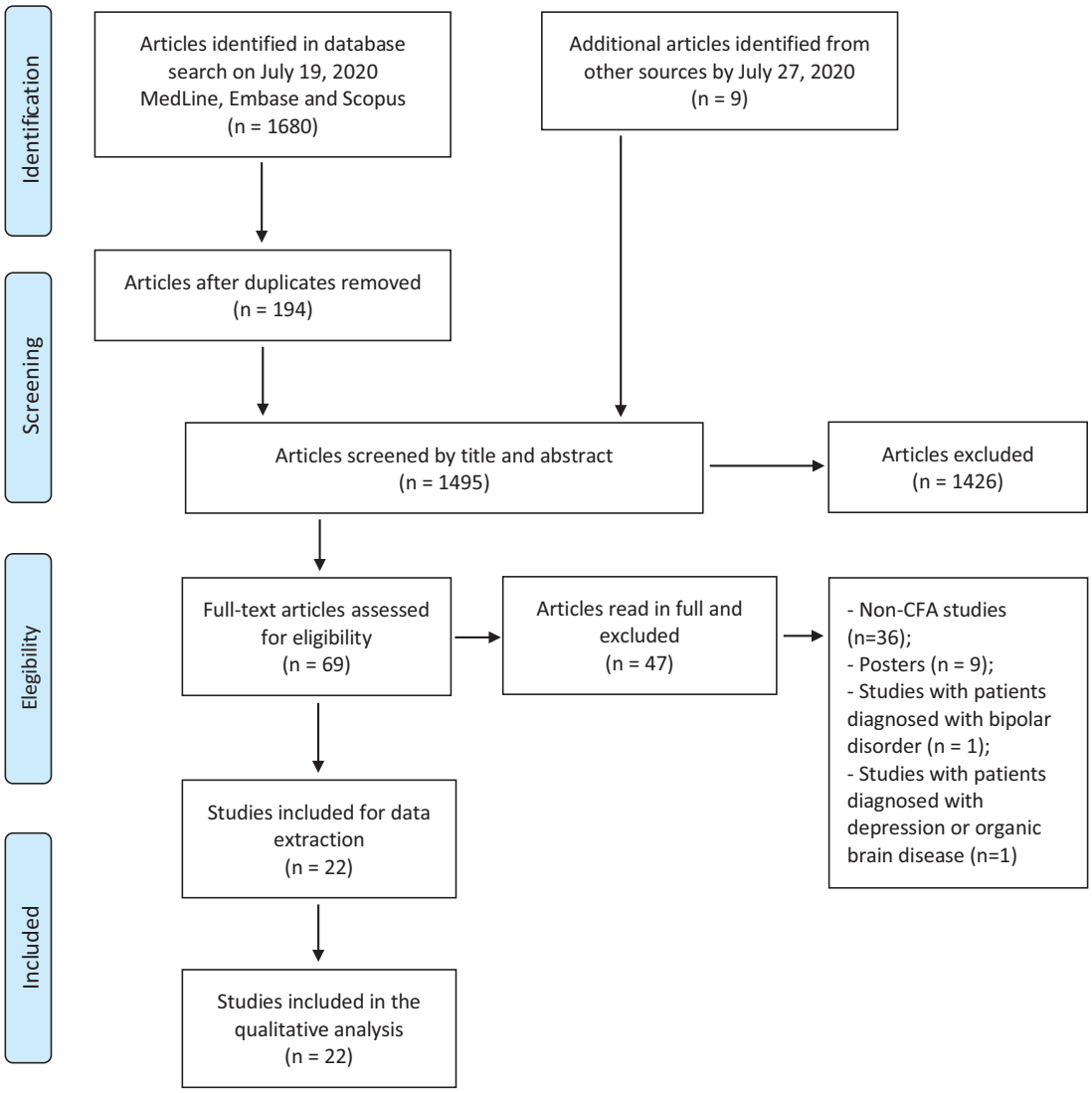

Fig. 1. Review flow diagram.

and 2020. Altogether, 17 086 individuals aged 16–79 years were included (range *n* = 86–6889), from 9 different countries (United States = 7, Singapore = 4, South Korea = 3, China = 2, Netherlands = 2, Germany = 1, Israel = 1, Italy = 1, and Spain = 1). Samples included chronic inpatients,[56–59] outpatients,[23,60–64] and patients in mixed settings,[24,65–74] mainly diagnosed not only with schizophrenia but also with schizoaffective disorder, schizophreniform disorder, delusional disorder, and psychosis not otherwise specified.

Study Design

Negative symptom assessment was hardly reliable in 7 studies,[56,57,59,61,67,70,73] reliable in 2 studies,[68,71] and highly reliable in 13 studies.[23,24,58,60,62–66,69,72,74,75]

Negative symptom scales varied greatly. Four models were based on the NSA,[56,57,63,66] 7 on the BNSS,[23,24,62,64,73] and the others were equally divided between the SANS,[23,58,59,65,67] the PANSS,[60,61,68,70,71] and the CAINS.[23,69,72,74,75]

Regarding models origins, most of the studies based the CFA on a PCA, EFA, and/or CFA from previous studies;[56,61,62,64,65,68–70,73–75] 5 studies obtained the CFA model from a PCA or EFA performed in a different sample,[24,60,63,66,72] while 3 studies used the same sample of the CFA;[59,67,71] and 3 studies performed the CFA based on a theoretical model not preceded by a PCA or EFA.[23,57,58]

The most employed estimation method was the maximum likelihood,[56–58,60,68,73,74] followed by a combination of robust maximum likelihood and robust weighted least squares,[23,24] and the robust weighted least squares alone.[61] Twelve studies did not provide information about the estimator.[59,62–67,69–72,75]

The competing model's approach was adopted by 18 studies. The most used discriminative criteria were based on the analysis of descriptive fit indices and information criteria,[23,24,64,67,73,75] followed by the chi-square

Table 1. Summary of the Main Findings

		Study Design							
				Assessment		Modeling Strategy			Model Fit/Psychometrics	
Study (Year)	Best Model	CFA Sample Size		Reliability of Assessment	Appropriateness of the Instrument	Model Origin	Estimator	Competing Models**	Fit Indices	Shortest Number of Items With Loadings > 0.5 per Factor
Keefe et al (1992)[65]	1. Diminished expression 2. Social dysfunction 3. Disorganization	130		ICC > 0.8	Original: SANS-20 Tested: SANS-13	PCA from previous study [81]	Not informed	Criteria: χ^2 difference test + GFI $M_0 \times M_1 \times M_2 \times \mathbf{M_3}$	$\chi^2 = 151.38$ *df* = 62 GFI = 0.865	2
Axelrod et al (1993)[57]	1. Communication 2. Emotion/affect 3. Social involvement 4. Motivation. 5. Retardation	223	276 (cross-validation sample)	ICC not informed/ performed	Original: NSA-25 Tested: NSA-16	Theoretical (NSA-25 original structure)	ML	Criteria χ^2 difference test + descriptive fit indices: $M_0 \times M_1 \times M_2 \times M_3 \times M_4 \times \mathbf{M_5}$	χ^2:*df* = 2.76 NFI = 0.88 NNFI = 0.89 CFI = 0.92 Cross-validation χ^2:*df* = 2.76 NFI = 0.89 NNFI = 0.91 CFI = 0.93	Unknown
Axelrod et al (1994)[56]	1. Communication 2. Emotion/affect 3. Social involvement 4. Motivation 5. Gross cognition 6. Retardation	223		ICC not informed/ performed	Original: NSA-26 Tested: NSA-25	PCA from previous study[82]	ML	Criteria: χ^2 difference test + descriptive fit indices $M_0 \times M_1 \times M_5 \times \mathbf{M_6}$	X^2:*df* = 2.5 AIC = 160 Satorra-Bentler χ^2:*df* = 2.2 NFI = 0.81 NNFI = 0.85 CFI = 0.87	3 or more
Peralta and Cuesta (1995)[58]	1. Affective flattening 2. Alogia 3. Avolition-apathy 4. Anhedonia-asociality 5. Inattention	253		ICC > 0.8	Original: SANS-20 Tested: SANS-19	Theoretical (SANS original structure)	ML	1) Criteria: χ^2 difference $M_0 \times \mathbf{M_5}$ (SANS-19) 2) Criteria: NNFI difference: M_3 (SANS-20) $\times \mathbf{M_5}$ (SANS-19)	$\chi^2 = 471$ *df* = 142 GFI = 0.828 NNFI = 0.849	2
Sayers et al (1996)[59]	1. Diminished expression 2. Social amotivation 3. Inattention-alogia	401		ICC not informed/ performed	Original: SANS-19 Tested: SANS-18	EFA + CFA with same sample	Not informed	Criteria: Factor definition quality $\mathbf{M_3}$ **(SANS-18)** $\times M_5$ (SANS-19)	$\chi^2 = 250.58$ *df* = 129 (*P* < .001) NNFI = 0.95 CFI = 0.96 $\chi^2 = 263.32$ *df* = 139 (*P* < .001) NNFI = 0.95 CFI = 0.96	3 or more (EFA factor loadings)
Levine and Leucht (2013)[67]	1. Affective-flattening 2. Asociality 3. Alogia-inattentiveness	487		ICC not informed / performed	Original: SANS-20 Tested: SANS-11	PCA using same sample	Not informed	Criteria: Descriptive fit indices + information criteria $\mathbf{M_3}$ **(SANS-11)** $\times M_3$ (SANS-20)	$\chi^2 = 81.61$ *df* = 41 (*P* < .01) AIC = 131.61 (CI: 0.27–0.40, 95%) GFI = 0.92 CFI = 0.93 NFI = 0.91 NNFI = 0.90 SRMR = 0.07 RMSEA = 0.05 (CI: 0.03-0.06, 95%)	2
Liemburg et al (2013)[60]	1. Core negative symptoms 2. Social emotive withdrawal	2172	845 423	ICC > 0.8	Original: PANSS Tested: PANSS (items N1, N2, N3, N4, N6, G5, G7, G13, and G16)	PCA using different sample	ML	No competing models	Diagnosis: non-affective psychotic disorder CFI = 1.0 GFI = 1.0 RMSEA = 0.015 RMR = 0.01 Diagnosis: schizophrenia CFI = 1.0 GFI = 1.0 RMSEA = 0.0082 RMR = 0.019 50% of the SCZ sample with the highest summed negative symptoms CFI = 0.98 GFI = 0.99 RMSEA = 0.052 RMR = 0.038	3 or more
Jang et al (2016)[68]	1. Expressive deficits 2. Experiential deficits	220		0.5 < ICC < 0.8	Original: PANSS Tested: PANSS (items N1, N2, N3, N4, N6, G7, and G16)	PCA from previous study[83]	ML	Criteria: χ^2 difference test $M_0 \times \mathbf{M_2}$	$\chi^2 = 28.887$ *df* = 11 (*P* < .01) CFI = 0.981 NNFI = 0.964 SRMR = 0.026 RMSEA = 0.086 (CI: 0.048-0.125, 90%)	3 or more
Jung et al (2016)[69]	1. Motivation and pleasure 2. Expression	119		ICC > 0.8	Original: CAINS Tested: K-CAINS	PCAs from previous studies[76,84]	Not informed	No competing models	$\chi^2 = 212.776$ *df* = 64 (*P* < .001) AIC = 292.776 CFI = 0.863 NFI = 0.820 NNFI = 0.806 IFI = 0.867 RMSEA = 0.140	3 or more

Table 1. Continued

		Study Design															
					Assessment		Modeling Strategy			Model Fit/Psychometrics							
Study (Year)	Best Model	CFA Sample Size			Reliability of Assessment	Appropriateness of the Instrument	Model Origin	Estimator	Competing Models**	Fit Indices			Shortest Number of Items With Loadings > 0.5 per Factor				
Lim et al (2016)[70]	1. Diminished emotional expression 2. Social avolition	687			ICC not informed/performed	Original: PANSS Tested: PANSS (items N1, N2, N3, N4, N6, G7, G13, and G16)	CFA from previous study[85]; PCAs from previous studies[60,83]	Not informed	Criteria: Descriptive fit indices $M_1 \times \mathbf{M_2}$	CFI = 0.978 NNFI = 0.967 WRMR=0.928 RMSEA = 0.077 (CI: 0.062-0.093, 90)			3 or more				
Stiekema et al (2016)[61]	1. Social amotivation 2. Expressive deficits	1157			ICC not informed/performed	Original: PANSS Tested: PANSS (items N1, N2, N3, N4, N6, G5, G7, G13, and G16)	PCA and CFA from previous study[60]	WLSMV	No competing models	CFI = 0.99 NNFI = 0.98 WRMR= 0.86 RMSEA = 0.06 (CI: 0.05-0.07, 90%)			3 or more				
Khan et al (2017)[71]	1. Expressive deficits 2. Experiential deficits	6689			0.5 < ICC < 0.8	Original: PANSS Tested: PANSS (items N1, N2, N3, N4, N6, G7, and G16)	EFA using same sample	Not informed	Criteria: χ^2 difference test + descriptive fit indices $M_1 \times \mathbf{M_2}$	$\chi^2 = 99.215$ $P < .001$ CFI = 0.923 NNFI = 0.913 RMSEA = 0.045 GFI = 0.906			3 or more				
Strauss et al (2018)[23]	NIMH 5-factor model	BNSS 192	CAINS 400	SANS 268	ICC > 0.8	BNSS, CAINS, and SANS	Theoretical	WLSMV and MLR	Criteria: descriptive fit indices + information criteria SANS/BNSS/CAINS: $M_1 \times M_2 \times \mathbf{M_5} \times M_H$	BNSS $\chi^2 = 39.87$ $df = 19$ $P < .01$ AIC = 5401.90 BIC = 5708.10 aBIC = 5410.34 CFI = 0.994 NNFI = 0.997 RMSEA = 0.076 WRMR = 0.43	CAINS $\chi^2 = 76.52$ $df = 19$ $P < .01$ AIC = 12 986.03 BIC = 13 285.39 aBIC = 13 047.41 CFI = 0.986 NNFI = 0.988 RMSEA = 0.077 WRMR = 0.89	SANS $\chi^2 = 161.51$ $df = 46$ $P < .001$ AIC = 11 331.08 BIC = 11 721.27 aBIC =11 375.68 CFI = 0.949 NNFI = 0.972 RMSEA = 0.097 WRMR = 1.00	BNSS 2	CAINS 1	SANS 2
Xie et al (2018)[75]	1. Motivation and pleasure 2. Expression	185			ICC > 0.8	CAINS	PCAs from previous studies[16,76]	Not informed	Criteria: descriptive fit indices + information criteria $M_1 \times \mathbf{M_2}$ **(Kring et al, 2013)** $\times M_3$ (Chan et al, 2015)	$\chi^2 = 85.02$ $df = 56$ AIC = 155.02 CFI = 0.99 NNFI = 0.99 RMSEA = 0.053			3 or more				
Ahmed et al (2019)[24]	NIMH 5-factor model	1681			ICC > 0.8	BNSS	EFA using different sample	WLSMV and MLR	1) Calibration sample Criteria: descriptive fit indices + information criteria $M_1 \times M_2 \times M_3 \times \mathbf{M_5} \times \mathbf{M_H}$ 2) Cross-cultural validation in 5 samples Criteria: descriptive fit indices + information criteria $\mathbf{M_5} \times M_H$ 3) Assessment of invariance of $\mathbf{M_5}$ across multinational samples Criteria: descriptive fit indices **configural invariance** model × metric invariance × scalar invariance × residual invariance	$\chi^2 = 252.36$ $df = 87$ $P < .0001$ CFI = 0.996 NNFI = 0.999 RMSEA = 0.074			2 (EFA factor loadings)				

Table 1. Continued

			Study Design						
			Assessment		Modeling Strategy			Model Fit/Psychometrics	
Study (Year)	Best Model	CFA Sample Size	Reliability of Assessment	Appropriateness of the Instrument	Model Origin	Estimator	Competing Models**	Fit Indices	Shortest Number of Items With Loadings > 0.5 per Factor
Ang et al (2019)[62]	Hierarchical model NIMH 5-factor model	274	ICC > 0.8	BNSS	EFA and CFA from previous study[24]	Not informed	Criteria: descriptive fit indices M_2 v $\mathbf{M_5 \times M_H}$	Hierarchical model: CFI = 0.999 NNFI = 0.998 WRMR = 0.512 RMSEA = 0.035 (CI = 0.000-0.056) / 5-factor model: CFI = 0.998 NNFI = 0.997 WRMR = 0.453 RMSEA = 0.043 (CI = 0.019-0.064)	2
Mucci et al (2019)[73]	Hierarchical model NIHM 5-factor model	249	ICC not informed/ performed	BNSS	EFA and CFAs from previous studies[23,24]	ML	Criteria: descriptive fit indices + information criteria $M_1 \times M_2 \times \mathbf{M_5 \times M_H}$	Hierarchical model: χ^2 = 116.1 *df* = 48 AIC = 200.083 BIC = 204.730 CFI = 0.974 NNFI = 0.957 RMSEA = 0.076 / 5-factor model: χ^2 = 108.1 *df* = 44 AIC = 200.100 BIC = 205.190 CFI = 0.975 NNFI = 0.956 RMSEA = 0.077	2
Rekhi et al (2019)[72]	1..MAP social 2. MAP vocational 3. MAP recreational 4. EXP	141	ICC > 0.8	CAINS	EFA using different sample	Not informed	Criteria: descriptive fit indices $M_2 \times \mathbf{M_4}$	CFI = 0.955 NNFI = 0.940 WRMR = 0.793 RMSEA = 0.078	1
Rekhi et al (2019)[63]	1. Restricted speech 2. Poor quality speech 3. Affective blunting 4. Amotivation	141	ICC > 0.8	Original: NSA-16 Tested: NSA-12	EFA using different sample	Not informed	Criteria: Descriptive fit indices M_5 (NSA-16: Axelrod et al, 1993b) × $\mathbf{M_4}$ **(NSA-12)**	CFI = 0.981 NNFI = 0.973 RMSEA = 0.064 (CI: 0.034-0.090, 90%)	2
Richter et al (2019)[74]	1. Motivation and pleasure 2. Expression	105	ICC > 0.8	CAINS	PCAs from previous studies[16,84]	ML	No competing models	χ^2 = 93.28 *df* = 61 AIC = 153.28 CFI = 0.95 RMSEA = 0.07 (CI: 0.04–0.10, 90%)	3 or more
Huang et al (2020)[66]	1. Communication 2. Emotion 3. Motivation	86	ICC > 0.8	Original: NSA-16 Tested: NSA-15	PCA using different sample	Not informed	Criteria: descriptive fit indices M_5 (NSA-16: Axelrod et al, 1993) × $\mathbf{M_3}$ (NSA-15)	χ^2: *df* = 1.468 CFI = 0.948 RMSEA = 0.074	2
Jeakal et al (2020)[64]	Hierarchical model	173	ICC > 0.8	Original: BNSS Tested: K-BNSS	EFA and CFAs from previous studies[23,24]	Not informed	Criteria: Descriptive fit indices + information criteria $M_1 \times M_2 \times M_3 \times \mathbf{M_5 \times M_H}$	χ^2 = 87.28 *df* = 48 *P* < .001 AIC = 5,724.46 BIC = 5856.90 aBIC = 5,723.91 CFI = 0.972 NNFI = 0.961 SRMR = 0.045 RMSEA = 0.077	2

Table 1. Continued

		Study Design							
			Assessment		Modeling Strategy			Model Fit/Psychometrics	
Study (Year)	Best Model	CFA Sample Size	Reliability of Assessment	Appropriateness of the Instrument	Model Origin	Estimator	Competing Models**	Fit Indices	Shortest Number of Items With Loadings > 0.5 per Factor
Total 22 studies (United states = 7, Singapore = 4, South Korea = 3, China = 2, Netherlands = 2, Germany = 1, Israel = 1, Italy = 1, and Spain = 1)	Range: 2-6 (median = 4) PANSS 2-factor model: 5 studies BNSS 5-factor model: 4 studies BNSS Hierarchical model: 3 studies CAINS 2-factor model: 3 studies SANS 3-factor model: 3 studies SANS 5-factor model: 2 studies CAINS 4-factor model: 1 study CAINS 5-factor model: 1 study NSA 3-factor model: 1 study NSA 4-factor model: 1 study NSA 5-factor model: 1 study NSA 6-factor model: 1 study	17 086 subjects, mostly diagnosed with schizophrenia	ICC > 0.8: 13 studies ICC not informed/performed: 7 studies 0.5 < ICC < 0.8: 2 studies	BNSS: 5 studies CAINS: 5 studies PANSS: 5 studies SANS: 5 studies NSA: 4 studies	PCA/EFA/CFAs from previous studies: 11 studies PCA/EFA using different sample: 5 studies PCA/EFA/CFA using same sample: 3 studies Theoretical: 3 studies	Not informed: 12 studies ML: 7 studies MLR + WLSMV: 2 studies WLSMV: 1 study	Descriptive indices + information criteria: 5 studies χ^2 difference test + descriptive fit indices: 5 studies Descriptive fit indices: 5 studies No competing models: 4 studies Factor definition quality: 1 study Competing models with null model: 1 study Models with the same number of factors derived from different items of the same scale: 1 study	CFI: 20 studies (14 with adequate fit) Range: 0.87-1.0 (median = 0.975). NNFI: 18 studies (11 with adequate fit) Range: 0.806-0.999 (median = 0.961). RMSEA: 17 studies (6 with adequate fit) Range: 0.0082-0.140 (median = 0.074). WRMR: 5 studies (all with adequate fit) Range: 0.43-1.0 (median = 0.827). GFI: 5 studies (3 with adequate fit) Range: 0.828-1 (median = 0.92). NFI: 4 studies (none with adequate fit) Range: 0.81-0.91 (median = 0.88).	Well-defined: 9 models Adequate: 14 models Poorly defined: 3 models

Note: AIC, Akaike information criterion; aBIC, adjusted Bayesian information criterion; BIC, Bayesian information criterion; aBIC: BNSS, Brief Negative Symptom Scale; CAINS, Clinical Assessment Interview for Negative Symptoms; CAINS-item 1, Motivation for Close Family/Spouse/Partner Relationships; CAINS-item 2, Motivation for Close Friendships/Romantic Relationships; CAINS-item 5, Motivation for Work and School Activities; CAINS-item 6, Expected Pleasurable Work and School Activities—Next Week; CFA, confirmatory factor analysis; ECVI, expected cross-validation index; CFI, comparative fit index; EXP, Expression; G5, Mannerisms and Posturing; G7, Motor Retardation; G13, Disturbance of Volition; G16, Active Social Avoidance; GFI, goodness-of-fit index; ICC, intraclass correlation coefficient; IFI, incremental fit index; K-CAINS, Korean version of the Clinical Assessment Interview for Negative Symptoms; N1, Blunted Affect; N2, Emotional Withdrawal; N3, Poor Rapport; N6, Lack of Spontaneity and Flow of Conversation; MAP, Motivation-pleasure; ML, maximum likelihood; MLR, robust maximum likelihood; NFI, normed fit index; NIMH, National Institute of Mental Health; NNFI, non-normed fit index; NOS, non otherwise specified; NSA, Negative Symptom Assessment; NSA-item 6, Affect: reduced modulation of intensity; NSA-item 16, Slowed movements; PANSS, Positive and Negative Syndrome Scale; PCA, principal component analysis; RMR, root mean square residual; RMSEA, root mean square error; SANS, Scale for the Assessment of Negative Symptoms; SANS-item 4, Poor eye contact; SANS-item 6, Inappropriate affect; SANS-item 9, Poverty of speech; SANS-item 10, Poverty of Content of Speech; SANS-item 11, Blocking; SANS-item 14, Grooming and Hygiene; SANS-item 19, Sexual activity; SANS-item 23, Social Inattentiveness; SRMR, standardized root mean square residual; WLSMV, weighted least square mean and variance; WRMR, weighted root mean square.
**Models in bold outperformed other models.

difference test combined with the analysis of descriptive fit indices,[56–58,65,71] the analysis only of descriptive fit indices,[62,63,66,70,72] and factor definition quality.[59]

Psychometrics and Model Fit

Dimensional models ranged from 2[60,61,68–71,74,75] to 6[56] dimensions (median = 4). Two and 5-factor[23,24,57,58,62,73] models were the most depicted among studies, followed by hierarchical,[62,64,73] 3-factor,[59,65–67] and 4-factor[63,72] structures. The hierarchical model consists of 2 second-order factors reflecting EXP and MAP and 5 first-order factors reflecting the domains of the NIMH consensus development conference.[12]

We observed high heterogeneity in terms of models and factors composition. For example, none of the studies that evaluated the 3-factor model using the SANS presented a common factor composed of the same indicators. In addition, none of the models derived from the NSA presented the same number of factors. On the other hand, models obtained from the PANSS, BNSS, and CAINS showed greater uniformity.

Fit indices used and their values varied. To assess model fit, older studies used the chi-square test and few descriptive fit indices, whereas more recent articles used modern fit measures and information criteria. The summary of descriptive fit indices is presented in table 2. Regarding factor definition quality, 9 models were considered well-defined,[56,60,61,68–71,74,75] 14 were adequate,[23,24,58,59,62–67,73] and 3 were poorly defined.[23,57,72]

Six models presented poor fit,[56–58,65,66,69] 14 presented acceptable fit,[23,24,63,64,67,68,70–74] and 6 models met requirements for excellent fit.[59–62,75]

The Best Latent Structure of Negative Symptoms

Table 3 summarizes the studies that used the SANS, BNSS, or CAINS to assess competing models of negative symptoms. The 5-factor and hierarchical models based on the NIMH-MATRICS Consensus outperformed other models in all 5 comparisons using the BNSS and in a single comparison both with the CAINS and the SANS.

Among these studies that directly compared the Consensus 5-factor and hierarchical models with other models, only Ang et al[62] obtained models with excellent fit according to the fit criteria proposed, whereas the others presented models with acceptable fit. Ang et al[62] used the BNSS and, based on models generated by Strauss et al[23] and Ahmed et al[24], concluded that the hierarchical model and the 5-factor model outperformed the 2-factor model, with the hierarchical model having an advantage over the 5-factor model. Regarding the studies with acceptable fit, according to the predefined criteria, the superiority of the hierarchical model in relation to the 5-factor model was not unanimous.

After contrasting the best 2 models, Strauss et al[23] concluded that the 5-factor model obtained a better fit both with the BNSS and the CAINS, based on the information criteria, but found that the hierarchical model performed slightly better with the SANS. Ahmed et al[24] considered both models adequate, based on descriptive fit indices and information criteria, but extended their analyzes of cultural invariance of the BNSS exclusively to the 5-factor model. Mucci et al[73] and Jeakal et al[64] also used the BNSS and considered both models equally adequate, but Jeakal et al[64] only provided the data about the hierarchical model.

Regarding the studies that did not directly compared the Consensus 5-factor and hierarchical models with other models, only Sayers et al[59] and Xie et al[75] presented models with excellent fit. Sayers et al[59] obtained excellent fit both with a 3-factor and a 5-factor SANS model but considered the latter less adequate because it had a factor composed of 2 items. Xie et al[75] used the CAINS and concluded that the 2-factor model by Kring et al[16] outperformed the unitary model and the 2-factor model by Chan et al[76].

Discussion

We analyzed 22 publications addressing the dimensionality of negative schizophrenia symptoms, 12 of which used the SANS, the BNSS, or the CAINS to assess competing models of negative symptoms. The Consensus 5-factor and hierarchical models outperformed unitary, 2-factor, and 3-factor models on direct comparisons, and our results suggest that they currently represent the

Table 2. Summary of Fit Indices

Fit Index	Range	Median	Reference	Used (*n*° of Studies)	Adequate (*n*° of Studies)	Adequacy Rate (%)
CFI	0.870-1.0	0.975	>0.95	20	14[23,24,59–64,68,70,72–75]	70.0
NNFI	0.806-0.999	0.961	>0.95	19	11[23,24,59,61–64,68,70,73,75]	61.1
RMSEA	0.0082-0.140	0.074	<0.06	17	6[60–62,67,71,75]	35.3
WRMR	0.430–1.0	0.827	<1.0	5	5[23,61,62,70,72]	100.0
GFI	0.828-1.0	0.920	>0.9	5	3[61,67,71]	60.0
NFI	0.810-0.910	0.880	>0.95	4	0	0

Note: CFI, confirmatory factor analysis; GFI, goodness-of-fit index; NFI, normed fit index; NNFI, non-normed fit index; RMSEA, root mean square error of approximation; WRMR, weighted root mean square.

Table 3. Competing models of the SANS, CAINS, and BNSS

Scale	Study (Year)	Model						
		M_0	M_1	M_2	M_3	M_4	M_5	M_h
SANS	Keefe et al (1992)[65]	×	×	×	•			
	Peralta & Cuesta (1995)[58]	×			×		•*	
	Sayers et al (1996)[59]				○		○*	
	Strauss et al (2018)[23]		×	×			•	•
BNSS	Strauss et al (2018)[23]		×	×			•	•
	Ahmed et al (2019)[24]		×	×	×		•	•
	Ang et al (2019)[62]			×			○	○
	Mucci et al (2019)[73]		×	×			•	•
	Jeakal et al (2020)[64]		×	×	×		•	•
CAINS	Strauss et al (2018)[23]		×	×			•	•
	Xie et al (2018)[75]		×	○				
	Rekhi et al (2019)[72]			×		•		

Note: BNSS, Brief Negative Symptom Scale; CAINS, Clinical Assessment Interview for Negative Symptoms; M_0, null model; M_1, 1-factor model; M_2, 2-factor model; M_3, 3-factor model; M_4, 4-factor model; M_5, 5-factor model; M_h, hierarchical model; SANS, Scale for the Assessment of Negative Symptoms. Green models ("•") outperformed red models ("×") according to each study's fit criteria; "○," model with excellent fit according to the criteria proposed; "*," 5-factor model not based on the NIMH-MATRICS Consensus Statement on Negative Symptoms.

best latent structure of negative symptoms. However, the reduced number of available studies may affect the robustness of this conclusion.

Twelve studies directly compared models using appropriate instruments to assess the latent structure of negative symptoms, of which only 5 compared the Consensus 5-factor and hierarchical structure to other models, resulting in 5 comparative analysis derived from the BNSS, 1 from the CAINS, and 1 from the SANS. Of note, Strauss et al[23] assessed competing models from the BNSS, CAINS, and SANS using a different sample for each instrument. Peralta and Cuesta[58] and Sayers et al[59] also tested SANS 5-factor models, but both studies presented latent structures with an inattention factor (no longer included in the negative dimension) and grouped anhedonia and asociality into a single factor. Moreover, among the 5-factor and hierarchical Consensus-based models, only those obtained by Ang et al[62] showed excellent fit according to the proposed criteria. In sum, direct comparisons favored the Consensus 5-factor and hierarchical model, but psychometric properties are limited.

Indeed, few studies fulfilled all the standard fit requirements proposed, and the RMSEA was the fit index with the lowest adequacy. Since it assesses how well a model fits in the population,[18] the RMSEA determines how a proposed model is apart from a perfect model.[77] It represents one of the most informative fit indices[78] for being sensitive to the number of estimated parameters in the model and relatively little influenced by sample size.[18] RMSEA seems to improve—indicating better fit—as more variables are added to the model, which means that the lower the number of items per factor, the worse the RMSEA.[79] As most of the instruments have few items to assess each factor, the psychometric criteria did not fully favor the 5-factor and hierarchical model, which should not be considered a criterion for rejecting these models but rather an example of the contrast between requirements for good fit and available instruments. In fact, the only fit index that separated the models obtained by Ahmed et al,[24] Strauss et al,[23] Mucci et al,[73] and Jeakal et al[64] from an excellent fit, according to the proposed quality criteria, was the RMSEA, whose cutoff value of 0.08 is also accepted by some authors.[18,80]

Despite the superiority of the 5-factor and hierarchical models over the 2-factor model in studies that performed comparative analyzes, the neurobiological and clinical evidence for specific constructs of the 5-factor model is still limited, although promising. Previous investigations regarding the underlying neurobiology of negative symptoms may have failed to find neural correlates for the 5 domains due to assessments based on scores of constructs with fewer factors.[25] Similarly, past unsuccessful attempts to promote targeted treatments for negative symptoms may, in fact, have been effective for 1 or more of the 5 domains, with positive results being masked by assessments not specific enough for each domain.[25] Further studies could fill these gaps assessing negative symptoms with greater granularity, which could, in turn, provide domain-specific therapeutic and neurobiological advances.

Although the present study strengthens the thesis that negative symptoms should be evaluated with greater granularity matching clinical experience, conclusions derived from only 5 studies. Therefore, the field may benefit from additional studies investigating unanswered questions: the 5-factor/hierarchical models are invariant across different disease stages or over time? Other 5-factor models could fit similarly or best? Can a multilevel structure affect the results?

Psychometrics, clinical experience, and neurobiology have convergences and tensions. No single source of evidence can be sufficient. The balance between seeking psychometric and clinical relevance has been especially controversial. We acknowledge that the rejection of models solely based on one single fit index can be excessive. However, merely disregarding fit indices as inadequate can overlook opportunities for new research questions. In the end, we believe that data and clinical utility provide the best guides to the current discussion. In this sense, our results identify some caveats when framing available evidence on the dimensional nature of negative symptoms.

Future investigations of latent structures for negative symptoms should adopt strict procedures regarding the study design, such as calculating the minimum sample size, using modern instruments (BNSS and CAINS), ensuring the highest possible reliability for the assessment of negative symptoms by performing and reporting the ICC, not using the CFA sample in a previous EFA or PCA, reporting the estimation method and justifying its use, and adopting the competing models' approach. Additionally, we encourage future studies to provide comparative analysis between the CAINS 2-factor, 5-factor, and hierarchical models and, ideally, also between the BNSS and the CAINS using samples jointly assessed by these instruments.

The limitations of this review include the high heterogeneity among studies' samples and designs (eg, assessment tools and modeling strategies), the small number of studies that analyzed competing latent models of negative symptoms using appropriate instruments, and the lack of a validated method to measure the quality of factor analysis studies. Nonetheless, we operationalized a unified quality assessment according to criteria previously recommended by methodological papers, which enabled a systematic comparison among the included studies. We adopted standard and widely used cutoff points but that may be considered overly rigorous and not completely unanimous in literature. Psychometrics represents an important perspective to validate a construct but does not provide the final answer alone. Thus, we added additional perspectives, ie, the study design and instruments used, to bring a comprehensive view of the field.

Overall, we conclude that the 5-factor and hierarchical models are currently the best conceptualization of negative symptoms. We believe that these results may guide future psychometric studies and facilitate the search for biological and clinical validity of the negative dimension in schizophrenia.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin Open* online.

Acknowledgments

The authors would like to thank the team of the Interdisciplinary Laboratory of Integrative Neurosciences (Linc) for their support and participation in this research. The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Dr Correll has been a consultant and/or advisor to or has received honoraria from Acadia, Alkermes, Allergan, Angelini, Axsome, Gedeon Richter, Gerson Lehrman Group, IntraCellular Therapies, Janssen/J&J, LB Pharma, Lundbeck, MedAvante-ProPhase, Medscape, Neurocrine, Noven, Otsuka, Pfizer, Recordati, Rovi, Sumitomo Dainippon, Sunovion, Supernus, Takeda, and Teva. He has provided expert testimony for Janssen and Otsuka. He served on a Data Safety Monitoring Board for Lundbeck, Rovi, Supernus, and Teva. He received grant support from the Berlin Institute of Health (BIH), Janssen, the National Institute of Mental Health (NIMH), Patient Centered Outcomes Research Institute (PCORI), Takeda, and the Thrasher Foundation. He received royalties from UpToDate and is also a stock option holder of LB Pharma. Dr Lacerda reports personal fees from Daiichi Sankyo Brasil, Pfizer, Janssen Pharmaceutical, Cristalia Produtos Químicos e Farmacêuticos, Mantecorp Indústria Química e Farmacêutica, Sanofi-Aventis, Libbs Farmacêutica, and Eli Lilly and grants from Janssen Pharmaceutical, H Lundbeck A/S, Cristalia Produtos Químicos e Farmacêuticos, Eli Lilly, H Lundbeck A/S, Servier Laboratories, FQM Farma, Hoffman-La Roche, and Forum Pharmaceuticals, outside the submitted work. Dr Gadelha has been a speaker and/or advisor to or has received honoraria from Janssen, Daiichi-Sankyo, Aché, Cristália, and Torrent.

References

1. Bleuler E. *Dementia Praecox or the Group of Schizophrenias*, translated by Zinkin J. New York, NY: International Universities Press; 1950.
2. Kraepelin E. *Dementia praecox and paraphrenia (1919)*, translated by Barclay RM. Huntingdon. New York, NY: Robert E Publishing Co.; 1971.
3. Correll CU, Schooler NR. Negative symptoms in schizophrenia: a review and clinical guide for recognition, assessment, and treatment. *Neuropsychiatr Dis Treat*. 2020;16:519–534.
4. Marder SR, Galderisi S. The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry*. 2017;16(1):14–24.
5. Brian Kirkpatrick BK. Developing concepts in negative symptoms: primary vs secondary and apathy vs expression. *J Clin Psychiatry*. 2014;75:3–7.
6. Fervaha G, Foussias G, Agid O, Remington G. Amotivation and functional outcomes in early schizophrenia. *Psychiatry Res*. 2013;210(2):665–668.

7. Konstantakopoulos G, Ploumpidis D, Oulis P, *et al*. Apathy, cognitive deficits and functional impairment in schizophrenia. *Schizophr Res*. 2011;133(1–3):193–198.
8. Harvey PD, Strassnig M. Predicting the severity of everyday functional disability in people with schizophrenia: cognitive deficits, functional capacity, symptoms, and health status. *World Psychiatry*. 2012;11(2):73–79.
9. Galderisi S, Rossi A, Rocca P, *et al*.; Italian Network For Research on Psychoses. The influence of illness-related variables, personal resources and context-related factors on real-life functioning of people with schizophrenia. *World Psychiatry*. 2014;13(3):275–287.
10. Correll CU, Rubio JM, Inczedy-Farkas G, Birnbaum ML, Kane JM, Leucht S. Efficacy of 42 pharmacologic cotreatment strategies added to antipsychotic monotherapy in schizophrenia: systematic overview and quality appraisal of the meta-analytic evidence. *JAMA Psychiatry*. 2017;74(7):675–684.
11. Krogmann A, Peters L, von Hardenberg L, Bödeker K, Nöhles VB, Correll CU. Keeping up with the therapeutic advances in schizophrenia: a review of novel and emerging pharmacological entities. *CNS Spectr*. 2019;24(S1):38–69.
12. Kirkpatrick B, Fenton WS, Carpenter WT Jr, Marder SR. The NIMH-MATRICS consensus statement on negative symptoms. *Schizophr Bull*. 2006;32(2):214–219.
13. Henry JD, Green MJ, de Lucia A, Restuccia C, McDonald S, O'Donnell M. Emotion dysregulation in schizophrenia: reduced amplification of emotional expression is associated with emotional blunting. *Schizophr Res*. 2007;95(1–3):197–204.
14. Kirkpatrick B. Progress in the study of negative symptoms. *Schizophr Bull*. 2014;40(Suppl 2):S101–S106.
15. Meehl PE. Hedonic capacity: some conjectures. *Bull Menninger Clin*. 1975;39(4):295–307.
16. Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. *Am J Psychiatry*. 2013;170(2):165–172.
17. Foussias G, Remington G. Negative symptoms in schizophrenia: avolition and Occam's razor. *Schizophr Bull*. 2010;36(2):359–369.
18. Brown TA. *Confirmatory Factor Analysis for Applied Research*. 2nd ed. New York, NY: The Guilford Press; 2015.
19. Blanchard JJ, Cohen AS. The structure of negative symptoms within schizophrenia: implications for assessment. *Schizophr Bull*. 2006;32(2):238–245.
20. Kirkpatrick B, Strauss GP, Nguyen L, *et al*. The Brief Negative Symptom Scale: psychometric properties. *Schizophr Bull*. 2011;37(2):300–305.
21. Mucci A, Galderisi S, Merlotti E, *et al*.; Italian Network for Research on Psychoses. The Brief Negative Symptom Scale (BNSS): independent validation in a large sample of Italian patients with schizophrenia. *Eur Psychiatry*. 2015;30(5):641–647.
22. Bischof M, Obermann C, Hartmann MN, *et al*. The Brief Negative Symptom Scale: validation of the German translation and convergent validity with self-rated anhedonia and observer-rated apathy. *BMC Psychiatry*. 2016;16(1):415.
23. Strauss GP, Nuñez A, Ahmed AO, *et al*. The latent structure of negative symptoms in schizophrenia. *JAMA Psychiatry*. 2018;75(12):1271–1279.
24. Ahmed AO, Kirkpatrick B, Galderisi S, *et al*. Cross-cultural validation of the 5-factor structure of negative symptoms in schizophrenia. *Schizophr Bull*. 2019;45(2):305–314.
25. Strauss GP, Ahmed AO, Young JW, Kirkpatrick B. Reconsidering the latent structure of negative symptoms in schizophrenia: a review of evidence supporting the 5 consensus domains. *Schizophr Bull*. 2019;45(4):725–729.
26. Peralta V, Cuesta MJ. How many and which are the psychopathological dimensions in schizophrenia? Issues influencing their ascertainment. *Schizophr Res*. 2001;49(3):269–285.
27. Galderisi S, Mucci A, Buchanan RW, Arango C. Negative symptoms of schizophrenia: new developments and unanswered research questions. *Lancet Psychiatry*. 2018;5(8):664–677.
28. Blanchard JJ, Kring AM, Horan WP, Gur R. Toward the next generation of negative symptom assessments: the collaboration to advance negative symptom assessment in schizophrenia. *Schizophr Bull*. 2011;37(2):291–299.
29. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–269, W64.
30. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62(10):e1–34.
31. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Publishing; 1980.
32. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Publishing; 1994.
33. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Publishing; 2013.
34. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychol Methods*. 1999;4:84–99.
35. Kyriazos TA. Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*. 2018;9:2207–2230.
36. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis, Andover*. Hampshire, United Kingdom: Cengage Learning; 2019.
37. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420–428.
38. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1:30–46.
39. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River, NJ: Pearson; 2000.
40. Marder SR, Kirkpatrick B. Defining and measuring negative symptoms of schizophrenia in clinical trials. *Eur Neuropsychopharmacol*. 2014;24(5):737–743.
41. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Ass Res Eval*. 2005;10(1):7.
42. Henson RK, Roberts JK. Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educ Psychol Meas*. 2006;66:393–416.
43. Worthington RL, Whittaker TA. Scale development research: a content analysis and recommendations for best practices. *Couns. Psychol*. 2006;34:806–838.
44. Akaike H. Factor analysis and AIC. *Psychometrika*. 1987;52:317–332.

45. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res Online*. 2003;8:23–74.
46. Bollen KA. *Structural Equations with Latent Variables*. New York, NY: Wiley; 1989.
47. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107(2):238–246.
48. Hu LT, Bentler PM. Evaluating model fit. In: Hoyle RH, ed. *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks, CA: Sage Publications. Inc; 1995:76–99.
49. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model A Multidiscip J*. 1999;6:1–55.
50. Marsh HW, Hau KT, Balla JR, Grayson D. Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behav Res*. 1998;33(2):181–220.
51. Boomsma A, Hoogland JJ. The robustness of LISREL modeling revisited. *Struct Equ Model Present Futur*. 2001:139–168.
52. Thompson B. *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. Washington, DC: American Psychological Association; 2004.
53. Dimitrov DM. *Statistical Methods for Validation of Assessment Scale Data in Counseling and Related Fields*. Alexandria, VA: American Counseling Association; 2012.
54. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. Boston, MA: Pearson Education, Inc.; 2013.
55. Hayduk LA, Littvay L. Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Med Res Methodol*. 2012;12:159.
56. Axelrod BN, Goldman RS, Woodard JL, Alphs LD. Factor structure of the negative symptom assessment. *Psychiatry Res*. 1994;52(2):173–179.
57. Axelrod BN, Goldman RS, Alphs LD. Validation of the 16-item negative symptom assessment. *J Psychiatr Res*. 1993;27(3):253–258.
58. Peralta V, Cuesta MJ. Negative symptoms in schizophrenia: a confirmatory factor analysis of competing models. *Am J Psychiatry*. 1995;152(10):1450–1457.
59. Sayers SL, Curran PJ, Mueser KT. Factor structure and construct validity of the Scale for the Assessment of Negative Symptoms. *Psychol. Assess*. 1996;8:269–280.
60. Liemburg E, Castelein S, Stewart R, van der Gaag M, Aleman A, Knegtering H; Genetic Risk and Outcome of Psychosis (GROUP) Investigators. Two subdomains of negative symptoms in psychotic disorders: established and confirmed in two large cohorts. *J Psychiatr Res*. 2013;47(6):718–725.
61. Stiekema APM, *et al*. Confirmatory factor analysis and differential relationships of the two subdomains of negative symptoms in chronically ill psychotic patients. *PLoS One*. 2016;11:e0149785.
62. Ang MS, Rekhi G, Lee J. Validation of the Brief Negative Symptom Scale and its association with functioning. *Schizophr Res*. 2019;208:97–104.
63. Rekhi G, Alphs L, Ang MS, Lee J. Clinical utility of the Negative Symptom Assessment-16 in individuals with schizophrenia. *Eur Neuropsychopharmacol*. 2019;29(12):1433–1441.
64. Jeakal E, Park K, Lee E, Strauss GP, Choi KH. Validation of the Brief Negative Symptom Scale in Korean patients with schizophrenia. *Asia Pac Psychiatry*. 2020;12(3):e12382.
65. Keefe RS, Harvey PD, Lenzenweger MF, *et al*. Empirical assessment of the factorial structure of clinical symptoms in schizophrenia: negative symptoms. *Psychiatry Res*. 1992;44(2):153–165.
66. Huang BJ, Wang Y, Miao Q, Yu X, Pu CC, Shi C. Validation of the Chinese version of the 16-item negative symptom assessment. *Neuropsychiatr Dis Treat*. 2020;16:1113–1120.
67. Levine SZ, Leucht S. Psychometric analysis in support of shortening the Scale for the Assessment of Negative Symptoms. *Eur Neuropsychopharmacol*. 2013;23(9):1051–1056.
68. Jang SK, Choi HI, Park S, *et al*. A two-factor model better explains heterogeneity in negative symptoms: evidence from the Positive and Negative Syndrome Scale. *Front Psychol*. 2016;7:707.
69. Jung SI, Woo J, Kim YT, Kwak SG. Validation of the Korean-version of the Clinical Assessment Interview for Negative Symptoms of Schizophrenia (CAINS). *J Korean Med Sci*. 2016;31(7):1114–1120.
70. Lim J, Lee SA, Lam M, *et al*. The relationship between negative symptom subdomains and cognition. *Psychol Med*. 2016;46(10):2169–2177.
71. Khan A, Liharska L, Harvey PD, Atkins A, Ulshen D, Keefe RSE. Negative symptom dimensions of the Positive and Negative Syndrome Scale across geographical regions: implications for social, linguistic, and cultural consistency. *Innov Clin Neurosci*. 2017;14(11–12):30–40.
72. Rekhi G, Ang MS, Yuen CKY, Ng WY, Lee J. Assessing negative symptoms in schizophrenia: validity of the clinical assessment interview for negative symptoms in Singapore. *Schizophr Res*. 2019;206:177–182.
73. Mucci A, Vignapiano A, Bitter I, *et al*. A large European, multicenter, multinational validation study of the Brief Negative Symptom Scale. *Eur Neuropsychopharmacol*. 2019;29(8):947–959.
74. Richter J, Hesse K, Schreiber L, *et al*. Evidence for two distinct domains of negative symptoms: confirming the factorial structure of the CAINS. *Psychiatry Res*. 2019;271:693–701.
75. Xie DJ, Shi HS, Lui SSY, *et al*. Cross cultural validation and extension of the Clinical Assessment Interview for Negative Symptoms (CAINS) in the Chinese context: evidence from a spectrum perspective. *Schizophr Bull*. 2018;44(suppl_2):S547–S555.
76. Chan RCK, *et al*. Validation of the Chinese version of the Clinical Assessment Interview for Negative Symptoms (CAINS): a preliminary report. *Front. Psychol*. 2015:6:1–5.
77. Xia Y, Yang Y. RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: the story they tell depends on the estimation methods. *Behav Res Methods*. 2019;51(1):409–428.
78. Diamantopoulos A, Siguaw JA. *Introducing LISREL: A Guide for the Uninitiated*. Thousand Oaks, CA: SAGE Publications; 2000.
79. Kenny DA, McCoach DB. Effect of the number of variables on measures of fit in structural equation modeling. *Struct Equ Model A Multidiscip. J*. 2003;10:333–351.
80. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociol Methods Res*. 1992;21:230–258.

81. Liddle PF. The symptoms of chronic schizophrenia: a re-examination of the positive-negative dichotomy. *Brit J Psychiat*. 1987;151(2):145–151.
82. Alphs LD, Summerfelt A, Lann H, Muller RJ. The negative symptom assessment: a new instrument to assess negative symptoms of schizophrenia. *Psychopharmacol Bull*. 1989;25:159–163.
83. Fervaha G, Foussias G, Agid O, Remington G. Motivational and neurocognitive deficits are central to the prediction of longitudinal functional outcome in schizophrenia. *Acta Psychiatr Scand*. 2014;130:290–299.
84. Engel M, Fritzsche A, Lincoln TM. Validation of the German version of the clinical assessment interview for negative symptoms (CAINS). *Psychiatry Res*. 2014;220: 659–663.
85. Jiang J, Sim K, Lee J. Validated five-factor model of positive and negative syndrome scale for schizophrenia in Chinese population. *Schizophr Res*. 2013;143(1):38–43.