

## Factor Analysis of Scale for Assessment of Negative Symptoms using SAS Software

Ben Adeyi, Shire Pharmaceuticals, Chesterbrook, PA

David Shen, Independent Consultant, Chesterbrook, PA

### ABSTRACT

The objective is to study the psychometric properties of the SANS (Scale for the Assessment of Negative Symptoms) using factor analysis, and investigate the feasibility of shortening the current version of the SANS. SANS data were examined with exploratory factor analysis to identify the principle components of SANS. The results showed that the SANS consisted of three factors, suggested a short SANS with 3 response cluster options of 10 items. Confirmatory factor analysis demonstrated the reliability of the short version of SANS.

### INTRODUCTION

Negative Symptoms are associated with deficits in cognitive, social and real-world functioning, and were used for the early description of the destruction of the personality. A consensus group, formed on negative Symptoms with participants from FDA, industry and academia, highlighted methodology and assessment problems in clinical trials of negative symptoms. One of the most widely used measures to assess negative symptoms is the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983).

Currently available treatment options have been shown to have only modest impact on negative and cognitive symptoms. There are intensive initiatives that aim to address the antipsychotic treatment of negative symptoms. For instance, new medications are being developed to treat negative symptoms such as SPD489. Some reports in the literature suggest psychostimulant-based pharmacological therapy may be a viable option for the treatment of negative symptoms and cognitive impairment (Barch et al. 2005).

The SANS is a 20-minute clinician-administered scale that assesses 5 symptom complexes to rate the negative symptoms of subjects. These five domains consist of

- Affective flattening or blunting
- Alogia
- Avolition-apathy
- Anhedonia-asociality
- Disturbance of attention

A recent consensus statement on negative symptoms noted that the SANS is appropriate to use in clinical trials of negative symptoms but empirically the symptom factors are inconsistently identified (Kirkpatrick et al., 2006).

Factor analysis can identify over-lapping 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. In psychiatric epidemiology it has been used to reduce the length of the Epidemiologic Studies-Depression scale. In the study of schizophrenia, reanalysis of clinical trial data using factor analysis has shown that a PANSS (Positive and Negative Syndrome Scale) reduced in length and with fewer rating options is feasible. In clinical trials, factor analysis has been used to shorten PROs (Patient Reported Outcomes).

### MATERIALS AND METHODS

There were about 70 subjects in the full analysis set with observed valid values at the scheduled time points.

#### Data Collection

To explore the efficacy of XYZ on negative symptoms in adult subjects with predominant negative symptoms, the SANS total scores of all non-global item ratings of affective flattening, alogia, avolition-apathy, and anhedonia-

asociality, were measured at the scheduled visits. Each item describes a specific impact. Responses were made on a six-point Likert-type scale (from 0=not at all, to 5=severe). The global items were omitted so that the focus was on individual rather than global domains to ensure that the assessment was more sensitive to any signal in this population. The attention items were also omitted so that any signal was not inflated because XYZ is known to affect attention. Thus, the primary efficacy variable of the study was the change in SANS total score of all nonglobal item ratings of affective flattening, alogia, avolition-apathy, and anhedonia-asociality, which is referred as SANS-18, from baseline to the end of the treatment.

Partial example data of original analysis dataset (SANS) as listed below

```
-----
```

Obs	usubjid	avisitn	avisit	paramn	pparam	aval
1	STUDYXYZ-XYX-001	0	Baseline	1	01_Unchanging Facial Expression	4
2	STUDYXYZ-XXX-001	0	Baseline	2	02_Decreased Spontaneous Movements	2
3	STUDYXYZ-XXX-001	0	Baseline	3	03_Paucity of Expressive Gestures	3
4	STUDYXYZ-XXX-001	0	Baseline	4	04_Poor Eye Contact	3
5	STUDYXYZ-XXX-001	0	Baseline	5	05_Affective Nonresponsivity	2
6	STUDYXYZ-XXX-001	0	Baseline	6	06_Lack of Vocal Inflections	4

.....

### Methods and SAS Approaches

To derive the latent dimensions, the structure of the instrument (SANS) was conducted by factor principal-components analysis. Retained principal components were varimax-rotated. Items were assigned to retained rotated principal components when they had a loading on these components (i.e., dimensions) of 0.5 or greater in absolute value. The number of principal components was then determined for all SANS items. Cronbach's correlation coefficients were calculated to refine and validate the reliability of dimensions and interpret dimensions retained rotated principal components were found to be stable. Then confirmatory factor analysis (CFA) was performed to test the hypothesized three-factor model.

### RESULTS AND DISCUSSION

To allow factor analysis be used to explore multivariable relationships between SANS parameters, SANS items need to be assigned as variables. The following codes perform such data structure conversion from vertical to horizontal. The data displayed below only shows 6 of 18 variables of SANS.

```
proc transpose data=sans out=sans1;
  by usubjid ;
  id pparam;
  var aval;
run;

data sans2;
  set sans1 (drop=_label_ _name_);
  keep _:;
run;
```

Obs	_01_Unchanging_Facial_Expression	_02_Decreased_Spontaneous_Moveme	_03_Paucity_of_Expressive_Gestur	_04_Poor_Eye_Contact	_05_Affective_Nonresponsivity	_06_Lack_of_Vocal_Inflections
1	4	3	4	4	4	4
2	4	3	4	4	3	3
3	4	4	5	3	4	4
4	3	3	4	2	3	4
5	3	4	3	4	4	3
6	4	4	4	4	2	4
7	4	4	3	4	4	3
8	4	4	4	4	3	3

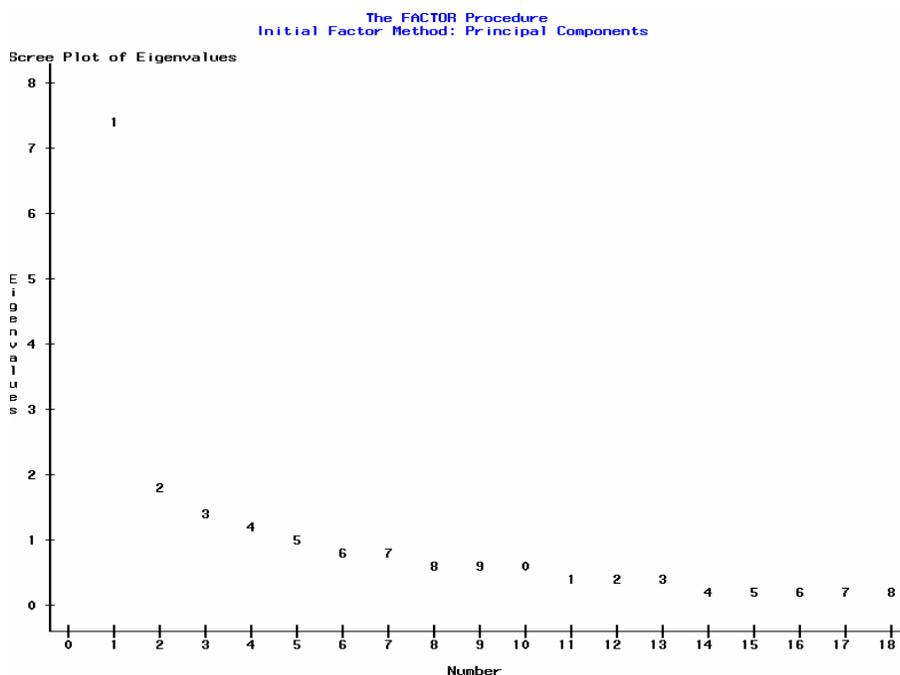
.....

### 1). Exploratory Factor Analysis (EFA)

PROC FACTOR is used to determine the factor structure (model) and explain a maximum amount of variance. Factor analysis seeks to discover common factors. The technique for extracting factors attempts to take out as much common variance as possible in the first factor. Subsequent factors are, in turn, intended to account for the maximum amount of the remaining common variance until, hopefully, no common variance remains.

```
proc factor data=sans2 method=principal nfactors=3 reorder rotate=v scree;
run;
```

Input Data Type		Raw Data			
Number of Records Read		70			
Number of Records Used		70			
N for Significance Tests		70			
Eigenvalues of the Correlation Matrix: Total = 18 Average = 1					
	Eigenvalue	Difference	Proportion	Cumulative	
1	7.32094499	5.59971293	0.4067	0.4067	
2	1.72123206	0.35866120	0.0956	0.5023	
3	1.36257085	0.25824220	0.0757	<b>0.5780</b>	
4	1.10432866	0.05217983	0.0614	0.6394	
5	1.05214882	0.16652808	0.0585	0.6978	
6	0.88562074	0.12543034	0.0492	0.7470	
7	0.76019040	0.07688230	0.0422	0.7893	
8	0.68330810	0.13526589	0.0380	0.8272	
9	0.54804221	0.04686439	0.0304	0.8577	
10	0.50117782	0.03154944	0.0278	0.8855	
11	0.46962839	0.07198111	0.0261	0.9116	
12	0.39764727	0.07844197	0.0221	0.9337	
13	0.31920530	0.05259866	0.0177	0.9514	
14	0.26660664	0.07066365	0.0148	0.9663	
15	0.19594299	0.02362927	0.0109	0.9771	
16	0.17231372	0.04166069	0.0096	0.9867	
17	0.13065303	0.02221502	0.0073	0.9940	
18	0.10843800		0.0060	1.0000	



The factors are not correlated. Eigenvalues represent amount of variance accounted for by each factor. In common practice, factor scores are calculated with a mean or sum of measured variables that load on a factor. The result from factor analysis can identify the most parsimonious number of dimensions. This is complemented by the scree plot generated by principal components analysis of the actual data super imposed on values from SANS-18 variables to identify the number of SANS dimensions. Looking for an elbow in the scree plot to explore the number of factors, the scree plot suggests that we take three factors.

Factors may not provide direct interpretation. Structural coefficients from the principal components analysis with promax rotation are used to interpret the dimensions. The results of rotation methods are sometimes referred to as derived solution because they are obtained as a second stage from the results of direct solutions. Rotation applied to the reference axes means the axes are turned about the origin until some alternative position has been reached. The simplest case is when the axes are held at 90 degree to each other, orthogonal rotation. So the factor loading will illustrate the correlations between items and factors. The REORDER option arranges factors loading by factor from largest to smallest value.

Rotated Factor Pattern

	Factor1	Factor2	Factor3
<b>_02_Decreased_Spontaneous_Moveme</b>	<b>0.86121</b>	0.04101	0.05592
<b>_03_Paucity_of_Expressive_Gestur</b>	<b>0.84083</b>	0.15462	0.03732
<b>_01_Unchanging_Facial_Expression</b>	<b>0.78065</b>	0.19709	0.12239
<b>_06_Lack_of_Vocal_Inflections</b>	<b>0.77193</b>	0.24116	0.21040
_05_Affective_Nonresponsivity	0.67898	0.48603	0.12073
_09_Poverty_of_Speech	0.67083	0.37194	0.02217
_04_Poor_Eye_Contact	0.64334	0.46364	0.03265
_16_Physical_Anergia	0.59897	0.15645	0.51543
_21_Relationships_with_Friends_a	0.47897	0.38104	0.19648
<b>_11_Blocking</b>	-0.00930	<b>0.78794</b>	-0.13560
<b>_12_Increased_Latency_of_Respons</b>	0.28725	<b>0.66239</b>	-0.02347
<b>_20_Ability_to_Feel_Intimacy_and</b>	0.25222	<b>0.62964</b>	0.24160
<b>_10_Poverty_of_Content_of_Speech</b>	0.35914	<b>0.52294</b>	0.02590
_08_Inappropriate_Affect	0.35653	0.43731	-0.34000
<b>_15_Impersistence_at_Work_or_Sch</b>	-0.09513	-0.02809	<b>0.68197</b>
<b>_18_Recreational_Interests_and_A</b>	0.41660	0.44411	<b>0.59401</b>
<b>_19_Sexual_Interest_and_Activity</b>	0.18481	-0.13266	<b>0.55978</b>
<b>_14_Grooming_and_Hygiene</b>	0.17612	0.46169	<b>0.49722</b>

## 2). Internal Consistency for Reliability (Cronbach's Alpha)

Interrelated items may be summed to obtain an overall score for each participant. Cronbach's coefficient alpha estimates the reliability of this type of scale by determining the internal consistency of the test or the average correlation of items within the test. With the ALPHA option, the CORR procedure computes Cronbach's coefficient alpha.

```
proc corr data = sans2 alpha nomiss nosimple nocorr ;
  var _02: _01: _03: _06: ;
run;
```

The larger the overall alpha coefficient, the more likely that items contribute to a reliable scale. To determine how each item reflects the reliability of the scale, you calculate a coefficient alpha after deleting each variable independently from the scale. The alpha coefficient provides information on how each variable reflects the reliability of the scale with variables. If the alpha decreases after removing a variable from the construct, then this variable is strongly correlated with other variables in the scale. On the other hand, if the alpha increases after removing a variable from the construct, then removing this variable from the scale makes the construct more reliable. Cronbach's Alpha helps to refine the items for each factor.

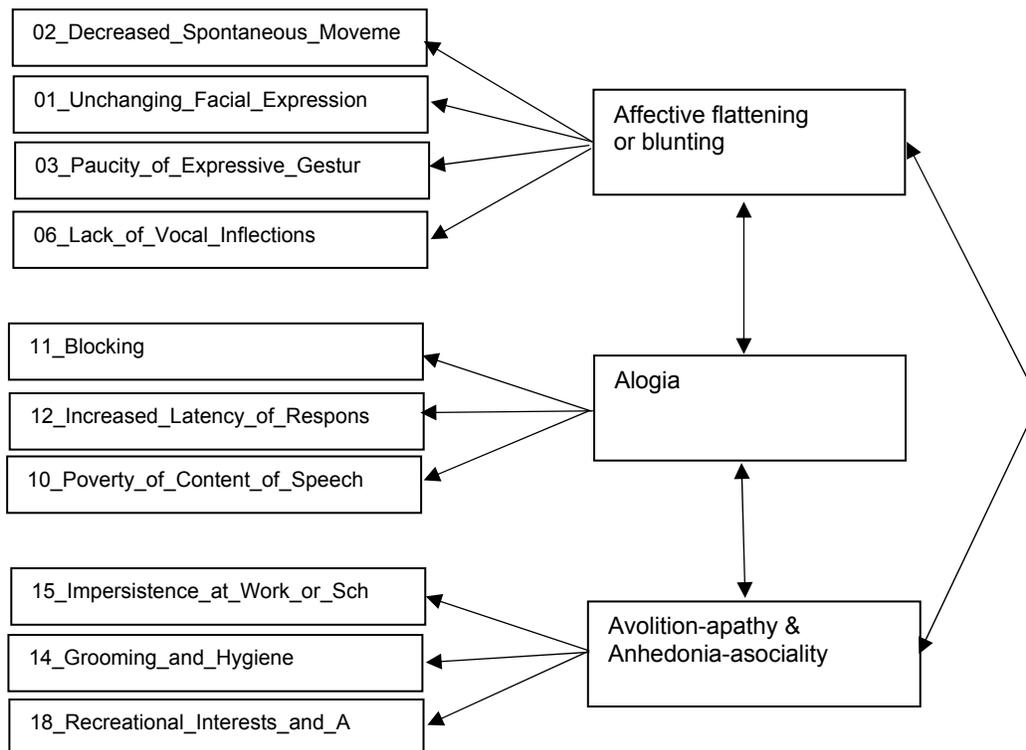
**Cronbach Coefficient Alpha for Selected SANS**

Factors	Variables	Cronbach' $\alpha$
Factor 1	_02_Deceased_Spontaneous_Move _01_Unchanging_Facial_Expression _03_Paucity_of_Expressive_Gestur _06_Lack_of_Vocal_Inflections	0.88
Factor 2	_11_Blocking _12_Increased_Latency_of_Respons _10_Poverty_of_Content_of_Speech	0.63
Factor 3	_15_Impersistence_at_Work_or_Sch _14_Grooming_and_Hygiene _18_Recreational_Interests_and_A	0.58

**3). Confirmatory Factor Analysis (CFA)**

**a). CFA Diagram**

The following figure illustrates the hypothesized model based on the exploratory findings.



### b). PROC CALIS Codes

SAS CALIS procedure (Covariance Analysis of Linear Structural Equations) estimates parameters and tests the appropriateness of structural equation models using covariance structural analysis. CFA for the 3-factor structure is run with PROC CALIS using the following SAS codes.

```

data tsans;
  set sans;
  pparamn=strip(put(paramn, z2.));
run;

proc transpose data=tsans out=tsans2;
  by usubjid ;
  id pparamn;
  var aval;
run;

title "Basic Confirmatory Factor Model";
proc calis data=tsans2 ;
  factor
    F1 ---> _01,
    F1 ---> _02,
    F1 ---> _03,
    F1 ---> _06,

    F2 ---> _10,
    F2 ---> _11,
    F2 ---> _12,

    F3 ---> _14,
    F3 ---> _15,
    F3 ---> _18;

  pvar
    F1 = 1.,
    F2 = 1.,
    F3 = 1.;

  fitindex noindextype on(only)=[chisq df probchi rmsea agfi gfi
    bentlercfi ];
run;

```

### c). Fit Statistics

PROC CALIS procedure first provides the number of observations, variables, estimated parameters, and information related to model specification. Notice measured variables may have different scales.

#### Optimization Results

Iterations	38	Function Calls	81
Jacobian Calls	40	Active Constraints	0
Objective Function	0.5099515623	Max Abs Gradient Element	0.0000187751
Lambda	0	Actual Over Pred Change	1.6367856763
Radius	0.0001566171		

Convergence criterion (GCONV=1E-8) satisfied.

Covariance Structure Analysis: Maximum Likelihood Estimation

Fit Summary

Chi-Square	46.4056
Chi-Square DF	32
Pr > Chi-Square	0.0479
Goodness of Fit Index (GFI)	0.9080
Adjusted GFI (AGFI)	0.8419
RMSEA Estimate	0.0703
Bentler Comparative Fit Index	0.8675

CFA relies on several statistical tests to determine the adequacy of model fit to the data.

Chi-square describes similarity of the observed and expected matrices. Acceptable model fit is indicated by a chi-square probability greater than or equal to 0.05. For this CFA model, the chi-square p = 0.0479, approximates the 0.05 value.

Root Mean Square Error of Approximation (RMSEA) is related to residual in the model. RMSEA values range from 0 to 1 with a smaller RMSEA value indicating better model fit. The RMSEA is 0.0703, which is close to the conventional 0.06 value for a good model fit.

The Comparative Fit Index (CFI) is equal to the discrepancy function adjusted for sample size. CFI ranges from 0 to 1 with a larger value indicating better model fit. Good model fit is indicated by a CFI value of 0.90 or greater. Bentler's comparative fit index is 0.8675, which indicates a not bad model fit.

The fit statistics indicate acceptable fit or close to indicating acceptable fit. Overall, the model seems to be quite reasonable for the data with sample size of 70 observations. The CFA analysis has confirmed the factor structure.

**d). Parameter Estimates**

When acceptable model fit is found, the next step is to determine significant parameter estimates. A t value is calculated by dividing the parameter estimate by the standard error,  $0.279 / 0.04945 = 5.642$ . Parameter estimates are significant at the 0.05 level if the t value exceeds 1.96 and at the 0.01 level if the t value exceeds 2.56. Parameter estimates for the confirmatory factor model are significant at the 0.01 level, except \_12 which is significant at 0.05 level.

Error Variances

Variable	Parameter	Estimate	Standard Error	t Value
_01	_Add04	0.27900	0.04945	5.64228
_02	_Add05	0.47278	0.07690	6.14839
_03	_Add06	0.26936	0.07263	3.70848
_06	_Add07	0.28368	0.05731	4.95000
_10	_Add08	0.84614	0.13606	6.21887
_11	_Add09	0.77346	0.15533	4.97964
_12	_Add10	0.18999	0.08336	2.27914
_14	_Add11	1.01224	0.15155	6.67944
_15	_Add12	0.70581	0.14093	5.00814
_18	_Add13	0.40336	0.11139	3.62117

**CONCLUSION**

With the aim of shortening the SANS, the psychometric properties ratings were examined in 70 participants with predominantly negative symptoms. Based on the SANS items exploratory factor analyses identified a short SANS with three symptom groupings consisting of 10 items of

- Affect 01. Unchanging Facial Expression
- Affect 02. Decreased Spontaneous Movements

Affect 03. Paucity of Expressive Gestures  
Affect 06. Lack of Vocal Inflections  
Alog 10. Poverty of Content of Speech  
Alog 11. Blocking  
Alog 12. Increased Latency of Response  
Apat 14. Grooming and Hygiene  
Apat 15. Impersistence at Work or School  
Anhed 18. Recreational Interests and Activities

The confirmatory factor analysis showed that the short three and not the original five SANS factor solution was a good fit to the data. The results are consistent with suggestions in the literature that the SANS may need revision to lessen the number of response options, reduce item redundancy, provide a concrete response for the symptoms and elaborate the SANS to assess less extreme negative symptom severity.

The short version identified in this study goes some way to attain this goal for research purposes in patients with predominantly negative symptoms. The conclusion supported that a short SANS is feasible to assess predominantly negative symptoms.

However, the present findings are subject to limitations. The study may not accurately represent patients encountered in clinical studies due to inadequately large sample size for the analyses. Accordingly, one could envisage that the short SANS has utility in some research settings and not others. In other words, the short SANS may be appropriate in some and not other circumstances. More research will be needed to further examine possible modifications to the large data of SANS and to prospectively validate the short version of SANS.

## REFERENCES

- Kirkpatrick, B., Fischer, B., 2006. Subdomains within the negative symptoms of schizophrenia: commentary. *Schizophr. Bull.* 32, 246–249.
- Andreasen, N.C., 1983. *Scale for the Assessment of Negative Symptoms*. University of Iowa, Iowa City.
- Andreasen, N.C., Arndt, S., Alliger, R., Miller, D., Flaum, M., 1995. Symptoms of schizophrenia. Methods, meanings, and mechanisms. *Arch. Gen. Psychiatry* 52, 341–351.
- Barch DM, Carter CS. Amphetamine improves cognitive function in medicated individuals with schizophrenia and in healthy volunteers. *Schizophr Res.* 2005;77:43–58.
- Levine, S.Z., Leucht, S., Psychometric analysis in support of shortening the Scale for the Assessment of Negative Symptoms. *European Neuropsychopharmacology* (2012), <http://dx.doi.org/10.1016/j.euroneuro.2012.11.008>
- Diana D. Suhr, Exploratory or Confirmatory Factor Analysis? SUGI31, Paper 200-31.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please contact the authors at:

Ben Adeyi, Director of Biostatistics  
Shire Pharmaceuticals  
735 Chesterbrook Boulevard  
Chesterbrook, PA 19087  
Work Phone: 484-595-8752  
E-mail: [badeyi@shire.com](mailto:badeyi@shire.com)  
Web: [www.shire.com](http://www.shire.com)

David Shen, Consultant Statistician  
Email: [shenda168@hotmail.com](mailto:shenda168@hotmail.com)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.