Negative Symptom Assessment-4 (NSA-4) Manual

Version 1.0

© 2006

Larry Alphs, MD PhD
Rules for Rating the NSA-4

Purpose

The purpose of the Negative Symptom Assessment (NSA) is to permit the reliable rating of behaviors commonly associated with the concept of negative symptoms in schizophrenia. It should be emphasized that this scale rates behavior and not psychopathology. Thus, scores on this scale do not necessarily imply psychopathology or etiology. For instance, subjects could achieve high ratings on some items because of depression, medication regimen, institutionalization, uncooperativeness, deficit syndrome or, even personality style. Raters using this scale should not correct ratings because, in their judgment, the etiology is something other than that sought in the particular study where it is currently used.

Sources of Information

The source of information for the NSA is a semi-structured, clinical interview and observations of behavior during that interview. This information should be applied with good clinical judgment in making a final rating. The rater may be, and is encouraged to be, aware of information from sources outside of the interview. However, it is not a requirement that other informants be identified. If the subject's report is at variance with outside reports (of nurses, family, etc.), these discrepancies should be clarified. Similarly, if the patient provides contradictory information, or other information from the interview is at variance with the subject's stated performance in a particular area, the rater should use his/her best clinical judgment in assigning a severity score for this behavior.

Reference Population

The 'normal' reference population against which the subject is to be compared is a young person in their twenties. It is not (1) the same person at another point in time; (2) a healthy person of similar age, living under similar circumstances; nor (3) another hospitalized person.

Use of Anchors (Item Descriptors)

Each of the specific items is scored 1-6. The anchors that are provided for this scale should be used as guides and should be used insofar as they apply to the persons being rated. No set of anchors could possibly describe or be applicable to all persons. If the subject being rated is in special circumstances that make it difficult to use the anchors provided, the rater should make a best effort to extrapolate the anchors provided to the subject's situation.

In general, severity scores on the NSA can be interpreted as follows:
1. The behavior being assessed is not reduce or absent as compared to a healthy young human
2. The behavior being assessed is minimally reduced, significance is questionable
3. The behavior being assessed is mildly reduced, it might only be noted as reduced by a trained rater, but he/she notes a definite reduction
4. The behavior being assessed is moderately reduced, it reduction should be obvious to an untrained rater
5. The behavior being rated is markedly reduced, this behavior is easily observable and definitely interferes with the subject's functioning
6. The behavior being assessed is severely reduced or entirely absent, it is glaring and markedly interferes with functioning.

9. An item is scored as ‘9’ only if the behavior cannot be rated despite heroic efforts by the rater to get adequate information, make a best guess. The rater’s best guess is likely to be better than that of someone not involved with the person. The ‘9’ rating should NOT be part of any total score and is provided to mark that the rating for this item has not been skipped.

Use of the Semi-Structured Interview

The semi-structured interview should be used as a stimulus to help remind the rater of which questions must be asked. Some subjects may require that the question be asked in a different way than that provided. Most subjects will require that more detailed questions be asked to gather adequate information regarding the frequency and severity of reduction for behavior being assessed. The interview can probably be accomplished in 5-10 minutes for most patients. It may be efficiently done in conjunction with a PANSS, BPRS or similar interview.

Time Frame

The time frame during which items are to be rated should be defined by the particular study, however, for item 1 only behavior noted during the interview are to be rated. For items 2-4 the time frame is frequently defined as the previous 7 days. Information from other time periods may be useful for the rater and subject to gain perspective, but only reduction in behaviors noted during the specified time frame are to be rated.

Rating the Global Score

This item rates overall severity of negative symptoms when defined as the absence or reduction of behaviors normally present in a healthy, young person. Ratings should not depend on any specific item or items from the NSA or any other similar instrument. Instead, it should measure the rater’s gestalt of the interview and often is ratable within a few minutes of initiating the interview.

Raters

Raters should be trained in the use of the NSA before they begin a clinical study. To achieve high inter-rater reliability for several raters, it is probably best to use exactly the same stimulus (like a videotape or a common interview). Although raters should be trained mental health professional persons, previous experience has demonstrated that they need not have advanced degrees to rate this scale reliably.
Negative Symptom Assessment-4 (NSA-4)

Definition:

Negative symptoms represent the reduction or absence of behaviors normally present in a healthy person. These include, but are not limited to, behaviors related to personal, social and affective behavior. Specifically, negative symptoms of schizophrenia are considered to include reduction in emotional expression and perception, reduction in the fluency and productivity of thought and speech, reduced desire for social involvement and reduced social interaction with others and a loss or lack of goal-directed behavior. These symptoms contribute to substantial reductions in the functioning of schizophrenic patients as compared to others in their society.

1. **Restricted speech quantity**
   This item assesses the amount of speech the subject provides in the course of the interview. Ratings on this item suggest that the subject gives brief answers to questions and/or provides elaborating details only after the interviewer prods him.
   
   1. Normal speech quantity
   2. Minimal reduction in quantity, may be extreme of normal
   3. Speech seems reduced, but more can be obtained with minimal prodding
   4. Speech is maintained only by regularly prodding the subject
   5. Responses are usually limited to a few words and/or details are only obtained by prodding or bribing
   6. Responses are usually non-verbal or limited to 1 or 2 word answers (despite one’s best efforts to get the subject to elaborate)
   9. Not ratable (use only when all efforts to rate this item have failed)

2. **Emotion: Reduced range (specify time frame for this assessment)**

   Emotion is the feeling content of a person’s inner life. This item assesses the range of emotion experienced by the subject during the last week (or other specified time period). Base ratings on the subject’s answers to queries of whether he/she has felt happy, sad, etc. during the last week, as well as any reports of having these emotions later in the interview. A full range of emotions would include, but not be limited to happiness, sadness, pride, fear, surprise, and anger. This item should be distinguished from the capacity to display affect, which is rated elsewhere. (If you sense that a subject’s emotional life is autistic and not contextually validated, rate his/her emotional range according to your interpretations of his/her experience.)

   1. Normal range of emotion
   2. Minimal reduction in range, may be extreme of normal
   3. Range seems restricted relative to a normal person, but during the specified time frame subject convincingly reports at least 4 emotions.
   4. Subject convincingly identifies 2 or 3 emotional experiences
   5. Subject convincingly identifies only 1 emotional experience
   6. Subject reports little or no emotional range
   9. Not ratable (use only when all efforts to rate this item have failed)

3. **Reduced social drive**

   This item assesses how much the subject desires to initiate social interactions. Desire may be measured in part by the number of actual or attempted social contacts with others. To rate severity probes the type of social interactions, and their frequency. Remember the reference range is a normal 25 year old. Many subjects may be rated 2 to 3.
1. normal social drive
2. minimal reduction in social drive, may be extreme of normal desire for social interactions seems somewhat reduced
3. obvious reduction in desire to initiate social contacts, but a number of social contacts are initiated each week
4. marked reduction in desire to initiate social contacts, but a few contacts are maintained at subject's initiation (as with family)
5. no desire to initiate any social interactions
6. not ratable (use only when all efforts to rate this item have failed)

4. Reduced interests

This item assesses the range and intensity of the subject's interests.

1. Normal interests
2. Minimal reduction in interests, may be extreme of normal range of interests and/or commitment to them seems diminished
3. Range of interests is clearly diminished and is not particularly committed to interests held
4. Only 1 or 2 interests reported, and these pursued superficially
5. No identifiable goals
6. Not ratable (use only when all efforts to rate this item have failed)

Global Rating

1. No evidence of negative symptoms of schizophrenia
2. Minimal evidence of negative symptoms
3. Mild evidence of negative symptoms
4. Moderate evidence negative symptoms, apparent to the casual observer
5. Marked evidence negative symptoms, readily apparent to casual observer
6. Severe, not only obvious but has marked impact on functioning
7. Extremely severe symptoms, they are incapacitating for subject
SEMI-STRUCTURED INTERVIEW FOR RATING THE NSA-4

Restricted speech quantity
No specific questions are required. Use information from full period of contact with the subject.

Emotion: Reduced range (Specify time frame for this assessment, usually the past week)

1. How have you been feeling in the past week?
2. Have you felt anxious, nervous or worried during the past week? What has that been like for you? How bad was it? What kind of things make you feel anxious?
3. Have there been times during the past week when you felt happy? What made you feel this way? What kind of things make you feel happy? Have you felt proud about any accomplishments in the past week?
4. Have you felt sad or depressed during the past week? What has that been like for you? How bad was it? What kind of things make you feel sad or depressed? Have you felt guilty about something you did this past week?
5. Have you felt scared, surprised about anything this last week?
6. Have you felt angry about anything this last week?
7. If after the completion of this section, few emotions are identified, suggest a situation that might have caused emotion and ask how the subject would feel [e.g., getting lost, a car accident, an award, a quarrel, a party], but do not suggest the emotion that might have been experienced.

Social Drive (Specify time frame for this assessment, usually the past week)

1. Do you live alone, or with someone else?
2. Do you like to be around other people? Who are your friends? How often do you see them? What have you done with them in the past week?
3. Have you called them on the phone? Texted them? When you got together this past week, who decided what to do and where to go?
4. Would you say that your family or friends are concerned about your happiness and well-being? What makes you think that?
5. Do you have difficulty feeling emotionally close to your family and friends? Could you explain that?

Reduced interests (Specify time frame for this assessment, usually the past week)
1. What do you enjoy doing? What else do you enjoy? Have you done these things in the past week? What are your plans to do them again? Do you have a hobby? Tell me about it? What have you done in the last week?

2. Are you interested in what is going on in the world? Do you read the newspapers? Do you watch the news on TV or internet? Can you tell me about some of the important news stories of the past week?

3. Do you like sports? What is your favorite sport? Which is your favorite team? Who are the top players in this sport? Have you followed them in the past week? Have you played in any sport during the past week?

4. Do you like music? What is your favorite type of music? Why do you like it? Have you listen to it in the last week? What is your favorite music group? Why do you like them?
HOW DOES THE NSA-4 COMPARE TO THE NSA-16?
Janet B.W. Williams, Lori M. Garzio, Douglas Osman, Danielle Popp

Background. The 16-item Negative Symptom Assessment (NSA-16) is increasingly used as a validated measure to track response to treatment of negative symptoms in clinical trials of schizophrenia. The NSA-16, although reliable, takes up to 30 minutes to administer. As clinical trials have become more complex, a briefer assessment tool would be useful. Alphs et al have proposed a four-item version, the NSA-4, as a reliable and valid alternative to the NSA-16. Four of the 16-NSA items are included: restricted speech quantity, emotion: reduced range, reduced social drive, and reduced interests; in addition, both versions of the scale include an overall global rating of negative symptoms. Alphs et al examined the psychometric properties of the NSA-4 in two randomized clinical trials. The current study is an effort to replicate their findings in two other large schizophrenia trials.

Methods. Data are from two Phase 2 randomized double-blind studies comparing an antipsychotic medication with placebo in the treatment of subjects with DSM-IV-TR schizophrenia with prominent negative symptoms. Subjects were interviewed by live two-way videoconferencing at screen, baseline, and 11 more visits, including end point. Raters were from a centralized independent and blinded cohort who were uniformly trained via initial didactic and applied training, and were monitored throughout the studies to ensure calibration and prevent drift. At each visit, raters administered the PANSS immediately followed by the NSA-16. Correlation coefficients between the NSA-16 and the NSA-4 were calculated for the NSA global rating, the PANSS negative and positive subscales, and the Marder factors for PANSS negative symptoms, anxiety/depression, hostility/excitement, disorganized thought, and positive symptoms. Cronbach’s alpha and interrater reliability (calculated as an ICC) were determined for the NSA-16 and NSA-4.

Results. The NSA-16 was administered a total of 2804 times by 29 Central Raters, to a total of 483 subjects enrolled in two clinical trials. Overall, the correlation between the total scores of the NSA-4 and NSA-16 was high (0.86). Good convergent validity of the NSA-4 was demonstrated by correlations between the NSA-4 and the NSA global rating (r = 0.67), as well as the PANSS negative subscale (r = 0.73) and the PANSS negative symptoms Marder factor (r = 0.73). Divergent validity in our sample was demonstrated by low correlations between the NSA-4 and the following PANSS Marder factors: anxiety/depression (r = -0.11), disorganized thought (r = 0.29), hostility/excitement (r = 0.03), and PANSS positive symptoms (r = 0.13). Cronbach’s alpha was lower for the NSA-4 (α = 0.65) compared to the NSA-16 (α = 0.87) Finally, the interrater reliability estimates for the NSA-4 and NSA-16 were 0.94 and 0.97, respectively.

Discussion. The PANSS and NSA-16 in this study were not administered independently of one another, so the usefulness of the NSA-4 alone can only be evaluated in the context of its pairing with the PANSS. Overall, these results were very similar to those obtained by Alphs et al. In the hands of highly trained and calibrated Central Raters, the NSA-4 had very good overall agreement with the NSA-16, and even higher convergent and divergent validity with the selected PANSS subscales and interrater reliability than was demonstrated by Alphs et al.