The Positive and Negative Syndrome Scale and the Brief Psychiatric Rating Scale

Reliability, Comparability, and Predictive Validity

MORRIS BELL, PH.D.,[1,2] ROBERT MILSTEIN, M.D., PH.D.,[2] JOSEPH BEAM-GOULET, M.S.,[1,2] PAUL LYSAKER, PH.D.,[1,2] AND DOMENIC CICCHETTI, PH.D.[1,2]

In a psychiatric rehabilitation study, 154 concurrent ratings were performed using the 30-item Positive and Negative Syndrome Scale (PANSS) and the 18-item Brief Psychiatric Rating Scale (BPRS). Although both instruments had excellent interrater reliability, the PANSS was consistently better: on the 18 symptom items the two instruments share, the PANSS had higher intraclass r's on 14; for the syndromes, the PANSS was higher than the BPRS on positive, negative, and total. Weighted Kappas comparing shared items revealed that most were not interchangeable, with only three coefficients in the excellent range. However, syndrome scale scores were very highly correlated and resulted in similar classification for negative schizophrenia. Ten of the 12 items of the PANSS not included in the BPRS had low zero-order correlations with BPRS items, which suggests that they measure symptoms distinct from those measured by the BPRS and should add to clinical predictive power. This proved true in our study of rehabilitation of patients with schizophrenia. PANSS symptom ratings explained up to 55% of the variance on seven measures of work performance, whereas the BPRS had lower predictive power on six of the seven measures. We concluded that the PANSS may be superior to the BPRS in clinical research on schizophrenia and that most BPRS items are not interchangeable with identically named PANSS items.

— *J Nerv Ment Dis* 180:723–728, 1992

The Positive and Negative Syndrome Scale (PANSS) is a rating scale designed to provide a "trustworthy" instrument for the study of symptom phenomenology in schizophrenia (Kay et al., 1987). It is composed of the 18-item Brief Psychiatric Rating Scale (BPRS; Overall and Gorham, 1962) and 12 additional items from the Psychopathology Rating Scale (Singh and Kay, 1975). The authors believed this expansion of the BPRS would capture clinical phenomena not accounted for by the BPRS, including several negative symptoms and a number of general symptoms.

Other changes included more extensive anchor descriptors and identification of the first two points on the 7-point scale (absence, minimal) as being below clinical significance. The BPRS break point for the presence of a clinical entity is between the first and second points (not present, very mild) of the 7-point scale.

In a series of papers, the PANSS demonstrated strong psychometric properties, including excellent interrater reliability (Kay et al., 1989), good internal consistency,

and good to fair test-retest reliability (Kay et al., 1987). Validity of the PANSS has been supported in studies of psychopharmacological treatment (Kay and Opler, 1985; Singh et al., 1987), cognitive functioning (Kay, 1990), discriminant validity, and course of illness (Kay et al., 1987; Lindenmayer et al., 1986).

The PANSS authors did not, however, investigate the relationship between their instrument and the BPRS. The BPRS is widely used in psychiatric research and has been employed in investigations of course of positive and negative symptoms in schizophrenia (Kane et al., 1988). In considering whether to employ the PANSS, investigators want to know whether it will surpass the BPRS in its power to describe schizophrenic phenomenology.

A similar question was asked comparing the BPRS with the Scale for the Assessment of Negative Symptoms (SAN; Andreasen, 1982). The authors (Thiemann et al., 1987) determined that the BPRS withdrawal-retardation subscale was highly correlated with the SANS summary score and that SANS interrater reliabilities compared unfavorably with the BPRS interrater reliability. They concluded that there was little to be gained by using the SANS rather than the BPRS.

Given the robustness of the BPRS, it is desirable to learn whether, and in what ways, the additional effort

involved in using the PANSS would be justified. The present study compares the PANSS to the BPRS by examining interrater reliabilities and internal consistencies of the syndrome scales; it investigates whether the additional PANSS items capture clinical phenomena distinct from the BPRS items, and, if so, whether the PANSS has greater power to predict measures of work performance in a psychosocial treatment study.

In addition, Kay[3] suggested that we examine the comparability of the 18 items shared by the PANSS and the BPRS. If the modifications the PANSS authors have made in descriptors for the BPRS symptom items of the same name did not substantially alter their ratings, then it might be possible to convert BPRS scores into PANSS ratings. This would make the vast psychiatric literature using the BPRS available for comparison with PANSS studies.

Method

Subjects

A total of 154 concurrent symptom ratings were obtained from 56 subjects with DSM-III-R diagnoses of schizophrenia or schizoaffective disorder. Assessments of work performance were obtained from a subset of 30 of the initial 56 subjects who accepted work placements. Subjects were recruited from the Psychiatry Service of a Department of Veterans Affairs Medical Center for a study of the rehabilitative effects of productive activity on schizophrenia. Subject demographics and mean subscale values are presented in Table 1.

Instruments

Positive and Negative Syndrome Scale. The PANSS is a 30-item rating scale, completed by clinically trained research staff at the conclusion of chart review, and a semistructured interview. The PANSS items are grouped into three rationally derived categories: positive symptoms, or symptoms that should not be present in a normal mental status (*e.g.*, hallucinations and delusions); negative symptoms, symptoms that represent behaviors that should be present in a normal mental status but are not (*e.g.*, blunted affect and emotional withdrawal); and general symptoms (*e.g.*, anxiety and depression). The list of PANSS items is found in Table 2. Scores for positive, negative, and general scales are obtained by summing scores within each category. An overall total is obtained by summing all items.

The Brief Psychiatric Rating Scale. The BPRS (Overall and Gorham, 1962) is an 18-item rating scale filled out by clinically trained research staff. The BPRS is completed after chart review and a semistructured interview. Kane et al. (1988) have grouped the items of

[3]S. R. Kay (personal communication, 1989).

TABLE 1
Demographics and Mean Subscale Values

	Reliability ($N = 56$)	Predictive ($N = 30$)
Group Frequency		
Gender		
Female	4	2
Male	52	28
Race		
Black	6	2
Hispanic	2	1
Caucasian	48	27
Primary Diagnosis		
Disorganized	2	2
Paranoid	30	16
Residual	1	0
Schizoaffective	16	7
Undifferentiated	7	5
Course of Illness		
Episodic	11	4
Undetermined	8	4
Continuous	37	22
Age	40.2 ± 8.6	38.0 ± 8.4
Slosson IQ score	101.8 ± 14.8	103.4 ± 12.9
Years of education	12.5 ± 2.1	12.4 ± 2.3
Age at first hospitalization	23.3 ± 5.7	23.6 ± 6.3
Total no. (lifetime) of hospitalizations	10.0 ± 7.8	10.8 ± 9.1
Duration of illness (yrs)	16.9 ± 7.8	15.2 ± 9.2
No. of mos. in hospital, last 3 yrs	7.1 ± 6.8	7.8 ± 7.2
Years of work	3.6 ± 3.8	3.5 ± 4.2
BPRS negative	4.3 ± 3.1	4.4 ± 3.4
BPRS positive	12.3 ± 6.2	12.3 ± 5.9
BPRS general	8.4 ± 3.6	8.3 ± 3.6
BPRS total	25.1 ± 8.8	25.1 ± 8.9
PANSS negative	17.3 ± 4.5	17.6 ± 4.7
PANSS positive	19.0 ± 5.9	19.2 ± 5.5
PANSS general	38.3 ± 7.1	39.1 ± 7.2
PANSS total	74.7 ± 13.2	75.9 ± 12.6
GAF	35.6 ± 5.7	35.6 ± 4.2

$\bar{X}$ ± SD.

the BPRS into categories paralleling those in the PANSS: positive, negative, and general; they are presented in Table 2.

Work Personality Profile. The Work Personality Profile (WPP; Bolton and Roessler, 1986) is a 58-item work performance inventory completed after direct observation of a subject's work behavior and supervisor interview. From these ratings, five factor scores are generated: work skills, social skills, work motivation, self-control and judgment, and personal presentation. These scores are compared to norms from a rehabilitation population.

Minnesota Satisfactoriness Scale. The Minnesota Satisfactoriness Scale (MSS; Gibson et al., 1970) is a work performance inventory rated by supervisors. Four scores are derived: work performance, work confor-

TABLE 2
Interrater Reliabilities for the Positive and Negative Syndrome Scale and the Brief Psychiatric Rating Scale, for Three Raters

PANSS Score	Intraclass r	BPRS Score	Intraclass r
Positive Scales			
Delusions	.93	Unusual thoughts	.82
Conceptual disorganization	.86	Conceptual disorganization	.75
Hallucinations	.93	Hallucinations	.92
Excitement	.81	Excitement	.85
Grandiosity	.92	Grandiosity	.86
Suspiciousness	.88	Suspiciousness	.84
Hostility	.82	Hostility	.69
Positive total	.93	Mannerisms and posturing	.52
		Positive total	.87
Negative Scales			
Blunted affect	.88	Blunted affect	.71
Emotional withdrawal	.84	Emotional withdrawal	.52
Poor rapport	.72		
Passive withdrawal	.79	Uncooperativeness	.61
Difficulty in abstract thinking	.90	Motor retardation	.74
		Disorientation	.87
Lack of spontaneity	.80	Negative total	.90
Stereotyped thinking	.63		
Negative Total	.94		
General Scales			
Somatic concern	.92	Somatic concern	.84
Anxiety	.77	Anxiety	.70
Guilt feelings	.84	Guilt feelings	.82
Tension	.62	Tension	.67
Mannerisms and posturing	.84	Depressive mood	.86
		General total	.89
Depression	.80		
Motor retardation	.80		
Uncooperativeness	.57		
Unusual thoughts	.82		
Disorientation	.76		
Poor attention	.54		
Lack of insight and judgment	.74		
Disturbance of volition	.73		
Impulse control	.64		
Preoccupation	.65		
Active social avoidance	.74		
General total	.84		
Overall Scale Totals			
PANSS total	.91	BPRS total	.87

mance, dependability, and personal adjustment. Scores are compared with norms for general workers.

Procedures

Subjects were concurrently rated on the PANSS and BPRS as part of an initial screening battery or a weekly interview for a study on rehabilitation. Raters were a Ph.D. clinical psychologist and four master's level assistants. All had used PANSS training tapes prepared by Kay and participated in practice interviews before their ratings were included. Before the interviews, demographic information was collected. The DSM-III-R di-

agnoses were determined using the Structured Clinical Interview for DSM-III-R (Spitzer et al., 1989).

After the PANSS and BPRS interviews, subjects were offered job placement such as work in the hospital escort service, pharmacy, or blind center. Subjects' work performance was assessed weekly using the WPP and MSS.

Analysis

Agreement between scores of similarly named items on the PANSS and BPRS was calculated using Kappa (Cohen, 1968; Fleiss, 1981), by applying a weighing system recommended by Cicchetti (1976) and Cicchetti and Fleiss (1977). Concurrent symptom ratings were compared and resulted in 154 pairs of observations.

Interrater reliabilities within scales were assessed using the intraclass correlation coefficient (Bartko, 1966). Calculations were performed using a computerized scoring routine (Cicchetti and Showalter, 1988) and implementing an algorithm that takes into account both the number and specific set of judges that vary from subject to subject (*i.e.*, of three raters, all three or any combination of two). There were 55 ratings. Cronbach's alpha (Anastasi, 1988) was calculated item-by-item and overall for the three subscales of the BPRS and PANSS.

Symptom ratings from the initial interview were used to predict the first-week work performance measures in a series of multiple regression procedures (optimal subsets). First, symptom items were used to predict each of seven measures of work performance. The combination of symptoms with the highest R^2 were retained. The adjusted R^2's in Table 4 are the values found using the best combination of four symptom variables. Different combinations of symptoms predict each performance measure. For example, the four PANSS negative symptoms best predicting work skills resulted in an adjusted R^2 of .35.

Results

The PANSS and BPRS interrater reliabilities are presented in Table 2. Both instruments had high interrater reliability for each of the syndrome scales and the symptom total, although the PANSS reliability coefficients were higher on 14 of the 18 items the instruments share. Uncooperativeness and poor attention on the PANSS and emotional withdrawal and mannerisms on the BPRS were the only items below a good (.60 to .74 [Fleiss, 1981]) range of interrater reliability.

Internal consistency measures on PANSS positive and negative syndrome scales (alpha = .74, .69) were similar to those of BPRS syndrome scales (alpha = .69, .68) although the PANSS general scale was considerably more homogenous (alpha = .64) than the BPRS general scale (alpha = .46).

TABLE 3

Pearson Correlation Coefficients and Weighted Kappa Coefficients (K) *for Items Shared by the PANSS and BPRS*

PANSS Item	BPRS Item	r	K[a]
Delusions	Unusual Thoughts	.88	.72
Hallucinations	Hallucinations	.93	.86
Conceptual disorganization	Conceptual disorganization	.84	.65
Excitement	Excitement	.70	.56
Grandiosity	Grandiosity	.94	.82
Hostility	Hostility	.72	.53
Suspiciousness	Suspiciousness	.86	.70
Blunted affect	Blunted affect	.89	.86
Emotional withdrawal	Emotional withdrawal	.43	.24
Passive withdrawal	Emotional withdrawal	.47	.24
Somatic concern	Somatic concern	.84	.71
Anxiety	Anxiety	.57	.54
Tension	Tension	.75	.51
Guilt feelings	Guilt feelings	.89	.72
Depression	Depressive feelings	.82	.63
Motor retardation	Motor retardation	.78	.74
Mannerisms and posturing	Mannerisms and posturing	.68	.46
Uncooperativeness	Uncooperativenss	.51	.38
Disorientation	Disorientation	.76	.60
Unusual thoughts	Unusual thoughts	.81	.63
Active social avoidance	Emotional withdrawal	.20	.11

[a]Qualitative descriptors of the clinical or practical meaning of weighted Kappa or intraclass r coefficients (Cicchetti and Sparrow, 1981): ≥ .75, excellent; .60 to .74, good; .40 to .59, fair/moderate; ≤ .40, poor.

The PANSS items originating from the BPRS were compared with 18 BPRS items in Table 3. Twenty-one comparisons were made because emotional withdrawal on the BPRS contains descriptors appropriate to PANSS active and passive social avoidance as well as emotional withdrawal, and unusual thoughts from the BPRS contains descriptors that apply to the delusions and unusual thoughts items on the PANSS. Weighted Kappa is the appropriate procedure for determining comparability; Pearson coefficients r (which can inflate reliability estimates) allow reference to previous literature on the BPRS. Results indicated that only three items (blunted affect, hallucinations, and grandiosity) had kappas in the excellent range (≥ .75), which demonstrates that modifications have altered the meaning of the items so that they are not interchangeable with BPRS items of the same designation. Items with the poorest Kappas had the poorest interrater reliability, which suggests that error of measurement compounds discrepancies between instruments.

Pearson correlation coefficients between the 12 added symptom items of the PANSS and the 18 BPRS items revealed only modest relationships. Of the 216 correlations, only two were above .60 (passive withdrawal with blunted affect, $r = .61$; active social avoidance with suspiciousness, $r = .65$, $p < .0001$). Lack of spontaneity was moderately correlated with BPRS negative symptoms (motor retardation, $r = .52$; uncooperativeness, $r = .55$; and blunted affect, $r = .55$, $p < .0001$) and impulse control correlated with BPRS hostility (.55).

Correlations between PANSS and BPRS syndrome scales revealed a very close association. Negative syndromes were correlated .82 and positive syndromes .92. The general scales were moderately correlated at .61, whereas the scale totals were .84.

To demonstrate the equivalence of the negative syndrome scores on the two instruments, a typology was created of high negative and low negative subjects using the median as the cutoff score on each instrument (PANSS > 17; BPRS > 3). Agreement was 85.72%, with 48 subjects falling in the same category, three subjects rated high on the BPRS and low on the PANSS, and five subjects rated high on the PANSS and low on the BPRS (Kappa = .72, $p < .001$).

Table 4 compares the predictive power of the two instruments with measures of work performance. A maximum of four symptom items was used to predict each work score. The PANSS negative symptom items had significant predictive power for all seven work scores; the BPRS negative symptom items had significant predictive power for five. The PANSS positive symptom items predicted three work scores; the BPRS positive symptom items predicted four. The PANSS general symptom items proved significant predictors of six work scores; the BPRS general symptoms predicted one.

Discussion

Have the PANSS authors succeeded in developing an instrument that may surpass the BPRS in studies of

TABLE 4
Prediction of Work Performance Ratings from Subsets of PANSS and BPRS Positive, Negative, and General Symptom Ratings Using Multiple Regression

| | Adjusted R^2 | | | | | |
| | PANSS[a] | | | BPRS[a] | | |
Performance Variables	Negative subsets	Positive subsets	General subsets	Negative subsets	Positive subsets	General subsets
WPP Ratings						
Work skills	.35**	NS	.38**	.27**	.18*	NS
Social skills	.54***	NS	.35***	.39***	NS	NS
Work motivation	.24*	NS	.31**	NS	NS	NS
Work conformance	.15*	.16*	.44***	NS	.31**	.24**
Personal presentation	.23**	NS	.24**	.18**	.14*	NS
MSS Ratings						
Work quality	.55***	.33**	.41**	.21**	.20*	NS
Response to rules and authority	.19*	.21*	NS	.17*	NS	NS

[a]Each subset contained a maximum of four symptoms that best predicted each work performance rating (*Cf. Method: Analysis*).
*$p < .05$; ** $p < .01$; *** $p < .001$.

schizophrenia? The results of this study suggest they have. Although both instruments have high interrater reliability, we found superior reliability for the PANSS on the shared symptom items and the additional PANSS items. This finding suggests the symptom descriptors of the PANSS were superior to those of the BPRS. One advantage of the PANSS is that it distinguishes between severity and frequency of occurrence of a symptom, elements that are combined in a single anchor on the BPRS (*e.g.*, anxiety: mild/infrequent, mild/transient, upset most of the time and/or attacks of acute anxiety, continuously shaken). The PANSS interrater reliabilities may also have been enhanced by using videotapes as part of the training procedure. Our reliabilities for the BPRS were higher than those usually reported (Hedlund and Viewig, 1980) and may indicate that our reliabilities on the PANSS are unusually high as well. More experience by a variety of investigators will be needed before reaching a conclusion about the reliability of the PANSS.

The internal consistency of the PANSS syndrome scales, the general scale, and the instrument as a whole was greater than that of the BPRS. This suggests that the additional items on the PANSS and the modifications of shared items increased the homogeneity of the syndrome scales. The greatest improvement was on the general scale, which had been increased to 16 symptom items. These findings suggest that the PANSS rational categorization is psychometrically valid and yields more homogeneous syndromes than the BPRS categorizations used by Kane et al. (1988).

An exploration of the comparability of the shared items demonstrated that modifications had sufficiently altered the meaning of those items so as to make most of them noninterchangeable. A BPRS finding could be substituted for a PANSS finding for only three symptoms. However, the negative and positive syndrome

scores were a different matter. The close association between the two instruments approached theoretical limits based on reliability error variance. Our correlation for negative symptom syndrome scores of .82 actually exceeds the .70 between the BPRS and SANS as reported by Thiemann et al. (1987). The high agreement using high and low negative symptom categories supports the use of either instrument in dichotomization. Investigators can reasonably conclude that studies using BPRS negative syndrome scores can be compared with studies using the PANSS negative syndrome scores.

The extra effort in rating the 12 additional items of the PANSS is justified by their generally low correlations with BPRS items. The additional items also increased predictive power. Although positive symptoms for both instruments were not especially good predictors of work variables, PANSS negative and especially general symptoms proved much more useful than the BPRS in explaining variance in work behavior. The PANSS, although it required additional effort, resulted in more powerful predictions of work performance. Important findings may be lost by employing an instrument less sensitive to schizophrenic phenomenology than the PANSS.

Taken together, these results support continued use of the PANSS in schizophrenia research. At this point, the PANSS provides an instrument with considerable psychometric stability and predictive power, and we encourage investigators to consider its use.

References

Anastasi A (Ed) (1988) *Psychological testing.* New York: Macmillan.

Andreasen NC (1982) Negative symptoms in schizophrenia: Definition and reliability. *Arch Gen Psychiatry* 39:784–788.

Bartko JJ (1966) The intraclass correlation coefficient as a measure of reliability. *Psychol Rep* 19:3–11.

Bolton B, Roessler R (1986) *Manual for the Work Personality Profile.*

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.

fastcase®
Smarter legal research.