

## Sensitivity to Change of the Brief Psychiatric Rating Scale—Extended (BPRS–E): An Item and Subscale Analysis

Gary M. Burlingame, Scott Seaman,  
Jennifer E. Johnson, Jason Whipple, and  
Elizabeth Richardson  
Brigham Young University

Frank Rees, Dallas Earnshaw,  
Richard Spencer, Mark Payne, and  
Brock O’Neil  
Utah State Hospital

The Brief Psychiatric Rating Scale (BPRS) is the most commonly used outcome measure for the severely and persistently mentally ill (SPMI) population, possessing good interrater reliability, concurrent validity, and a strong factor structure. However, psychometric study of the extended version of the BPRS (the BPRS–E) is limited when compared with earlier versions (BPRS and BPRS–A). This study examined the item, factor, and diagnosis-specific sensitivity to change of the BPRS–E, the most recent version of this popular scale. Assessments were conducted at 90-day intervals with 201 adult psychiatric inpatients at the Utah State Hospital, yielding 786 symptom ratings. Of note was that ratings were conducted by independent assessors who were unaware of patients’ treatment status. All but 2 of the 24 BPRS–E items, all 4 factors, and the total score were found to be sensitive to change when comparing patients’ admission and discharge scores. Patient diagnosis was not associated with item, factor, or total score sensitivity to change. These findings extend the psychometric support for the BPRS–E and have implications for assessing outcome with the SPMI population.

Robust assessment of patient outcome is essential to the evidence-based treatment guidelines that are emerging from professional, governmental, and third-party payers (American Psychological Association, 2005; Institute of Medicine, 2005; Lambert & Ogles, 2004). In some cases, the lack of consensus regarding efficacious treatments for particular patient populations may be due, in part, to the psychometric features of the measures used to assess patient outcomes rather than treatment effectiveness per se. For instance, tracking the effectiveness of long-term care for a severely and persistently mentally ill (SPMI) patient population has unique psychometric challenges. Time constraints within the health care settings

servicing chronic populations require that administration and scoring of measures be brief. Because SPMI populations by definition require extended care, outcome measures must (a) be valid with repeated use, (b) capture multitrait domains applicable to the wide variety of symptoms, and (c) be sensitive enough to detect changes in the chronic course of clients’ illnesses.

Outcome assessment in the SPMI population also requires careful consideration of the source of outcome ratings. Self-report measures are attractive because they require less staff time. However, scores may not appear valid and can fluctuate greatly depending on the patient’s current condition, honesty, or insight (cf. Burlingame et al., 2005). Likewise, therapist report measures can be a cause for concern. Lambert and Hill (1994) presented data showing that therapists consistently report higher levels of improvement on clinician rating measures when compared with self-report measures. If true, such biased ratings are less helpful when assessing the actual progress or deterioration of a patient over time; that is, estimates of improvement may be magnified. Given the potential biases of both patient self-report and therapist report, surveyors who are independent of daily

---

Gary M. Burlingame, Scott Seaman, Jennifer E. Johnson, Jason Whipple, and Elizabeth Richardson, Department of Psychology and Clinical Psychology, Brigham Young University; Frank Rees, Dallas Earnshaw, Richard Spencer, Mark Payne, Hospital Administration (Executive Staff), Utah State Hospital, Provo, Utah; and Brock O’Neil, Quality Resources, Utah State Hospital, Provo, Utah.

Correspondence concerning this article should be addressed to Gary M. Burlingame, Department of Psychology and Clinical Psychology, Brigham Young University, 284 Taylor Building, Provo, UT 84602-8626. E-mail: gary\_burlingame@byu.edu.

treatment and case management may be the most reliable source of information about patient change.

The Brief Psychiatric Rating Scale (BPRS) is an independent evaluator-rated instrument that has been noted for its flexibility, simplicity, and usefulness as a repeat measure in an SPMI population (Lukoff, Liberman, & Nuechterlein, 1986; Morlan & Tan, 1998; Rhoades & Overall, 1988). The scale is widely adaptable to a variety of treatment settings (Bech, Larsen, & Andersen, 1988) and patient populations, including schizophrenic (Borison, Sinha, Haverstock, McLarnon, & Diamond, 1989; Gur et al., 1991) and depressed patients (Feighner, Merideth, & Claghorn, 1984).

The BPRS was first designed in 1962 by Overall and Gorham to provide an efficient, rapid evaluation of change resulting from treatment, as well as an overall description of major symptom characteristics. Since then, the BPRS has been revised twice, with the most recent version containing 24 items; this version is referred to as the Brief Psychiatric Rating Scale—Extended (BPRS–E; Dingemans, Linszen, Lenior, & Smeets, 1995). A few advantages of the BPRS–E over earlier versions (cf. Lachar, Espadas, & Bailey, 2004) include the addition of (a) item rating anchors, (b) items that capture more symptom domains (e.g., excitement and disorientation), and (c) six items that rely on rater observations (e.g., self-neglect, motor hyperactivity). These refinements have increased its clinical utility by broadening the symptoms captured while also improving interrater reliability.

In a recent meta-analysis (Burlingame et al., 2005), the BPRS was found to be the most commonly used outcome measure for assessing change in SPMI populations; a modest majority of studies used the predecessors of the BPRS–E (the BPRS and BPRS–A). BPRS versions have been found to have good reliability and validity (Hedlund & Viewberg, 1980; Newcomer, Faustman, Yeh, & Csernansky, 1990; Rhoades & Overall, 1988; Thieman, Csernansky, & Berger, 1987), and factor analytic studies have shown the BPRS to have a fairly consistent factor structure in general clinical patient populations (Burger, Calsyn, Morse, Klinkenberg, & Trusty, 1997; Burlingame & Lee, 2004; Dingemans et al., 1995; Long & Brekke, 1999; Morlan & Tan, 1998; Mueser, Curran, & Mc-

Hugo, 1997; Ownby & Seibel, 1994). Furthermore, the BPRS meets the criteria for brief administration and scoring, validity with repeated use (Inch, Crossley, Keegan, & Thora-rinson, 1997; Packer, Husted, Cohen, & Tomlinson, 1997; Swett, 1996; Varner, Chen, Swann, & Moeller, 2000), and multitrait content applicable to a wide variety of symptoms and independent observer ratings (Linszen, Dingeman, & Lenior, 1994; Marder et al., 1991; Subotnik & Nuechterlein, 1988; Van der Does, Linszen, Dingemans, Nugter, & Scholte, 1993).

Fewer studies have reported on the sensitivity to change of versions of the BPRS. Existing studies (e.g., Burlingame et al., 2005) have reported a large pre- to posttreatment effect size (1.21) by averaging patient gains from 18 studies that used the BPRS total score. However, because measuring sensitivity to change is a relatively new development in the outcome literature (Lambert & Hill, 1994), metrics and methodologies are rapidly evolving. As a result, we found no previous study that systematically used the most current recommendations for measuring sensitivity to change. For example, few studies reported change at the subscale level, and virtually no information was available at the item level. Demonstrating sensitivity to change of individual BPRS items is particularly important because items may provide specific symptom information that could be useful in patient care decisions. Furthermore, we found no previous study that has tested for differences in the sensitivity to change of the BPRS among varying patient populations (e.g., diagnosis-specific sensitivity to change). Doing so is important because recent factor analytic studies of the BPRS–E (Lachar et al., 2004; Thomas, Donnell, & Young, 2004) have suggested that its factor loadings may differ slightly for differing patient populations. In addition, no previous study has examined the sensitivity to change of the most recent (1993) BPRS version, the BPRS–E. The current study addresses these deficits in the literature by reporting diagnosis-specific sensitivity to change for the BPRS–E at the subscale and item levels, using factors recently derived (Thomas et al., 2004) for an inpatient population classified as SPMI (cf. Burlingame et al., 2005; Earnshaw, Rees, Dunn, & Burlingame, 2005).

In psychotherapy outcome, sensitivity to change can be viewed as the degree to which a

measure is likely to reflect changes that occur following a mental health intervention, for example, psychotherapy, medication, hospitalization, and so forth (Lambert & Hill, 1994). Sensitivity to change of an outcome measure is directly related to the instrument's ability to do what it purports to do—measure individual change over time. Therefore, the concept of sensitivity to change is best conceptualized as an issue of construct validity, with the most convincing evidence being documentation of longitudinal within-subject changes on the measure of interest after an effective intervention (Kirshner & Guyatt, 1985; Maruish, 2004).

Although sufficient psychometric evidence of other BPRS-E properties has accumulated (e.g., reliability, validity, and factor structure), less is known about the measure's sensitivity to change, especially on an item-by-item and factor-by-factor level. Accordingly, this study's purpose was to examine the clinical utility and psychometric soundness of the BPRS-E to detect treatment effects by evaluating (a) the sensitivity to change of each individual item, including identification of items that do not reflect changes in symptom status; (b) the usefulness of previously identified factors as general indicators of change (i.e., factor sensitivity to change); and (c) the usefulness of those factors as indicators of change for clients with specific diagnoses (i.e., diagnosis-specific sensitivity to change). More specifically, although the BPRS-E total score has been shown to be sensitive to change, little is known about how individual items contribute to this sensitivity, and even less is known about potential interactions between diagnosis and sensitivity to change of items and subscales.

## Method

### Participants

A total of 786 protocols were obtained from a database for 223 adult psychiatric inpatients at the Utah State Hospital (USH) in the intermountain west. Participant ratings were drawn from a set of patients who received services at the hospital between 1999 and 2001 and who met criteria for at least one major Axis I disorder as classified by the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994). Participants'

data were accepted into the study on the basis of the following three criteria: (a) the patient was administered both an admission and a discharge BPRS-E by a trained independent rater; (b) subsequent administrations occurred at 90-day intervals (plus or minus 1 week); and (c) no more than 1 item per BPRS subscale factor was missing and no more than 2 of the total 24 BPRS-E items were missing.

This adult psychiatric sample had a mean age of 39 ( $SD = 14.57$  years) with a range of 22 to 89 years and consisted of 131 (53%) men and 116 (47%) women. The ethnic-racial distribution was 89.2% Caucasian, 1.3% African American, 4.9% Hispanic, 2.2% American Indian, 1.3% Asian/Pacific Islander, and 1.1% were unreported. Fifty-seven percent of the sample had never been married, 26% were divorced, 11% were married, and 6% were separated, widowed, or married by common law. Patient diagnoses from this sample were grouped into the following five major categories for analytic purposes and appear to be consistent with the SPMI inpatient population census: schizophrenia (including all subtypes), 58%; bipolar disorder, 13%; comorbid mood and anxiety disorders, 17%; and delusional and other psychotic disorders, .07%. The remaining 5% include diagnoses that had too few subjects to be analyzed (e.g., AXIS II, substance-induced psychosis, etc.). Although bipolar disorder is classified as a mood disorder, it was analyzed separately because of the large number of patients presenting with this diagnosis. Owing to the limited number of personality disorders and to the fact that these typically occurred as secondary diagnoses, personality disorder diagnoses were dropped from the analyses. The mean number of hospitalizations for this sample was 1.6 ( $SD = 1.05$ , range = 1–13), with an average length of stay of 172 days ( $SD = 26.8$ , range = 32–926).

Previous research assessing sensitivity to change has often used untreated, equally distressed nonpatients as a control or comparison group (cf. Vermeersch, Lambert, & Burlingame, 2000). In these analyses, change in the treatment group on an outcome measure was contrasted with change in a comparable untreated group. However, ethical and logistic obstacles precluded the use of a similar comparison group for the SPMI population under study. Accordingly, we established a within-subject

control group for our study. More specifically, 63 of the 223 patients in our original sample were readmitted to USH after discharge, enabling a comparison between initial admission and readmission scores on the BPRS-E. The readmission group was older (mean age = 44 years,  $SD = 14.05$ ), less represented by African Americans (0%), and had a longer length of stay ( $M = 415$  days,  $SD = 372.48$ ) than the initial sample. Forty-seven percent of the sample had never married, 31.7% were divorced, 12.7% were married, and 8.6% were separated, widowed, or married by common law. Patient diagnoses were similar to the overall sample: schizophrenia (all subtypes), 59%; bipolar disorder, 15%; and mood and anxiety disorders, 17%. The remaining percentages are diagnoses with too few subjects to be analyzed.

### *Instruments*

The BPRS-E (Dingemans et al., 1995) includes 24 items addressing somatic concern, anxiety, emotional withdrawal, conceptual disorientation, guilt feelings, tension, mannerisms and posturing, grandiosity, depressive mood, hostility, suspiciousness, hallucinatory behavior, motor retardation, uncooperativeness, unusual thought content, blunted affect, excitement, disorientation, bizarre behavior, suicidality, self-neglect, motor hyperactivity, distractibility, and elevated mood.

Scoring of the BPRS-E is based on a 30-min assessment performed by a trained independent rater (Raskin, 1988). Each BPRS-E item is scored on a 7-point scale ranging from 1 to 7, with 1 = *not present* and 7 = *extremely severe*. A total score is derived by adding the scores for all items. Many item scores are based on the use of "nondirective interaction" to elicit information about the presence of symptoms (Overall & Gorham, 1962; Rhoades & Overall, 1988). However, several items (mannerisms and posturing, motor retardation, tension, emotional withdrawal, uncooperativeness, blunted affect, and excitement) are rated by the rater's observation of symptoms.

Interrater reliability on the BPRS has ranged between .67 and .88 in recent studies (Bech et al., 1988; Morlan & Tan, 1998; Tarcell & Schulz, 1988). Our experience using a standardized training protocol has resulted in a higher overall average interrater reliability of .85

(Earnshaw et al., 2005). The specific method used to administer the BPRS-E in this study is a manualized technique developed by the University of California, Los Angeles (UCLA), Department of Psychiatry and Biobehavioral Sciences and Mental Health Clinical Research Center (Ventura, Green, Shaner, & Liberman, 1993).

### *Procedure*

All adult patients admitted to the hospital were assessed using the BPRS-E within 72 hr of admission, again at subsequent 90-day intervals (within a 7-day window), and finally before discharge. Psychology staff and interns administered the BPRS-E using a semistructured interview after being trained to criterion (intraclass correlation coefficient  $\geq .80$ ) using consensus-coded tapes from the UCLA Department of Psychiatry and Biobehavioral Sciences and Mental Health Clinical Research Center (Ventura et al., 1993). Rater drift was assessed at 4-month intervals using previously unrated consensus-coded tapes and corrected through further training as needed. Ratings reflected the patient's psychiatric symptoms over the previous 2 weeks. After admission, raters were not aware of the patient's treatment status and of previous BPRS-E scores or the patient's length of stay in the hospital. When no more than 1 item per BPRS-E subscale factor was missing and no more than a total of 2 of the total 24 BPRS-E items were missing, the rounded average of the remaining items in the missing item's factor was used as a substitute value. Finally, an intake interview with the unit psychiatrist provided an opportunity to reevaluate the admission diagnoses.

### *Design and Analysis*

*Individual item sensitivity to change.* The two criteria for evaluating the degree of sensitivity of individual test items are (a) that the items change in the theoretically proposed direction following intervention and (b) that the changes measured are significantly greater in treated than in untreated individuals (Vermeersch et al., 2000). The first criterion was established using the SAS MIXED procedure (SAS Institute, 1996), a multilevel linear mod-

eling approach. In essence, this procedure allows for the statistical analysis of time series, repeated-measures data via the generation of an individual slope and y-intercept for each participant on the items, factor, and total scores. This procedure demonstrates whether the change is in the desired direction (i.e., a negative slope indicates patient improvement over time) and whether each individual item's slope is significantly different from zero (no change). Although there is a recent resurgence of discussion regarding the analysis of Likert data using standard parametric models (Harwell & Gatti, 2001; Wang, Yu, Wang, & Huang, 1999), most would agree that a priori tests of the assumptions are a requirement to ensure interpretable results. Accordingly, item-level analyses of normality were performed to assist in interpreting results from the multilevel analysis.

The second criterion for sensitivity to change (change in treatment significantly greater than change in a comparison group) was tested with a subsample of individuals who were readmitted to the hospital ( $n = 63$ ). We tested the assumption that BPRS scores for patients readmitted to the hospital would be similar to the same patients' initial admission scores. If patients' readmission scores are similar to their original admission scores, but significantly different than their discharge scores, we hypothesized that changes in BPRS scores were due to the effect of treatment. This was tested with an analysis of variance comparing all BPRS-E item scores at admission, discharge, and readmission. Alphas for all analyses were set at .05

*Factor sensitivity to change.* Factor analytic studies have shown that the BPRS-E consistently loads to four or five different factors when used with a general clinical patient population (Burger et al., 1997; Dingemans et al., 1995; Long & Brekke, 1999; Morlan & Tan, 1998; Mueser et al., 1997; Ownby & Seibel, 1994). The most commonly cited factors are Thought Disturbance, Anxiety-Depression, Withdrawal, Hostile-Suspicious and Activity/Mania (see Table 1). A recent study by Thomas et al. (2004), using the same USH database, revealed that 16 of the 24 BPRS-E items loaded on the following four factors: Thought Disturbance, Animation, Mood Disturbance, and Apathy (see Table 2). Because of the similarity of these factors to those found in previous studies (see Table 1) and to provide the best match to

Table 1  
*Common Factor Dimensions of the Brief Psychiatric Rating Scale—Extended (BPRS-E)*

Factor	BPRS-E items
Thought disturbance/positive symptoms	8. Grandiosity (9. Suspiciousness) 10. Hallucinations 11. Unusual thought content 15. Conceptual disorganization
Anxiety-depression	1. Somatic concern 2. Anxiety 3. Depression 5. Guilt
Withdrawal, negative symptoms	14. Disorientation 16. Blunted affect 17. Emotional withdrawal 18. Motor retardation
Hostile-suspicious/paranoid	6. Hostility (9. Suspiciousness) (20. Uncooperativeness)
Activity/mania	19. Tension (20. Uncooperativeness) 21. Excitement 24. Mannerisms and posturing

*Note.* Items in parentheses were grouped with different factors in other studies.

the USH population, these factors were used in this study. As with the individual item sensitivity analyses, we evaluated these four factors and the BPRS total score direction of change and nonzero slopes.

*Diagnosis-specific sensitivity to change.* The BPRS-E total score, factor scores, and the remaining eight items were tested to see if they were more sensitive to change for one diagnosis versus another (e.g., schizophrenia vs. bipolar disorder). A hierarchical linear modeling procedure similar to the one used for the individual item sensitivity and the factor sensitivity analyses described above was used. The four diagnostically grouped categories (i.e., schizophrenia, bipolar disorder, mood/anxiety, and psychotic/delusional) were evaluated as independent variables in six comparisons using patient change as the dependent variable. We used Bonferroni corrections ( $p < .008$ ) to correct for familywise error rate inflation.

Table 2  
*Utah State Hospital Factors (Thomas, Donnell, & Young, 2004)*

Factor name	BPRS-E item number
Thought disturbance	11, 8, 9, 10
Animation	23, 21, 19, 7
Mood disturbance	3, 4, 5, 2
Apathy	17, 16, 18, 13

Note. BPRS-E = Brief Psychiatric Rating Scale—Extended.

## Results

Descriptive analyses of each wave indicated little restriction of range in rater values assigned on the 24 items. Specifically, minimum and maximum values were reached for the majority (75%) of items. Modest restriction (1 point) was found with minimum values for 4 items (Items 6, 8, 9 and 11) and with maximum values for 2 items (Items 13 and 21). More important, 2 of the 24 items violated standard metrics used to assess distribution normality. Specifically, 1 item (Item 24) produced a skewed distribution (skewness = 4.2), and two items (Items 7 and 24) failed the kurtosis standard (i.e., kurtosis = 7.3 and 20.5, respectively, by item).

### *Individual Item Sensitivity to Change*

Results of the data analysis for the psychiatric hospital population indicated that all 24 BPRS-E item slopes met the first criterion for change sensitivity in that they changed in the theoretically proposed direction (i.e., a negative slope suggested patient improvement over time). However, only 22 of 24 items demonstrated a slope that was significantly different from zero (see Table 3). The two items that failed to meet the first criterion for change sensitivity assessed elevated mood (Item 7) and mannerisms and posturing (Item 24). The most parsimonious reason lies in the descriptive analysis results showing a nonnormal distribution. Specifically, a nonnormal distribution would hamper sensitivity to change indices as they rely on predictions from the general linear model. Furthermore, it is possible that a floor effect differentially influenced these two items in that their average intercept scores (1.4241 and 1.2067, respectively) were significantly lower than the other 22 items at intake.

### *Readmission Comparison*

The results of the analysis of variance confirmed significant differences between admission, discharge, and readmission on all but 2 items of the BPRS-E for the 61 patients who were readmitted to the hospital. Post hoc analyses demonstrated that there were significant differences between admission and discharge groups on all but Items 19 and 24. These same analyses confirmed that there were no differences between the initial admission and readmission scores on any of the 24 items. This supports the assumption that the observed changes in slope are most likely due to the effect of treatment.

### *Factor Sensitivity to Change*

As expected from the results of the item sensitivity analysis, each of the four factors demonstrated both change in the theoretically proposed direction and slopes that were significantly different from zero. The total BPRS-E

Table 3  
*Average Slopes of Change for Brief Psychiatric Rating Scale—Extended Items*

Item	Slope
1. Somatic concern	-0.0019*
2. Anxiety	-0.0025*
3. Depression	-0.0041*
4. Suicidality	-0.0036*
5. Guilt	-0.0013*
6. Hostility	-0.0021*
7. Elevated mood	-0.0003
8. Grandiosity	-0.0021*
9. Suspiciousness	-0.0044*
10. Hallucinations	-0.0031*
11. Unusual thought	-0.0069*
12. Bizarre behavior	-0.0029*
13. Self-neglect	-0.0021*
14. Disorientation	-0.0021*
15. Conceptual disorganization	-0.0032*
16. Blunted affect	-0.0021*
17. Emotional withdrawal	-0.0018*
18. Motor retardation	-0.0010*
19. Tension	-0.0008*
20. Uncooperativeness	-0.0014*
21. Excitement	-0.0020*
22. Distractibility	-0.0010*
23. Motor hyperactivity	-0.0010*
24. Mannerisms and posturing	-0.0004

\*  $p < .01$ .

Table 4  
Average Slopes of Change for Brief Psychiatric Rating Scale—Extended Factors and Total Score

Factor	Slope
Thought disturbance	-0.0217*
Animation	-0.0005*
Mood disturbance	-0.0169*
Apathy	-0.0097*
Total score	-0.0935*

\*  $p < .01$ .

score also evidenced both change in the expected direction and a significant slope as compared with zero. The latter suggested that over a 90-day period there was an average change of nine total score points. Table 4 contains the slope estimates (i.e., average growth rate) for each factor and for the total BPRS-E score.

#### Diagnosis-Specific Sensitivity to Change

Factor, item, and total score sensitivities to change were not found to vary significantly across diagnostic groups in any case (see Table 5). This finding indicates that the BPRS-E total score, factor scores, and the eight items not included in the four factors were not better predictors of overall patient change for one diagnosis than for another. In other words, the rate

of patient improvement did not appear to be associated with patient diagnosis.

#### Discussion

The results support 22 of 24 BPRS-E items, each of the four factors, and the total score as being sensitive to change. These findings underscore the usefulness of the BPRS-E as a repeat measure for the SPMI population (Burlingame & Lee, 2004; Lukoff et al., 1986; Morlan & Tan, 1998; Rhoades & Overall, 1988). In addition to robust psychometric properties, the BPRS-E appears to be sensitive to change at the item and factor level for this population. Heretofore, practitioners have relied on individual items and factors in tracking patient symptom distress and change with little empirical support. Although composite scores are unquestionably more reliable and valid, the change found herein provides one more measure of confidence for the clinical value of the BPRS-E at the item and factor level.

Surprisingly, diagnosis was unrelated to change at all levels of analysis (item, factor, and total score). One would expect larger gains on the Mood Disturbance factor for the bipolar and mood/anxiety diagnostic groupings. Similarly, greater change would be expected on the Thought Disturbance factor for the schizo-

Table 5  
Differences in Sensitivity to Change by Diagnosis

Factor or item	Schiz vs. bipolar	Schiz vs. mood/anx	Schiz vs. delusional	Bipolar vs. mood/anx	Bipolar vs. delusional	Mood/anx vs. delusional
Total score	1.67	0.78	0.33	-0.86	-0.99	-0.25
Factors						
TD	0.96	-0.86	-0.09	-1.39	-0.79	0.51
AN	2.00	-0.06	1.24	-1.74	-0.63	1.07
MD	0.44	2.49	0.89	1.38	0.36	-0.94
AP	0.51	0.97	-0.29	0.23	-0.60	-0.91
Item						
1	2.15	-0.75	0.22	-2.31	-1.47	0.69
6	-1.19	-0.59	0.74	0.63	1.44	1.01
12	-0.91	0.01	1.21	0.77	1.59	1.01
14	0.18	1.76	0.76	1.01	0.42	-0.56
15	1.47	-0.96	0.16	-1.89	-0.98	0.79
20	-0.65	-0.60	-0.11	0.15	0.40	0.30
22	0.55	0.75	0.91	-0.00	0.24	0.28
24	0.04	-0.76	0.61	-0.54	0.42	1.03

Note.  $t$  values are reported;  $n = 223$ . Schiz = schizophrenia; Mood/anx = mood or anxiety disorder; Delusional = psychotic/delusional disorders. TD = Thought Disorder; AN = Animation; MD = Mood Disturbance; AP = Apathy.

phrenic group. Neither was found. The most plausible explanation lies in the reliability of the diagnostic information. Indeed, review of medical records reveals dynamic diagnoses, with the same patient classified in different groupings depending on an episode of care or the course of his or her illness. Although diagnostic imprecision may explain some of the variability, another consideration may be the nature of diagnostic presentation in this chronic, treatment-resistant population.

A second and equally plausible explanation for the absence of a relationship between subscale-item change and diagnosis may lie in sample size. The bipolar and mood/anxiety groupings contained 26 and 34 patients (respectively), which may be insufficient to tease apart differential symptom presentation and change patterns. Indeed, the largest difference between diagnostic groupings (see Table 5) was with the comparison between the mood/anxiety and schizophrenic groups on the Mood Disturbance factor. Clearly, further work in this area is warranted before accepting the finding of no interaction between subscale-item change and patient diagnosis. Finally, the patients studied herein received an unusually long dose of inpatient care given comparable extant studies. Although the reduced scores were reflective of the effects of inpatient treatment, there are certainly other nonspecific factors that may explain some of this change. For instance, adequate nutrition, hydration, and appropriate social interactions cannot be discounted as important explanations of change. Furthermore, time alone for some patients (e.g., those with bipolar disorder) may have led to some symptom remission. In combination, these nonspecific factors may have attenuated differences that may exist between diagnostic groupings.

One strength of this study is the high degree of interrater reliability achieved. We attribute the high level of agreement with the use of a "gold standard" (Ventura et al., 1993) and periodic reliability checks to control rater drift. As described elsewhere (Earnshaw et al., 2005), USH is principally focused on service delivery, with few resources devoted to research. However, with modest commitment of resource (i.e., psychology interns) and ongoing rater calibration, our experience suggests that reliable BPRS-E ratings are well within reach of similar service delivery systems. More important, a by-

product of an interrater reliability protocol is clinical staff (psychiatry, psychology, social work) confidence in the meaning and integrity of outcome assessments—a must if outcome scores are to be used for decision making.

Several limitations affect the strength of this study's findings. The inability to use a no-treatment control group would have been more problematic if more of the items and factors had been found to be insensitive to change. However, the addition of the within-subject comparison group lends support to the conclusion that the improvement in BPRS-E scores is likely a result of actual treatment effects. Furthermore, although we have no reason to believe that our hospital differs on any important demographic from other state hospitals, these results would be clearly more generalizable if replicated. If estimates are replicated, a logical next step might be the use of longitudinal factor analytic procedures to test for factors related to patient change. More specifically, the focus of longitudinal factor analysis is to determine if there is a common structure to the change exhibited across a particular cluster of patients.

Several other directions for future research are suggested by this study's results. Our most recent work has indicated that a very small number of outcome instruments are recurrently used in the literature to assess change in the SMPI population (e.g., Global Assessment of Function, Positive and Negative Symptoms, and the Scale for the Assessment of Negative Symptoms; cf. Burlingame et al., 2005). A useful next step would be to empirically evaluate each of the most commonly used instruments in a similar manner to determine differential sensitivity to change. If one instrument were found to be superior in capturing the limited change produced from the treatment provided to this population, obvious clinical, administrative, and research implications would follow.

If reliable estimates and models for average change were available for the BPRS-E, one could use these to track patient change at an aggregate level. A very recent trend in outcome research is to establish average change trajectories for patient properties using instruments with well-known psychometric properties and large change samples (Lambert, Gregerson, & Burlingame, 2004). Knowledge of average change trajectories allows one to determine when symptom deviation (e.g., deterioration vs.

improvement) is within allowable clinical limits and when the clinician should be concerned (Finch, Lambert, & Schaalje, 2001; Lambert et al., 2002). Such information is currently being used by some managed care companies to calibrate outpatient care on a session-by-session basis (Brown, Burlingame, Lambert, Jones, & Vacarro, 2001). The utility of this method for inpatient care has yet to be assessed.

Finally, most of the extant BPRS-E change studies focus on the inpatient status of a SMPI patient (cf. Burlingame et al., 2005; Lachar et al., 2004). Given the high financial, personal, and societal cost of relapse, change trajectories on the BPRS-E (or other instruments with empirically established sensitivity to change) could be extended to tracking the discharge status of the SMPI population, that is, to outpatient settings. Our own clinical work in this area has convinced us of the difficulty in tracking this patient population after discharge. More specifically, our earlier attempts at tracing these patients across a continuum of care (Burlingame et al., 2002) were hampered by a number of factors. For instance, the transient nature of this population makes it difficult to obtain meaningful serial measures of outcome. Transience also increases the likelihood that these patients will present at multiple inpatient and outpatient settings for treatment. The use of diverse outcome measures by different publicly funded outpatient settings—a common practice throughout North America—exacerbates the difficulty of accumulating comparable outcome assessments.

The lack of comparable outcome assessment data for this population has led at least one state—Utah—to provide outpatient substance and mental health facilities (both in- and outpatient) with a common self-report outcome measure for adults as well as children and adolescents. Although multisource outcome assessment with this population is ideal (cf. Burlingame et al., 2005), Utah's move to provide a common self-report measure ensures comparable outcome profiling irrespective of which mental health facility the patient presents at for treatment. It behooves researchers, clinicians, and policymakers alike to systematically tackle this challenging problem given the costs associated with the “revolving door” phenomena common with this population.

## References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychological Association. (2005). *Policy statement on evidence-based practice in psychology*. Retrieved December 23, 2005, from [www.apa.org/practice/ebpstatement.pdf](http://www.apa.org/practice/ebpstatement.pdf)
- Bech, P., Larsen, J., & Andersen, J. (1988). The BPRS: Psychometric developments. *Psychopharmacology Bulletin*, 24(1), 118–121.
- Borison, R. L., Sinha, D., Haverstock, S., McLarnon, M., & Diamond, B. (1989). Efficacy and safety of tiospirone vs. haloperidol and thioridazine in a double-blind, placebo-controlled trial. *Psychopharmacology Bulletin*, 25(2), 190–193.
- Brown, G., Burlingame, G., Lambert, M., Jones, E., & Vacarro, J. (2001). Pushing the quality envelope: A new outcomes management system. *Psychiatric Services*, 52, 925–934.
- Burger, G., Calsyn, R., Morse, G., Klinkenberg, W., & Trusty, M. (1997). Factor structure of the Expanded Brief Psychiatric Rating Scale. *Journal of Clinical Psychology*, 53(5), 451–454.
- Burlingame, G., Dunn, T., Chen, S., Lehman, A., Axman, R., Earnshaw, D., & Rees, F. M. (2005). Selection of outcome assessment instruments for patients with severe and persistent mental illness (SPMI). *Psychiatric Services*, 56, 444–451.
- Burlingame, G. M., Earnshaw, D., Hoag, M., Barlow, S. H., Richardson, E. J., Donnell, A. J., & Villani, J. (2002). A systematic program to enhance clinician group skills in an inpatient psychiatric hospital. *International Journal of Group Psychotherapy*, 52(4), 555–587.
- Burlingame, G., & Lee, J. (2004). Brief Psychiatric Rating Scale. In B. Strauss & J. Schumacher (Eds.), *Clinical interviews and ratings* (pp. 50–55). Goettingen, Germany: Hogrefe.
- Dingemans, A., Linszen, D., Lenior, M., & Smeets, R. (1995). Component structure of the Brief Psychiatric Rating Scale (BPRS-E). *Psychopharmacology*, 122, 263–267.
- Earnshaw, D., Rees, F., Dunn, T., & Burlingame, G. (2005). Implementing a multisource outcome assessment protocol in a state psychiatric hospital: A case study. *Psychiatric Services*, 56, 411–413.
- Feighner, K., Merideth, C., & Claghorn, J. (1984). Multicenter placebo-controlled evaluation of nomifensine treatment in depressed outpatients. *Journal of Clinical Psychiatry*, 45(4), 47–51.
- Finch, A., Lambert, M., & Schaalje, B. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration in an early warning system. *Clinical Psychology and Psychotherapy*, 8, 231–242.

- Gur, R. E., Mozley, P. D., Resnick, S. M., Levick, S., Erwin, R., Saykin, A., & Gur, R. C. (1991). Relations among clinical scales in schizophrenia. *American Journal of Psychiatry, 148*, 472–478.
- Harwell, M., & Gatti, G. (2001). Rescaling of ordinal data to interval data in educational research. *Review of Educational Research, 71*(1), 105–131.
- Hedlund, J., & Viewberg, B. (1980). The Brief Psychiatric Rating Scale (BPRS): A comprehensive review. *Journal of Operational Psychiatry, 11*, 49–65.
- Inch, R., Crossley, M., Keegan, D., & Thorarinson, D. (1997). Use of the Brief Psychiatric Rating Scale to measure success in psychosocial day programs. *Psychiatric Services, 48*, 1195–1197.
- Institute of Medicine. (2005). *Improving the quality of health care for mental and substance-use conditions: Quality Chasm Series (Free executive summary)*. Retrieved December 23, 2005, from <http://www.iom.edu/CMS/3809/19405/30836.aspx>
- Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of Chronic Diseases, 38*, 27–36.
- Lachar, D., Espadas, A., & Bailley, S. (2004). The brief psychiatric rating scale: Contemporary applications. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (3rd ed., pp. 153–190). Mahwah, NJ: Erlbaum.
- Lambert, M., Gregerson, A., & Burlingame, G. (2004). The Outcome Questionnaire-45. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (3rd ed., pp. 191–234). Mahwah, NJ: Erlbaum.
- Lambert, M., & Hill, C. (1994). Assessing psychotherapy outcomes and processes. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 309–330). New York: Wiley.
- Lambert, M., & Ogles, B. (2004). The efficacy and effectiveness of psychotherapy. In M. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York: Wiley.
- Lambert, M., Whipple, J., Bishop, M., Vermeersch, D., Gray, G., & Finch, A. (2002). Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clinical Psychology and Psychotherapy, 9*, 149–164.
- Linszen, D. H., Dingeman, P. M., & Lenior, M. E. (1994). Cannabis abuse and the course of recent-onset schizophrenic disorders. *Archives of General Psychiatry, 51*, 273–279.
- Long, J., & Brekke, J. (1999). Longitudinal factor structure of the Brief Psychiatric Rating Scale in schizophrenia. *Psychological Assessment, 11*, 498–506.
- Lukoff, D., Liberman, R., & Nuechterlein, K. (1986). Symptom monitoring in the rehabilitation of schizophrenic patients. *Schizophrenia Bulletin, 12*, 578–602.
- Marder, S., Mintz, J., Van Patten, T., Lebell, M., Wirsching, W., & Johnston-Crank, K. (1991). Early prediction of relapse in schizophrenia: An application of receiver operating characteristic (ROC) methods. *Psychopharmacology Bulletin, 27*, 79–82.
- Maruish, M. (2004). *The use of psychological testing for treatment planning and outcomes assessment* (3rd ed.). Mahwah, NJ: Erlbaum.
- Morlan, K., & Tan, S. (1998). Comparison of the Brief Psychiatric Rating Scale and the Brief Symptom Report. *Journal of Clinical Psychology, 54*(7), 885–894.
- Mueser, K., Curran, P., & McHugo, G. (1997). Factor structure of the Brief Psychiatric Rating Scale in schizophrenia. *Psychological Assessment, 9*, 196–204.
- Newcomer, J., Faustman, W., Yeh, W., & Csernansky, J. (1990). Distinguishing depression and negative symptoms in unmedicated patients with schizophrenia. *Psychiatry Research, 31*, 243–250.
- Overall, J., & Gorham, D. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports, 10*, 799–812.
- Owaby, R., & Seibel, H. (1994). A factor analysis of the Brief Psychiatric Rating Scale in an older psychiatric population: Exploratory and confirmatory analyses. *Multivariate Experimental Clinical Research, 10*(3), 145–156.
- Packer, S., Husted, J., Cohen, S., & Tomlinson, G. (1997). Psychopathology and quality of life in schizophrenia. *Journal of Psychiatry Neurosciences, 22*, 231–234.
- Raskin, A. (1988). Discussion: Recent developments in ascertainment and scaling of the BPRS. *Psychopharmacology Bulletin, 24*(1), 122–124.
- Rhoades, H., & Overall, J. (1988). The semistructured BPRS interview and rating guide. *Psychopharmacology Bulletin, 24*(1), 101–104.
- SAS Institute. (1996). *SAS/STAT: Changes and enhancements through Release 6.11*. Cary, NC: Author.
- Subotnik, K., & Nuechterlein, K. (1988). Prodromal signs and symptoms of schizophrenia relapse. *Journal of Abnormal Psychology, 97*, 405–412.
- Swett, C. (1995). Symptom severity and number of previous psychiatric admissions as predictors of readmission. *Psychiatric Services, 46*, 482–485.
- Tarcell, J., & Schulz, S. (1988). Nursing assessment using the BPRS: A structured interview. *Psychopharmacology Bulletin, 24*(1), 105–111.

- Thieman, S., Csernansky, J., & Berger, P. (1987). Rating scales in research: The case of negative symptoms. *Psychiatry Research, 20*, 47-55.
- Thomas, A., Donnell, A., & Young, T. (2004). Factor structure and differential validity of the expanded Brief Psychiatric Rating Scale. *Assessment, 11*(2), 177-187.
- Van der Does, A., Linszen, D., Dingemans, P., Nugter, M., & Scholte, W. (1993). A dimensional and categorical approach to the symptomatology of recent-onset schizophrenia. *Journal of Nervous and Mental Disease, 181*, 744-749.
- Varner, R., Chen, R., Swann, A., & Moeller, F. (2000). The Brief Psychiatric Rating Scale as an acute inpatient outcomes measurement tool: A pilot study. *Journal of Clinical Psychiatry, 61*, 418-421.
- Ventura, J., Green, M., Shaner, A., & Liberman, R. (1993). Training and quality assurance with the Brief Psychiatric Rating Scale: "The Drift Busters." *International Journal of Methods in Psychiatric Research, 3*, 221-244.
- Vermeersch, D., Lambert, M., & Burlingame, G. (2000). Outcome questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74*(2), 242-261.
- Wang, S., Yu, M., Wang, C., & Huang, C. (1999). Bridging the gap between the pros and cons of treating ordinal scales as interval scales from an analysis point of view. *Nursing Research, 84*(4), 226-229.

Received December 29, 2004

Revision received January 11, 2006

Accepted February 8, 2006 ■

## E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!