

Observed Ward Behavior Strongly Associated with Independent Living Skills: An Analysis of Convergent and Criterion-Related Validity of the NOSIE and the ILSI

Elizabeth A. Cook · Charlie A. Davidson ·
Jeffrey R. Nolting · William D. Spaulding

© Springer Science+Business Media, LLC 2010

Abstract Ratings of ward behavior and skill acquisition are important criteria for monitoring an individual's recovery trajectory during inpatient psychiatric rehabilitation. This study compared data from the Nurses' Observational Scale for Inpatient Evaluation (NOSIE) and the Independent Living Skills Inventory (ILSI) to identify relationships between ward behaviors and independent living skills and to support the convergent validity of these instruments. The criterion-related validity of the instruments in their relationships with time in program (TIP) and number of previous hospitalizations (#HOSP) was also investigated. NOSIE "positive" subscales were more strongly related to living skills than NOSIE "negative" subscales. The NOSIE and the ILSI predicted TIP and postdicted #HOSP equally well. Implications for assessment and recommendations for future research are discussed.

Keywords Ward behavior, Independent living skills · NOSIE · ILSI · Convergent and criterion-related validity

Introduction

Ratings of ward behavior and skill acquisition are important criteria by which an individual's recovery trajectory may be

monitored. Inpatient staff involved in providing psychiatric rehabilitation services observe and quantify patient ward behavior for the purposes of informing treatment decisions and determining preparedness for discharge (Paul et al. 1986). Furthermore, ward behavior is predictive of outcome measures such as length of hospitalization (Hopko et al. 2001; Lyons et al. 1997; Burdock et al. 1960; Aumack 1962) and number of hospital admissions (Lyons et al. 1997; Burdock et al. 1960).

Traditional ward behavior instruments have assessed domains such as hygiene and grooming, cooperativeness, communication, affect and mannerisms, interpersonal relations and social interest, hostility, irritability, and aggressiveness, behaviors indicative of psychiatric symptoms, and task competency (Paul et al. 1986; Honigfeld and Klett 1965; Lorr et al. 1960; Aumack 1962). These assessments are theoretically similar to scales of independent living ability, which typically measure domains such as physical health care, social skills, social supports and relationships, personal safety, and work skills (Dickerson 1997). Despite this similarity, few studies have examined the extent to which observed behaviors on the ward relate to living skills that are necessary for successful community transition and "real world" functioning. One purpose of this study is to compare data from the Nurses' Observational Scale for Inpatient Evaluation, 30-item version (NOSIE), a ward behavior instrument, and the Independent Living Skills Inventory (ILSI), a functional assessment instrument, in order to explore this relationship.

The NOSIE is a widely used observational measure that identifies the behavioral strengths and weaknesses of inpatients (Honigfeld and Klett 1965). It has been used for the purposes of identifying acute psychosis and medication effects, and it is also a valuable qualitative assessment of behavioral change (Spaulding et al. 2003;

E. A. Cook (✉) · C. A. Davidson · J. R. Nolting ·
W. D. Spaulding
Department of Psychology, University of Nebraska-Lincoln,
238 Burnett Hall,
Lincoln, NE 68588-0308, USA
e-mail: psych.cook@gmail.com

J. R. Nolting
Fulton State Hospital,
Fulton, MO, USA

Volavka et al. 2004; Baker et al. 1983; Rada and Kellner 1976; Hollister et al. 1975; Bankier and Mathewson 1972). In addition to a Total Assets score, its thirty items form six subscales: Social Competence (COM), Social Interest (SOC), Personal Neatness (NEA), Irritability (IRR), Manifest Psychosis (PSY), and Retardation (RET). The first three subscales (COM, SOC, and NEA) reflect “positive” characteristics or assets, while IRR, PSY, and RET reflect “negative” characteristics or liabilities. Previous studies have provided evidence for the NOSIE’s construct validity (Carbonara et al. 1983; Farrell and Mariotto 1982; Lentz et al. 1971). The NOSIE has good inter-rater reliability, with reliability coefficients ranging from .68 to .76 (Lyall et al. 2004).

The ILSI (Sanchez 1986) is a functional assessment used to determine a person's treatment and support needs in the post-discharge environment. It has been utilized primarily in inpatient populations (Sanchez 1986; Keefe et al. 2006a, b; Menditto et al. 1999), and one study has used it in an outpatient population (Sanchez 1986). It was developed because many of its predecessors either inappropriately emphasized psychiatric symptoms and skill deficits or were too narrowly or broadly focused on social skills to be used as practical tools for clinical decision making (Sanchez 1987). It has been suggested that the ILSI is a measure of “real world” living skills, and it is correlated with cognitive abilities (Keefe et al. 2006a, b). It is sensitive to treatment effects, as measured by treatment team decisions and discharge placement, and can be readily administered by moderately experienced professionals and paraprofessionals without extensive special training (Sanchez 1987). The ILSI is comprised of seventy items reflecting ten domains relevant to independent community living: Personal Management (PE), Hygiene (HY), Clothing (CL), Basic Skills (BA), Interpersonal Skills (IN), Home Maintenance (HO), Money Management (MO), Cooking (CO), Resource Utilization (RE), and General Occupational Skills (GE). A total score may be calculated by summing the means of all subscales (Sanchez 1987). The ILSI is multidimensional in that its rating scales take into account both competence and actual performance (Sanchez 1987). It is a reliable instrument, as indicated by one study which found that subscale coefficient alphas fell above or very close to .70 (Sanchez 1987) and another study which found that subscale coefficient alphas fell above or very close to .90 (Menditto et al. 1999). In addition, subscale intraclass correlation coefficients have been shown to average .71 (Menditto et al., 1999).

Theoretically, NOSIE and ILSI measurement constructs overlap in their behavioral assessment of functional abilities with the goal of quantifying rehabilitation progress. Nevertheless, they diverge in many ways: the ILSI is “strength-based” while the NOSIE also includes psychiatric

symptoms in “negative” subscales; the ILSI focuses on specific skill areas while the NOSIE assesses positive and negative behaviors more generally; the ILSI is thorough and accommodates multiple raters while the NOSIE is optimized for practicality and frequent assessment; and the ILSI is intended to reflect “real world” skills while the NOSIE measures current behavior (Sanchez 1987). Although these areas of divergence are interrelated and far from polar, they make the two instruments clinically valuable in tandem and differentially useful in outcomes research.

Comparison of the NOSIE and the ILSI will provide evidence for convergent validity of these measures’ assessment of similar behaviors. Convergent validity is defined as strong interrelationships between measures of similar constructs (Sattler 2008). According to the Outcomes Roundtable, a key feature of instruments that measure treatment outcome is that they have “demonstrated validity” (Smith et al. 1997, p. 1034). The NOSIE and the ILSI are valuable outcome instruments, so it is imperative that they possess adequate validity in order to effectively serve this purpose. In addition, recent research efforts have attempted to organize and evaluate scales of functional abilities in schizophrenia to improve, among other things, comparability across studies (Green et al. 2008; McKibbin et al. 2004; Keefe et al. 2006a). Thus, comparison of the NOSIE and the ILSI addresses this goal by promoting a deeper understanding of the interrelationships between two measures that are frequently used in clinical and research settings.

This study is a partial replication of Sanchez (1987), who analyzed concurrent and construct validity between NOSIE and ILSI scales in an inpatient population at the Lincoln Regional Center (LRC), a Nebraska state psychiatric hospital. Through a series of multiple regression and correlational analyses using NOSIE and ILSI subscales, Sanchez found that significant and unique NOSIE predictors primarily came from NOSIE positive subscales. Also, contrary to her hypothesis, NOSIE negative subscales typically did not demonstrate significant inverse relationships with ILSI subscales.

In the present study, we have partially replicated these analyses at the same institution over 5 years later with a larger sample. This replication is important for several reasons. First, the population at the LRC has changed over time due to changes in mental health board policy. For example, the number of referrals to outpatient services instead of inpatient settings increased over time, resulting in a population with more severe symptomatology and functional deficits being admitted to the LRC. This study will either provide converging evidence for the relationships found in Sanchez (1987) or will reveal different relationships, which could indicate that population and time

have important implications for associations between the NOSIE and the ILSI.

Second, we wish to address a methodological limitation from Sanchez (1987), in which all data were analyzed in raw form using parametric statistics. The assumption that all variables are normally distributed must be met when regression and correlational analyses are being used for inferential purposes rather than simply descriptive reasons (Prymachuk and Richards 2007). Furthermore, research has demonstrated that outliers can bias any statistical analysis that is based upon sample means and variances, such as correlation or regression (Prymachuk and Richards 2007; High 2000). Our data were carefully cleaned before analyses in order to obtain the most accurate conclusions possible. We present results from analyses after data cleaning and also note any significant discrepancies with analyses that were performed with raw data, as suggested by Stevens (1984).

Third, we address the criterion-related validity of the NOSIE and the ILSI in their predictive relationships with Time in Program (TIP) and their postdictive relationships with Number of Previous Hospitalizations (#HOSP). Efficient rehabilitation and discharge as well as reduced re-hospitalizations are important goals for both providers and people in the recovery process, and thus the amount of time spent in an inpatient treatment program and number of previous hospitalizations may be considered useful proxies for outcome or recovery. Although number of previous hospitalizations is being postdicted in this study, it has been shown to be predictive of number of future hospitalizations (Carr et al. 2008), and may be useful for clinical decision making, as it may indicate rehabilitation non-responsiveness. In addition, we have chosen these criterion variables because they are related to ward behavior (Lyons et al. 1997; Hopko et al. 2001; Burdock et al. 1960; Aumack 1962).

As this study is a partial replication of Sanchez (1987), some of our hypotheses are the same. Specifically, (a) NOSIE ratings of positive subscales (Social Competence, Social Interest, and Personal Neatness) are expected to be significantly and positively related to ILSI total, and (b) NOSIE ratings of negative subscales (Manifest Psychosis, Irritability, and Retardation) are expected to be significantly and negatively related to ILSI total. Although this second hypothesis was not supported in Sanchez (1987), results may have been limited by non-normal raw data and small sample size, two methodological issues which we address in this study.

Additionally, we hypothesize that (c) NOSIE positive subscales will be more strongly related to ILSI total than negative subscales. According to Sanchez (1987), living skill competence appears to be strength based rather than deficit based; that is, a person's abilities rather than

functional deficits are the focus of assessment and treatment. In addition, recent research suggests that positive symptoms may not be unique predictors of functional outcomes (Smith et al. 2002; Green et al. 2000; Heinrichs et al. 2009), so NOSIE Manifest Psychosis and Irritability may not be strongly related to overall living skill ability.

It is also hypothesized that (d) NOSIE positive subscales and ILSI subscales will be negatively related to TIP (Hopko et al. 2001; Lyons et al. 1997; Burdock et al. 1960; Aumack 1962) and #HOSP (Lyons et al. 1997; Burdock et al. 1960). It is expected that NOSIE negative subscales will be positively related to TIP and #HOSP. Finally, we hypothesize that (e) NOSIE and ILSI subscales will predict TIP and postdict #HOSP equally well.

Method

Participants

One hundred sixty three participants were included in this study. The purposive sampling frame from which the data were drawn was all patients who stayed at least 3 months at the LRC's Community Transition Program between 1992 and 2004. A small number of participants had Not Guilty by Reason of Insanity/Not Responsible by Reason of Insanity (NGRI/NRRI) legal status. As the characteristics and support tactics associated with these individuals are distinct and conceptual outliers from the general inpatient population with serious mental illness (SMI) (Nolting et al. 2005), and the size of this group was too small to analyze separately, participants with NGRI/NRRI status were excluded from the analyses. Four additional cases were excluded due to missing NOSIE or ILSI data at the time point assessed. The mean age of the participants at admission to the program was 39.7 years ($SD=11.76$). The mean number of years of education was 12.5 ($SD=2.15$). Participants had spent an average of 672.2 days in the program ($SD=451.64$), and had an average of 8.3 previous hospitalizations ($SD=6.42$). Additional participant characteristics are summarized in Table 1.

Treatment Setting

The Community Transition Program (CTP) was a 40-bed unit psychiatric rehabilitation program in the Lincoln Regional Center, a state hospital. The CTP served people with severe and disabling mental illness who were under civil commitment and who could not be safely discharged to less restrictive or secure settings from conventional psychiatric inpatient units. This created a population over time that included a mix of people with years of continuous institutionalization (referred from other state hospitals) and

Table 1 Demographic and clinical characteristics (*N*=163)

Variable	M	SD	n
Sex			
Men	–	–	86
Women	–	–	77
Ethnicity			
White	–	–	144
African American	–	–	12
Hispanic	–	–	2
Asian-American	–	–	2
Other	–	–	3
Marital status (at admission)			
Single	–	–	105
Married	–	–	7
Divorced/widowed/separated	–	–	51
Axis I diagnosis			
Schizophrenia	–	–	106
Psychotic disorder NOS	–	–	1
Bipolar	–	–	7
Personality changes due to head injury	–	–	1
Impulse control disorder	–	–	1
Other	–	–	11
Second axis I diagnosis			
Substance abuse/dependence	–	–	37
PTSD	–	–	2
OCD	–	–	8
Pervasive developmental disorder	–	–	1
Anxiety disorder NOS	–	–	1
Dementia	–	–	1
Other	–	–	9
No diagnosis/missing	–	–	103
Axis II diagnosis			
Borderline	–	–	11
Schizoid	–	–	1
Paranoid	–	–	11
Antisocial	–	–	6
Histrionic	–	–	1
Mixed	–	–	1
Personality disorder NOS	–	–	30
Borderline intellectual functioning	–	–	1
Mild mental retardation	–	–	1
Other	–	–	13
No diagnosis/missing	–	–	87
COM	3.622	.834	147
SOC	20.843	6.863	147
NEA	3.404	.753	147
PSY	1.081	1.022	147
IRR	2.022	1.297	147
RET	6.109	3.614	147
TOT	6.632	1.737	147
POS	75.267	15.218	147

Table 1 (continued)

Variable	M	SD	n
NEG	13.981	9.233	147
ILSIPE	2.153	.853	112
ILSIHY	2.218	1.056	112
ILSICL	.361	.346	112
ILSIIN	3.883	1.245	113
ILSIHO	2.253	.773	111
ILSICO	2.604	.845	111
ILSIRE	3.583	.617	111
ILSIGE	12.950	7.651	99
ILSIBA	.263	.282	111
ILSIMO	6.315	3.908	111
ILSITOT	2.440	.707	98

COM NOSIE Social Competence, SOC NOSIE Social Interest, NEA NOSIE Neatness, PSY NOSIE Manifest Psychosis, IRR NOSIE Irritability, RET NOSIE Retardation, TOT NOSIE total, POS sum of NOSIE positive subscales, NEG sum of NOSIE negative subscales, ILSIPE ILSI Personal Management, ILSIHY ILSI Hygiene, ILSICL ILSI Clothing, ILSIIN ILSI Interpersonal Skills, ILSIHO ILSI Home Maintenance, ILSICO ILSI Cooking, ILSIRE ILSI Resource Utilization, ILSIGE ILSI General Occupational Skills, ILSIBA ILSI Basic Skills, ILSIMO ILSI Money Management, ILSI TOT ILSI total. Dash marks (–) indicate data that were not obtained

people with a severe “revolving door” syndrome of multiple hospitalizations and unstable community functioning. All had severe deficits in multiple domains of behavioral functioning. Although the average length of treatment varied considerably, there were three modal groups: a group that was in the program for 12 to 30 months, a slightly smaller group that was in the program for less than 12 months, and an even smaller group that was in the program for longer than 30 months (Rothmann 2005). The CTP employed a comprehensive biopsychosocial approach to psychiatric rehabilitation, including collaborative pharmacotherapy, training in social, living, occupational and illness/wellness management skills, contingency management and behavioral family therapy. Program participants generally experienced 35 to 40 h per week of active treatment and rehabilitation, provided on a highly structured daily schedule. More than 90% of the participants were discharged to community settings (e.g., community-based residential programs, supported independent living settings, independent living settings), with an overall reduction in hospital use of about 60% on long term follow-up.

Measures

NOSIE and ILSI data collected 3 months post-admission were compared.

The NOSIE was collected weekly by direct care staff (psychiatric technicians) and reported as monthly averages. Each observed ward behavior was rated on a Likert-type scale: Never = 0; Sometimes = 1; Often = 2; Usually = 3; and Always = 4.

During the period under study, the ILSI was completed at 3 months post-intake and reflected performance during the most recent month. It was collected by multiple inpatient staff (e.g., treatment team coordinators, social workers, teachers, or other relevant mental health care professionals) according to their areas of contact and expertise. ILSI items are rated on a scale of 0 to 3 according to levels of proficiency, 'No Competence' (NC), 'Partial Competence,' (PC) or 'Independent Competence' (IC).

All staff were blind to the research hypotheses, as these scales are clinical measures routinely administered as part of the CTP assessment and treatment system.

Procedures

Variables with skewness greater than ± 1 were cleaned using square root and log transformations, and were winsorized where applicable to further minimize skewness (Tukey 1962).

In order to assess the convergent validity of the NOSIE and the ILSI, bivariate and multivariate relationships were explored using Pearson's product-moment correlation and multiple regression analyses. First, each NOSIE subscale was correlated with ILSI total. In order to investigate differences between NOSIE positive and NOSIE negative subscales in their relationships with ILSI total, Steiger's *Z* test for correlated correlations was used. Signed *r* values were first included in the Steiger's *Z* test in order to determine whether NOSIE positive and NOSIE negative subscales had significantly different correlations with ILSI total. Next, absolute *r* values were entered into the Steiger's *Z* test to determine whether NOSIE positive or NOSIE negative subscales were significantly better predictors of ILSI total.

Next, multiple regression analyses were performed using a full NOSIE model and two reduced models to predict ILSI total. The full model contained all NOSIE subscales, and the reduced models contained only NOSIE positive subscales and only NOSIE negative subscales, respectively. In order to assess whether there was a significant difference in the unique predictive utility of a model containing only NOSIE positive subscales and a model containing only NOSIE negative subscales, a nested model comparison was performed using an *R*-square change test.

To assess the criterion-related validity of the NOSIE and the ILSI, Pearson's product-moment correlation and multiple regression analyses were performed using TIP and

#HOSP as criterion variables. Time in program was defined as the number of days from admission to the CTP program to discharge from the program. Number of previous hospitalizations included all mental health hospitalizations prior to the most recent admission. Multiple regression analyses using non-nested models were performed (i.e., all NOSIE subscales were included in one model and all ILSI subscales were included in the other model). Beta weights were examined to investigate the direction, strength, and extent of the relationship between NOSIE and ILSI subscales and the two criterion variables. Finally, Steiger's *Z* test for non-nested models was completed in order to determine whether NOSIE or ILSI models accounted for more variance in TIP and #HOSP.

All analyses were performed using SPSS version 17.

Results

The analyses presented below use different sample sizes due to limited availability of data and in order to maximize power. For intra-scale analyses, all available NOSIE ($n=147$) and ILSI ($n=98$) data were used. Complete data sets, which were produced by selecting cases in which no data were missing, were used for inter-scale analyses.

In order to assess the convergent validity of the NOSIE and the ILSI, Pearson's product-moment correlation coefficients were first examined. All correlations between NOSIE subscales and ILSI total were significant at the $p < .01$ level. NOSIE positive subscales were positively correlated with ILSI total, while NOSIE negative subscales were negatively correlated with ILSI total. NOSIE Social Competence had the strongest correlation with ILSI total, while Manifest Psychosis and Irritability had the smallest correlations with ILSI total.

The relationships between NOSIE positive versus NOSIE negative subscales with ILSI total were then examined using Steiger's *Z* test for correlated correlations. Results from Steiger's *Z* test using signed *r* values in the computation revealed that higher ILSI total was associated with higher values on NOSIE positive subscales, $r(80) = .624$, $p < .01$, and lower values on NOSIE negative subscales, $r(80) = -.497$, $p < .01$. These correlations with ILSI total were significantly different, $Z = 6.506$, $p < .01$. Results from Steiger's *Z* test using absolute *r* values in the computation showed that NOSIE positive subscales were better predictors of ILSI total than NOSIE negative subscales, $Z = 2.030$, $p < .05$.

Next, multivariate relationships between NOSIE subscales and ILSI total were examined. Tables 2 and 3 show correlation coefficients from each variable's correlation with ILSI total and regression weights for the various models, respectively. The full model had an

Table 2 Intercorrelations between NOSIE subscales and time in program (TIP), number of previous hospitalizations (#HOSP), and ILSI total (ILSI TOT) (N=147)

Subscale	1	2	3	4	5	6	7	8	9	10	11	12
1. TIP ^a	-	.188* ^b	-.430**	-.532**	-.235*	-.326**	.406**	.296**	.405**	-.508**	-.449**	.508**
2. # HOSP ^c		-	-.020	-.040	.147	-.081	.266*	.331**	-.060	-.092	.015	.227*
3. ILSI TOT ^d			-	.577**	.381**	.571**	-.285**	-.378**	-.498**	.602	** .624**	-.497**
4. COM				-	.423**	.755**	-.550**	-.578**	-.688**	.871**	.867**	-.788**
5. SOC					-	.425**	-.292**	-.210*	-.499**	.709**	.768**	-.417**
6. NEA						-	-.483**	-.517**	-.562**	.818**	.839**	-.662**
7. PSY							-	.515**	.314**	-.639**	-.529**	.736**
8. IRR								-	.249**	-.666**	-.507**	.830**
9. RET									-	-.708**	-.703**	.619**
10. TOT										-	.956**	-.876**
11. POS											-	-.746**
12. NEG												-

COM NOSIE Social Competence, SOC NOSIE Social Interest, NEA NOSIE Neatness, PSY NOSIE Manifest Psychosis, IRR NOSIE Irritability, RET NOSIE Retardation, TOT NOSIE total, POS sum of NOSIE positive subscales, NEG sum of NOSIE negative subscales

^a NOSIE*TIP n=82
^b TIP*#HOSP n=162
^c NOSIE*#HOSP n=79
^d NOSIE*ILSITOT n= 82
 *p<.05, **p<.01

R²=.401, F(6,75)=8.364, p<.001, with no NOSIE subscales having significant regression weights. The reduced model containing only NOSIE positive subscales had an R²=.391, F(3,78)=16.662, p<.001, with Social Compe-

Table 3 Summary of hierarchical regression analysis for NOSIE subscales predicting ILSI total (ILSI TOT) (N=82)

Variable	B	SE B	β
Step 1			
COM	.160	.153	.202
SOC	.014	.011	.132
NEA	.254	.133	.277
PSY	.060	.085	.084
IRR	-.053	.076	-.095
RET	-.024	.027	-.129
Step 2			
COM	.233	.114	.294*
SOC	.016	.010	.155
NEA	.260	.130	.284*
Step 3			
PSY	.026	.088	.036
IRR	-.158	.068	-.282*
RET	-.081	.019	-.432**

COM NOSIE Social Competence, SOC NOSIE Social Interest, NEA NOSIE Neatness, PSY NOSIE Manifest Psychosis, IRR NOSIE Irritability, RET NOSIE Retardation
 *p<.05, **p<.01

tence and Neatness having significant and unique contributions to the model. This model performed as well as the full model, R² change=-.010, F(3,75)=.431, p=.731. The reduced model containing only NOSIE negative subscales had an R²=.311, F(3,78)=11.755, p<.001, with Irritability and Retardation having significant and unique contributions to the model. This model did not perform as well as the full model, R² change=-.090, F(3,75)=3.736, p=.015.

Bivariate and multivariate relationships between NOSIE and ILSI subscales as predictors and TIP as the criterion variable were then investigated. Results from Pearson’s product-moment correlations revealed that all correlations, except for ILSI Money Management*TIP, were significant at the p<.05 level. NOSIE positive subscales and ILSI subscales were negatively correlated with TIP, and NOSIE negative subscales were positively correlated with TIP. NOSIE Social Competence had the strongest significant correlation with TIP, while ILSI Home Maintenance had the weakest significant correlation. Results from non-nested regression models revealed that the model containing only NOSIE subscales had an R²=.339, F(6,75)=6.419, p<.001, with Social Competence having a unique contribution to the model, and the model containing only ILSI subscales had an R²=.222, F(10,71)=2.031, p=.042, with no subscales having significant and unique contributions to the model. The predictive utility of the two models was compared

Table 4 Intercorrelations between ILSI subscales, time in program (TIP), and number of previous hospitalizations (#HOSP) (N=82)

Subscale	1	2	3	4	5	6	7	8	9	10	11	12
1. TIP ^a	-	.188* ^b	-.268*	-.330**	-.310**	-.334**	-.233*	-.357**	-.259*	-.350**	-.297**	-.209
2. # HOSP ^c		-	.100	.065	.023	.066	-.178	-.099	-.031	.010	-.107	-.102
3. ILSIPE			-	.750**	.702**	.277*	.324**	.382**	.358**	.417**	.370**	.337**
4. ILSIHY				-	.814**	.344**	.388**	.490**	.430**	.379**	.428**	.435**
5. ILSICL					-	.335**	.453**	.646**	.463**	.393**	.488**	.518**
6. ILSIIN						-	.420**	.316**	.365**	.454**	.306**	.202
7. ILSIHO							-	.637**	.845**	.481**	.413**	.716**
8. ILSICO								-	.655**	.490**	.684**	.671**
9. ILSIRE									-	.471**	.601**	.796**
10. ILSIGE										-	.367**	.373**
11. ILSIBA											-	.630**
12. ILSIMO												-

ILSIPE ILSI Personal Management, *ILSIHY* ILSI Hygiene, *ILSICL* ILSI Clothing, *ILSIIN* ILSI Interpersonal Skills, *ILSIHO* ILSI Home Maintenance, *ILSICO* ILSI Cooking, *ILSIRE* ILSI Resource Utilization, *ILSIGE* ILSI General Occupational Skills, *ILSIBA* ILSI Basic Skills, *ILSIMO* ILSI Money Management

^a ILSI*TIP n=82

^b TIP*#HOSP n=162

^c ILSI*#HOSP n=79

*p<.05, **p<.01

using Steiger’s Z. The correlation between these two models was $r=.479, p<.01$. The predictive utility of the two models was not significantly different, $Z=-1.191, p>.05$. See Tables 2 and 4 for the correlation coefficients resulting from each variable’s correlation with TIP, and Table 5 for the regression weights for the various models.

Next, bivariate and multivariate relationships between NOSIE and ILSI subscales as predictors and #HOSP as the criterion variable were investigated. Results from Pearson’s product-moment correlations revealed that only NOSIE Manifest Psychosis and Irritability had significant (positive) correlations with #HOSP. After examining Pearson’s correlations, the two non-nested models that were used to predict #HOSP were examined. The model containing only NOSIE subscales had an $R^2=.196, F(6,72)=2.916, p=.013$, with only Irritability having a significant and unique contribution to the model. The model containing only ILSI subscales had an $R^2=.170, F(10,68)=1.396, p=.201$. The predictive utility of the two models was compared using Steiger’s Z. The correlation between the two models was $r=.303, p<.01$. There was not a significant difference between the predictive utility of the two models, $Z=.247, p>.05$. See Tables 2 and 4 for the correlation coefficients resulting from each variable’s correlation with #HOSP, and Table 5 for the regression weights for the various models.

Results from analyses using raw, uncleaned data did not reveal any substantial differences from these results.

Discussion

The convergent and criterion-related validity of the NOSIE and the ILSI were examined in an inpatient sample after 3 months of rehabilitative treatment.

In concordance with the first hypothesis, generally, NOSIE ratings of “positive” subscales (Social Competence, Social Interest, and Personal Neatness) were significantly and positively related to ILSI total. Interestingly, this was not the case when the full NOSIE model was used to predict ILSI total. This is because NOSIE subscales were highly correlated with one another, a condition known as multicollinearity. Multicollinearity produces regression coefficients with large standard error, making them unstable and increasing the likelihood that they will not be statistically significant (Allen 2007). When NOSIE negative subscales were removed from the model, most NOSIE positive subscales again became significant predictors of ILSI total, suggesting that NOSIE positive and NOSIE negative subscales are highly interrelated. This study provides convergent evidence for the relationships between NOSIE positive subscales and the ILSI that were demonstrated by Sanchez (1987), suggesting that changes in population and time do not change these relationships substantially.

Results also supported the second hypothesis. In general, NOSIE ratings of “negative” subscales (Manifest Psychosis, Irritability, and Retardation) were significantly and negatively related to ILSI total. Similar to NOSIE positive subscales, NOSIE negative subscales were not significantly

Table 5 Summary of Simultaneous Regression Analyses for NOSIE and ILSI Subscales Predicting Time in Program (TIP) and Number of Previous Hospitalizations (#HOSP) (NOSIE, ILSI, and TIP: $n=82$; NOSIE, ILSI, and #HOSP: $n=79$)

Variable	B	SE B	β
TIP			
COM	-4.770	1.485	-.650**
SOC	-.025	.107	-.026
NEA	2.010	1.298	.237
PSY	1.612	.827	.245
IRR	-.731	.740	-.141
RET	.030	.262	.017
ILSIPE	-.033	1.271	-.004
ILSIHY	-1.107	1.195	-.188
ILSICL	1.391	3.992	.076
ILSIIN	-.948	.675	-.183
ILSIHO	.940	1.862	.116
ILSICO	-1.823	1.440	-.248
ILSIRE	-.459	2.751	-.041
ILSIGE	-.130	.117	-.151
ILSIBA	-.804	3.823	-.036
ILSIMO	.106	.320	.066
#HOSP			
COM	.458	.321	.329
SOC	.032	.023	.169
NEA	-.087	.276	-.055
PSY	.195	.178	.156
IRR	.420	.164	.426**
RET	.017	.056	.052
ILSIPE	.180	.262	.126
ILSIHY	.000	.239	.000
ILSICL	.086	.807	.025
ILSIIN	.183	.140	.183
ILSIHO	-1.177	.378	-.759**
ILSICO	.095	.288	.069
ILSIRE	1.342	.552	.628*
ILSIGE	.004	.024	.026
ILSIBA	-1.476	.760	-.349
ILSIMO	.003	.063	.011

COM NOSIE Social Competence, SOC NOSIE Social Interest, NEA NOSIE Neatness, PSY NOSIE Manifest Psychosis, IRR NOSIE Irritability, RET NOSIE Retardation, ILSIPE ILSI Personal Management, ILSIHY ILSI Hygiene, ILSICL ILSI Clothing, ILSIIN ILSI Interpersonal Skills, ILSIHO ILSI Home Maintenance, ILSICO ILSI Cooking, ILSIRE ILSI Resource Utilization, ILSIGE ILSI General Occupational Skills, ILSIBA ILSI Basic Skills, ILSIMO ILSI Money Management

* $p < .05$, ** $p < .01$

related to ILSI total until positive subscales were removed from the regression model, suggesting the presence of multicollinearity. These results are different from Sanchez (1987), in which NOSIE negative subscales had either

positive or non-significant relationships with living skills. This discrepancy may have been due to methodological differences between the two studies.

Supporting the third hypothesis, NOSIE positive subscales were more strongly related to ILSI total than NOSIE negative subscales. This finding is consistent with Sanchez (1987).

The fourth hypothesis regarding the relationship between the NOSIE and the ILSI and time in program (TIP) was also supported. In general, NOSIE positive subscales and ILSI subscales were negatively related to TIP, while NOSIE negative subscales were positively related to TIP. However, once these predictors were entered into regression models, many of these relationships became non-significant and were differentially correlated with TIP (e.g., some beta weights had opposite signs of bivariate correlations). Again, these results appear to be an artifact of highly interrelated NOSIE and ILSI subscales. Only NOSIE Social Competence was unaffected by multicollinearity, suggesting that it is more strongly and uniquely related to TIP than any other NOSIE or ILSI subscale. Thus, of the domains assessed by the NOSIE and the ILSI, the ability to maintain awareness and memory for a daily schedule (e.g., not having to be reminded about what to do or to follow hospital routine) and to independently engage in tasks (e.g., not having difficulty completing simple tasks on own) is the best predictor of days spent in an inpatient psychiatric rehabilitation program.

The fourth hypothesis regarding the relationship between the NOSIE and the ILSI and number of previous hospitalizations (#HOSP) was partially supported. In concordance with this hypothesis, generally NOSIE negative subscales were positively related to #HOSP. However, contrary to this hypothesis, NOSIE positive subscales and most ILSI subscales were unrelated to #HOSP. The NOSIE model revealed that, due to a combination of multicollinearity and non-significant relationships with #HOSP, only NOSIE Irritability was significantly related to #HOSP when controlling for all other NOSIE subscales. This finding is not surprising considering that NOSIE Irritability was also the only significant predictor of duration of illness in our sample, after controlling for all other NOSIE subscales.¹ These findings appear to be consistent with previous research, which has demonstrated clear relationships between number of hospitalizations and psychiatric symptoms, such as aggressiveness (Grassi et al. 2006; Carr et al. 2008), but generally inconsistent results when using other predictors and this criterion variable (see Klinkenberg and Calsyn 1996 for a review). When the relationship between

¹ This was a follow-up analysis conducted in response to a reviewer's question about the relationship between the NOSIE and the ILSI and illness duration. We wish to thank the reviewer for bringing this question to our attention.

the ILSI model and #HOSP was examined, suppressor effects involving ILSI Home Maintenance (ILSIHO) and Resource Utilization (ILSIRE) were discovered. Suppressor effects indicate that a predictor-criterion relationship is different from the relationship that this predictor has with variance in the criterion variable that is not accounted for by the other predictors (Conger and Jackson 1972). Thus, the relationships between ILSIHO and #HOSP, and ILSIRE and #HOSP are different from the relationships that these predictors have with the variance in #HOSP that is not accounted for by the other ILSI subscales. This may indicate that other ILSI subscales are mediators or moderators of the relationship between ILSIHO and ILSIRE and #HOSP. For example, examination of correlation coefficients between ILSI variables revealed that ILSI Home Maintenance is highly related to ILSI Cooking, Resource Utilization, and Money Management. In addition, ILSI Resource Utilization appears to be highly associated with ILSI Cooking, Basic Skills, and Money Management. These variables may help to explain or may account for the relationship between ILSIHO and #HOSP and between ILSIRE and #HOSP. Future research might address this issue, but should at least assess for the presence of these suppressor effects, as these effects may be unstable when sample size is modest (Conger and Jackson 1972).

Finally, results supported the fifth hypothesis. NOSIE and ILSI subscales predicted TIP and postdicted #HOSP equally well, which provides further evidence that they are highly related instruments.

There were some notable limitations to this study that may be addressed in future research. First, the fact that multiple tests were conducted increases the probability of Type I error, and the current findings should be replicated. Nevertheless, these findings are convergent in their robust support for the validity of the NOSIE and the ILSI. Second, a post hoc power analysis revealed that the sample size used in this study was not large enough to provide enough power to support some of the smaller effects (e.g., relationships between NOSIE and ILSI subscales and #HOSP), and it is recommended that a larger sample size be used in a replication study. Third, this study did not examine mediators or moderators of the relationships between the NOSIE and the ILSI, or between ratings of these measures and outcome variables. In the future, possible covariates, such as symptom severity, should be included in the analyses. Fourth, this study compared relationships between the NOSIE and the ILSI at only one time point, and it would be helpful to examine whether these relationships are time variant or invariant longitudinally. Fifth, the purpose of examining convergent validity in this study was to provide evidence for construct validity, an essential criterion for behavioral assessment instruments. However, divergent validity is also necessary in order to

establish construct validity (Trochim 2006). Convergent and divergent validity of the NOSIE and the ILSI instruments should be addressed in future research. Sixth, the two outcome variables used in this study (time in program and number of previous hospitalizations) provide only coarse estimates of treatment response and recovery. In the future, NOSIE and ILSI subscales and totals should be examined as predictors for less proximal outcome measures such as discharge destination. Finally, we were unable to collect reliability data for the present sample, as individual item scores and multi-rater scores were unavailable. However, reliability analyses were periodically conducted as part of CTP policy with the requirement that reliability coefficients be above .70. Therefore, reliability of the data used in these analyses is likely to be good.

Despite these limitations, this study makes a significant contribution to the literature on behavioral assessment in inpatient psychiatric rehabilitation. Results provide evidence that observed ward behavior is highly related to living skill abilities that are necessary for successful community transition and “real world” functioning. Additionally, this study supports the convergent validity of the NOSIE and the ILSI, thereby helping to justify their continued use in clinical and research settings. Although further research regarding the criterion-related validity of the NOSIE and the ILSI is warranted, this study also demonstrated that these instruments are capable of predicting time spent in an inpatient psychiatric rehabilitation program and postdicting number of previous hospitalizations, which may provide useful information for hospital service providers. Finally, these scales provide excellent measures of functional behavioral abilities that serve as replicable and generalizable outcomes to complement lab assessments and more subjective outcome measures (Keefe et al. 2006a).

Acknowledgments We wish to acknowledge Myla Browne, Ph.D., who developed the dataset used in this study, and the current members of the Serious Mental Illness Research Lab at the University of Nebraska-Lincoln for their helpful comments on earlier drafts of this paper.

The findings and their interpretation are the sole responsibility of the authors and do not necessarily reflect the opinions or policy of the Nebraska Health and Human Services System.

References

- Allen, M. P. (2007). *Understanding regression analysis*. New York: Plenum.
- Aumack, L. (1962). A social adjustment behavior rating scale. *Journal of Clinical Psychology*, *18*, 436–441.
- Baker, L. A., Cheng, L. Y., & Amara, I. B. (1983). The withdrawal of benzotropine mesylate in chronic schizophrenic patients. *British Journal of Psychiatry*, *143*, 584–590.

- Bankier, R. G., & Mathewson, F. A. (1972). A clinical study of mesoridazine and chlorpromazine in relapsed schizophrenic patients. *Diseases of the Nervous System*, 33, 529–534.
- Burdock, E. I., Hardesty, A. S., Hakerem, G., & Zublin, J. (1960). A ward behavior rating scale for mental hospital patients. *Journal of Clinical Psychology*, 16, 246–247.
- Carbonara, D. P., Regan, B. K., Delaney, K. J., & Kumbar, M. (1983). The factorial invariance of the nurses' observation scale for evaluation of inpatients. *Psychological Reports*, 52, 843–848.
- Carr, V. J., Lewin, T. J., Sly, K. A., Conrad, A. M., Tirupati, S., Cohen, M., et al. (2008). Adverse incidents in acute psychiatric inpatient units: rates, correlates and pressures. *Australian and New Zealand Journal of Psychiatry*, 42, 267–282.
- Conger, A. J., & Jackson, D. N. (1972). Suppressor variables, prediction, and the interpretation of psychological relationships. *Educational and Psychological Measurement*, 32, 579–599.
- Dickerson, F. B. (1997). Assessing clinical outcomes: the community functioning of persons with serious mental illness. *Psychiatric Services*, 48, 897–902.
- Farrell, A. D., & Mariotto, M. J. (1982). A multimethod validation of two psychiatric rating scales. *Journal of Consulting and Clinical Psychology*, 50, 273–280.
- Grassi, L., Biancosino, B., Marmai, L., Kotrotsiou, V., Zanchi, P., Peron, L., et al. (2006). Violence in psychiatric units: a 7-year Italian study of persistently assaultive patients. *Social Psychiatry and Psychiatric Epidemiology*, 41, 698–703.
- Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the "right stuff"? *Schizophrenia Bulletin*, 26, 119–136.
- Green, M., Nuechterlein, K., Kern, R., Baade, L., Fenton, W., Gold, J., et al. (2008). Functional co-primary measures for clinical trials in schizophrenia: results from the MATRICS psychometric and standardization study. *American Journal of Psychiatry*, 165, 221–228.
- Heinrichs, R. W., Ammari, N., Miles, A., Vaz, S. M., & Chopov, B. (2009). Psychopathology and cognition in divergent functional outcomes in schizophrenia. *Schizophrenia Research*, 109, 46–51.
- High, R. (2000). *Dealing with 'outliers': How to maintain your data's integrity*. Retrieved July 1, 2009, from University of Oregon, Information Services Web site: <http://cc.uoregon.edu/cnews/spring2000/outliers.html>.
- Honigfeld, G., & Klett, C. J. (1965). The nurses' observation scale for inpatient evaluation: a new scale for measuring improvement in chronic schizophrenia. *Journal of Clinical Psychology*, 21, 65–71.
- Hollister, L. E., Davis, K. L., & Berger, P. A. (1975). Butaclamol hydrochloride in newly admitted schizophrenics. *Psychopharmacology Communications*, 1, 493–500.
- Hopko, D. R., Lachar, D., Bailey, S. E., & Varner, R. V. (2001). Assessing predictive factors for extended hospitalization at acute psychiatric admission. *Psychiatric Services*, 52, 1367–73.
- Keefe, R. S. E., Poe, M., Walker, T. M., & Harvey, P. D. (2006a). The relationship of the brief assessment of cognition in schizophrenia (BACS) to functional capacity and real-world functional outcome. *Journal of Clinical and Experimental Neuropsychology*, 28, 260–269.
- Keefe, R., Poe, M., Walker, T., Kang, J., & Harvey, P. (2006b). The schizophrenia cognition rating scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *American Journal of Psychiatry*, 163, 426–432.
- Klinkenberg, W. D., & Calsyn, R. J. (1996). Predictors of receipt of aftercare and recidivism among persons with severe mental illness: a review. *Psychiatric Services*, 47, 487–496.
- Lentz, R. J., Paul, G. L., & Calhoun, J. F. (1971). Reliability and validity of three measures of functioning with "hard core" chronic mental patients. *Journal of Abnormal Psychology*, 78, 69–76.
- Lorr, M., O'Connor, J. P., & Stafford, J. W. (1960). The psychotic reaction profile. *Journal of Clinical Psychology*, 16, 241–245.
- Lyall, D., Hawley, C., & Scott, K. (2004). Nurses' observation scale for inpatient evaluation: reliability update. *Journal of Advanced Nursing*, 46, 390–394.
- Lyons, J. S., O'Mahoney, M. T., Miller, S. I., Neme, J., Kabat, J., & Miller, F. (1997). Predicting readmission to the psychiatric hospital in a managed care environment: implications for quality indicators. *American Journal of Psychiatry*, 154, 337–340.
- McKibbin, C. L., Brekke, J. S., Sires, D., Jeste, D. V., & Patterson, T. L. (2004). Direct assessment of functional abilities: relevance to persons with schizophrenia. *Schizophrenia Research*, 72, 53–67.
- Menditto, A. A., Wallace, C. J., Liberman, R. P., Vander Wal, J. S., Jones, N. T., & Stuve, P. (1999). Functional assessment of independent living skills. *American Journal of Psychiatric Rehabilitation*, 3, 200–219.
- Nolting, J. R., Kleinlein, P., & Spaulding, W. D. (2005, November). *NRRI vs. non-NRRI: Difference of staff perceptions and implications for staff training and patient treatment decisions*. Poster session presented at the 39th annual meeting of the Association for Behavioral and Cognitive Therapies, Washington, DC.
- Paul, G. L., Mariotto, M. J., & Redfield, J. P. (1986). Sources and methods for gathering information in formal assessment. In G. L. Paul (Ed.), *Principles and methods to support cost-effective quality operations: Assessment in residential treatment settings, Part 1* (pp. 27–62). Champaign: Research.
- Pryjmachuk, S., & Richards, D. A. (2007). Look before you leap and don't put all your eggs in one basket: the need for caution and prudence in quantitative data analysis. *Journal of Research in Nursing*, 12, 43–54.
- Rada, R., & Kellner, R. (1976). Thiothixene in the treatment of geriatric patients with chronic organic brain syndrome. *Journal of the American Geriatrics Society*, 24, 105–107.
- Rothmann, T. L. (2005). *An evaluation of the community transition program: Psychiatric rehabilitation for people with severe, disabling, and treatment resistant psychiatric conditions in Nebraska's health & human services system*. Unpublished manuscript.
- Sanchez, D. A. (1986). *Development of an instrument to assess independent living skills*. Unpublished manuscript.
- Sanchez, D. A. (1987). *The independent living skills inventory: An examination of its validity and clinical use*. Unpublished doctoral dissertation, University of Nebraska, Lincoln.
- Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego: Jerome M. Sattler, Publisher, Inc.
- Smith, G. R., Manderscheid, R. W., Flynn, L. M., & Steinwachs, D. M. (1997). Principles for assessment of patient outcomes in mental health care. *Psychiatric Services*, 48, 1033–1036.
- Smith, T. E., Hull, J. W., Huppert, J. D., & Silverstein, S. M. (2002). Recovery from psychosis in schizophrenia and schizoaffective disorder: symptoms and neurocognitive rate-limitors for the development of social behavior skills. *Schizophrenia Research*, 35, 229–237.
- Spaulding, W. D., Sullivan, M. E., & Poland, J. S. (2003). *Treatment and rehabilitation of severe mental illness*. New York: Guilford.
- Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin*, 95, 334–344.
- Trochim, W. M. K. (2006). *Convergent & discriminant validity*. Retrieved August 1, 2009, from <http://www.socialresearchmethods.net/kb/convdisc.php>.
- Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, 33, 18.
- Volavka, J., Nolan, K. A., Kline, L., Czobor, P., Citrome, L., Sheitman, B., et al. (2004). Efficacy of clozapine, olanzapine, risperidone, and haloperidol in schizophrenia and schizoaffective disorder assessed with nurses observation scale for inpatient evaluation. *Schizophrenia Research*, 76, 127–129.