Journal of the International Neuropsychological Society (2007), **13**, 377–385.

DOI: 10.1017/S1355617707070592

The ecological validity of neuropsychological assessment and the role of depressive symptoms in moderate to severe traumatic brain injury

NAOMI CHAYTOR,[1,2] NANCY TEMKIN,[3–5] JOAN MACHAMER,[3] AND SUREYYA DIKMEN[3,4,6]

[1]Regional Epilepsy Center, Harborview Medical Center, Seattle, Washington
[2]Department of Neurology, University of Washington, Seattle, Washington
[3]Department of Rehabilitation Medicine, University of Washington, Seattle, Washington
[4]Department of Neurological Surgery, University of Washington, Seattle, Washington
[5]Department of Biostatistics, University of Washington, Seattle, Washington
[6]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington

(RECEIVED August 26, 2006; FINAL REVISION December 11, 2006; ACCEPTED December 12, 2006)

Abstract

Evaluating the ecological validity of neuropsychological tests has become an increasingly important topic. Previous research suggests that neuropsychological tests have a moderate level of ecological validity when predicting everyday functioning. The presence of depressive symptoms, however, may impact the relationship between neuropsychological tests and real world performance. The current study empirically tests this hypothesis in a sample of 216 participants with moderate to severe traumatic brain injury (TBI) who completed neuropsychological testing, self-report of mood symptoms, and report of everyday functioning six months post-injury. Contrary to some previous research and clinical lore, results indicated that depression was weakly related to neuropsychological test performance, although it was more strongly related to everyday functioning. Neuropsychological test performance was also significantly related to everyday functioning. The ecological validity of the neuropsychological tests together was not impacted by depressive symptoms, when predicting significant other ratings of functional status. However, patient self-report seems somewhat less related to neuropsychological performance in those with significant depressive symptoms. Neuropsychological test performance was equally related to self and other report of everyday functioning in patients without significant depressive symptoms. (*JINS*, 2007, *13*, 377–385.)

Keywords: Activities of daily living, Everyday functioning, Cognition, Self-report, Depressed mood, Validation studies

INTRODUCTION

In the context of neuropsychological testing, ecological validity refers to the degree to which test performance corresponds to real world performance. Validity does not apply to the test itself but to the inferences that are drawn from the test in a given context (Franzen & Arnett, 1997; Heinrichs, 1990). Therefore, tests that have adequate diagnostic validity (i.e., are associated with dysfunction in a particular brain area) do not necessarily have adequate ecological validity, and ecological validity may vary across patient groups and circumstances. A critical element of ecological validity research is how one defines "everyday functioning," as some researchers have focused on basic or instrumental activities of daily living or employment status, whereas others have focused more specifically on everyday cognitive failures. A recent review of the literature on the relationships between neuropsychological tests and measures of everyday cognitive functioning, found that the magnitude of these relationships is in the moderate range and many individual tests are unrelated or weakly related to measures of outcome (Chaytor & Schmitter-Edgecombe, 2003). More specifically, in a general neurological sample, Chaytor et al. (2006) found that neuropsychological testing accounted for 18% to 20% of the variance in everyday cognitive functioning, as assessed by informant ratings of everyday cognitive failures. Heaton

Correspondence and reprint requests to: Naomi S. Chaytor, Ph.D., Regional Epilepsy Center, Harborview Medical Center, P.O. Box 359745, 325 9th Avenue, Seattle, Washington 98104-2499. E-mail: chaytor@u.washington.edu

et al. (2004) demonstrated greater correspondence between neuropsychological measures and laboratory-based simulations of everyday tasks (e.g., cooking, managing medications) than with self-reported cognitive complaints or self-reported dependence in instrumental activities of daily living.

Taken together, regardless of how everyday functioning is assessed, much of the variance in everyday functioning remains unaccounted for by neuropsychological performance. Many researchers may consider this an indication that neuropsychological tests have inadequate ecological validity. However, several authors indicated that cognitive test scores alone are unlikely to account for a large amount of the variance in everyday functioning for various reasons (Chaytor & Schmitter-Edgecombe, 2003; Franzen & Arnett, 1997; Long & Collins, 1997; Long & Kibby, 1995; Sbordone, 1996; Sbordone & Guilmette, 1999; Silver, 2000). Factors that may influence the relationship between test performance and everyday performance thereby limiting ecological validity, include emotional problems, level of premorbid functioning, motor functioning, health problems, and varying environmental demands (Chaytor et al., 2006; Long & Kibby, 1995; Sbordone, 1997). Accounting for these variables, in addition to performance on cognitive tests, may allow neuropsychologists to better predict everyday functioning.

In addition, the International Classification of Functioning, Disability and Health (ICF) of the World Health Organization asserts that the relationship between a health condition and activity level is not linear, but rather modified by environmental and personal factors. Thus, we should expect neuropsychological test performance (an indicator of disease status assessed in a uniform environment) to be only partially related to everyday activity, with varying environmental and personal factors playing a significant role in determining the ultimate activity level of an individual.

A more complete understanding of the relationship between neuropsychological test performance and everyday functioning is especially critical following traumatic brain injury (TBI), because rehabilitation professionals often base recommendations about appropriateness for return to work and other activities on cognitive test performances. In addition to cognitive deficits, depression is a common sequela of TBI, affecting 14% to 49% within the first year post-injury (Bush et al., 2004; Dikmen et al., 2004; Jorge et al., 2004; Kreutzer et al., 2001). This is problematic, given that depressive symptoms have been linked to everyday functioning (Satz et al., 1998; Temkin et al., 2003). The impact of emotional difficulties in everyday environments has been referred to as the conditional neurological lesion (Sbordone, 1997; Sbordone & Guilmette, 1999). Thus, despite relatively intact test performances, the individual may have marked difficulties in everyday functioning caused by the effect of emotional symptoms on participation in daily activities. The variable nature of depressive symptoms may also cause fluctuations in everyday functioning, making the relationship between cognitive performance and everyday functioning inconsistent over time. The presence of depression may, therefore, reduce the strength of relationships between cognitive test performance and everyday functioning in TBI, although this has not been studied directly.

Further, while previous meta-analytic studies have demonstrated deficits on neuropsychological testing in neurologically normal individuals with depression (Burt et al., 1995; Veiel, 1997), the type and magnitude of deficits differ substantially across studies. In short, whereas depressed individuals perform statistically poorer on most neuropsychological tests on average, as few as 15% were impaired on any given test (Veiel, 1997). The meta-analysis by Veiel (1997) found no differences between groups (depressed and non-depressed) on measures of attention/concentration and found inconsistent memory differences. In a sample of patients involved in litigation, Rohling et al. (2002) failed to find differences in neuropsychological performance associated with depression alone, in patients putting forth adequate effort. This suggests that factors other than the direct physiological effects of depression may account for the association of depression with cognitive impairment. Whereas the effects of depression on cognition in healthy adults are better studied, cognitive deficits secondary to depression have also been documented in mild and moderate TBI (Rapoport et al., 2005). This study was small, with only 21 participants with depression (total sample size of 74) and the non-depressed TBI group had very little cognitive impairment (more than half the sample was intact on all the neuropsychological measures). It is also possible that the depressed group was depressed secondary to being more cognitively compromised or that this group put forth poorer effort on testing. Satz et al. (1998) demonstrated that whereas those with TBI and depression have equivalent cognitive functioning to those without depression, they do have a poorer functional outcome. Similarly, depression was not related to cognitive disturbance in a sample of HIV-infected individuals (Grant et al., 1993), although it was uniquely related to real world functional ability (Heaton et al., 2004).

The current study sought to add to previous research examining the ecological validity of neuropsychological tests in a TBI population, while also investigating the effect of depression on neuropsychological performance and ecological validity. The specific study questions are as follows:

1. Are neuropsychological tests related to self and significant other ratings of everyday functional ability? (i.e., what is the ecological validity of the neuropsychological tests individually and as a group?)
2. Are neuropsychological tests related to depression severity in a TBI sample?
3. Does depression severity add to prediction of everyday ability above and beyond neuropsychological tests?
4. Is the ecological validity of the neuropsychological measures lower in the depressed TBI patients compared to those without depression?

METHOD

Participants

The participants were selected from subjects enrolled in the Magnesium Sulfate Study (Temkin et al., 2007). Patients with moderate to severe TBI admitted to Harborview Medical Center, Seattle, WA, a Level I regional trauma center, between August, 1998, and October, 2004, were eligible for the study. Moderate to severe TBI was defined by intracranial surgery within 8 hours of injury or post-resuscitation Glasgow Coma Scale (GCS) score of 3–12 or, if intubated, GCS motor score of 1–5 without pharmacologic paralysis. Patients were excluded if they were under 14 years old, could not receive study drug within 8 hours of injury, had serum creatinine over 2.0 mg/dL, were pregnant, were prisoners, or were known to live overseas. Study treatment consisted of a loading dose of IV magnesium sulfate and continuous infusion for 5 days to obtain serum magnesium levels between 1.0 and 2.5 mmol/L in those randomized to supplemental magnesium. Additional selection criteria for this project included completion of the 6-month neuropsychological assessment and the Center for Epidemiological Studies-Depression Scale (CES-D). To have another measure of injury severity, the time to follow commands (TFC), as measured by the number of days until the participant was able to follow simple commands consistently as defined by the motor score of the GCS was determined. The Magnesium Sulfate study enrolled 499 subjects. One hundred and six participants died (21.2%). Of the remaining participants, 134 had not completed the CES-D and were therefore excluded. Forty-two additional participants were excluded because they did not complete the neuropsychological assessment. One participant was excluded because she was not following commands by the 6-month assessment, making her CES-D likely invalid. This resulted in a final sample of 216.

It is important to note that the neuropsychological assessment administered for this research project was not a clinical assessment. No reports were generated and the project had a certificate of confidentiality from the National Institutes of Health, that the subject was made aware of, which precluded the results from being subpoenaed without the consent of the subject. This was done in order to reduce any potential external incentives that may have affected test performance. Study personnel were also specifically trained to obtain the patient's best effort throughout all aspects of the assessment.

The protocol and procedures were approved by the University of Washington Human Subjects Division. The study was permitted to enter patients with waiver of consent under the regulation for emergency medical research.

MATERIALS

Neuropsychological Tests

The Trail Making Test (TMT), Part A and B (Reitan & Wolfson, 1995), Controlled Oral Word Association Test (COWAT; Spreen & Benton, 1977), The Selective Reminding Test (SRT; Buschke, 1973), and the Processing Speed Index (PSI) from the WAIS-III (Wechsler, 1997) were selected to broadly survey the areas of memory, executive functioning, processing speed, and language. The Wechsler Abbreviated Scales of Intelligence (WASI; Psychological Corporation, 1999) was administered to characterize the sample. These tests were administered at six months post injury.

Depression

The CES-D was used as a measure of depressive symptoms (Radloff, 1977). It consists of 20 symptoms rated in frequency of occurrence in the last week. This measure is based on the participant's self-report, with higher scores meaning endorsement of more depressive symptoms. A cutoff of $\geqq 16$ was used to indicate clinical levels of depressive symptoms (Bush et al., 2004). The CES-D is a well-validated measure of depressive symptoms and has 86% to 89% sensitivity and 79% to 90% percent specificity for a DSM diagnosis of depression in neurological samples using a cut-off score of 16 (Jones et al., 2005; Parikh et al., 1988). This measure was completed at six months post injury.

Everyday functioning

The Functional Status Examination (FSE) (Dikmen et al., 2001) is a measure developed to evaluate changes in functional status as a result of TBI. It covers physical, social, and psychological domains of functioning. Within each area, a rating of 0 signifies no change; 1 signifies increased difficulties in performing the activity, but still total independence; 2 indicates dependence on others some of the time or not performing some activities in that area; and 3 signifies that the individual is completely dependent on others or does not perform that activity at all. The score is the sum of the limitations over 10 areas of functioning (personal care, mobility/ambulation, mobility/travel, work/school, home management, leisure/recreation, social integration, cognitive/behavioral competency, standard of living and financial independence). Higher scores indicate greater everyday functional limitations. The measure is administered by structured interviews with the participant and a significant other (when available) at six months post injury.

Data Analysis

Five subjects were not administered the SRT and the COWAT because of poor English proficiency. Their scores on those tests were considered to be missing. Several participants were unable to complete individual neuropsychological tests due to the severity of their TBI. These data points were replaced with a value one less than the poorest observed value for that measure. This resulted in less than 1% of the data points for each variable being replaced.

Due to skewed distributions of several variables (i.e., CES-D, TMT A, TMT B, SRT), non-parametric statistics

https://doi.org/10.1017/S1355617707070592 Published online by Cambridge University Press

were employed for all analyses. First, the self and significant other FSE ratings were correlated (Spearman) with the neuropsychological tests to determine the ecological validity of each individual neuropsychological test. Second, to determine the individual relationships between the neuropsychological tests and depression, the neuropsychological test variables and the CES-D were correlated (Spearman). Next, ranked self and significant other FSE scores were each regressed on the ranked neuropsychological test scores in order to determine the ecological validity of the neuropsychological tests together. Fourth, the independent contributions of the neuropsychological tests and the CES-D in predicting everyday functioning, was explored using rank regression.

Lastly, the sample was separated into two groups based on CES-D scores: "depressed" (CES-D $\geqq$ 16) and "non-depressed" (CES-D $<$ 16). Whereas the label "depressed" was used, this actually refers to clinically significant depressive symptoms and should not be confused with a DSM-IV diagnosis of depression. Demographic information, TBI severity, FSE self and significant other ratings, and neuropsychological test performance were compared between the groups using Mann-Whitney U tests with correction for multiple comparisons. The correlations and regression models described previously were computed for each group separately and the extent of correlation was compared using the Fisher r-to-z transformation procedure.

RESULTS

A sample of 216 TBI survivors (age; M = 29.54, SD = 14.04, *range* = 14–87 years) participated in this study. Seventy-eight percent of the sample was male. The average level of education of the sample was 11.36 years (SD = 2.10, *range* = 6–20 years). The average full-scale IQ was 95.51 (SD = 16.62, *range* = 54–135), as estimated by the WASI (Psychological Corporation, 1999). Injury severity was quantified by the worst GCS score obtained within 24 hours of injury (M = 6.85, SD = 2.29, *range* = 3–15) and TFC (M = 10.55 days, SD = 18.54, *Median* = 3.59, *range* = 0–106).

To examine the ecological validity of the individual neuropsychological tests for the sample as a whole, the individual measures were correlated (Spearman) with the self and other FSE score (Table 1). The strong relationship between the self and other ratings of everyday functioning (r = .83, p <.001) suggests that both are measuring the same basic construct. Each neuropsychological measure was significantly related to both the self and other ratings of everyday functioning. The correlations between the neuropsychological tests and the self-FSE were not significantly different in magnitude from those with the significant other FSE, using Choi's test for the difference between dependent correlation coefficients (Choi, 1977).

The Spearman correlations between the neuropsychological measures and the CES-D, while statistically significant, were smaller in magnitude than the relationships between the neuropsychological variables and everyday functioning. These differences were statistically significant for PSI for the self and other FSE and for TMT A for the self FSE. Thus, these neuropsychological measures appear more highly related to everyday functioning than to depressive symptoms. The correlations between the CES-D and self FSE rating (r = .42, p < .001) and the CES-D and significant other FSE rating (r = .35, p < .001) were both significant.

Table 1. Spearman correlations between depressive symptoms, everyday functioning and neuropsychological test performance

Variable	CES-D N = 209–216	FSE self N = 207–214	FSE other N = 184–187
Trails A	.21**[a]	.41***[a]	.42***
Trails B	.20**	.31***	.37***
SRT	.22**	.35***	.40***
PSI	.20**[b,c]	.42***[b]	.50***[c]
COWAT	.16*	.29***	.35***

Note. CES-D = Center for Epidemiological Studies–Depression Scale, FSE self = Functional Status Exam self rating, FSE other = Functional Status Exam other rating, Trails A = Trail Making Test Part A, Trails B = Trail Making Test Part B, SRT = Selective Reminding Test, PSI = WAIS-III Processing Speed Index score, COWAT = Controlled Oral Word Association Test. Scores reflected where appropriate to make correlations positive.
[a,b,c] Significantly different correlations are indicated by the same letter.
*p < .05, **p < .01(significant with Bonferroni correction: p = .05/5), ***p < .001

To determine the ecological validity of the neuropsychological tests as a group, the self and other FSE scores were each regressed on the TMT A, TMT B, SRT, PSI and COWAT using rank ordered regression procedures. The group of neuropsychological measures accounted for 21 percent of the variance in self-FSE ratings (adjusted R^2 = .19, p < .001) and 29 percent of the variance in other-FSE ratings (adjusted R^2 = .27, p < .001). The neuropsychological measures as a group accounted for little of the variance in depression severity (R^2 = .06, adjusted R^2 = .03, p < .05), indicating that neuropsychological difficulty is only weakly predictive of depression severity.

Adding the CES-D scores to the regression model after the group of neuropsychological tests led to an R^2 change of .10 (p < .001) when predicting self-FSE (overall R^2 = .31), and an R^2 change of .05 (p < .001) when predicting other-FSE (overall R^2 = .34). Conversely, adding the group of neuropsychological tests to the CES-D resulted in an R^2 change of .14 (p < .001) when predicting self-FSE and an R^2 change of .22 (p < .001) when predicting other-FSE (see Table 2). Thus, both neuropsychological and emotional variables are significant independent predictors of everyday functioning, although neuropsychological measures appear to be relatively stronger predictors of everyday functioning.

For the following analyses, the sample was divided into two groups based on a CES-D cut off score of 16: depressed (N = 87) and non-depressed (N = 129). At six months

https://doi.org/10.1017/S1355617707070592 Published online by Cambridge University Press

Table 2. Regression of everyday functioning on the neuropsychological measures and depressive symptoms

Block	Dependent variable	R^2 change	Overall R^2	Adjusted R^2
NPT[a] alone	FSE self	—	.21***	.19***
	FSE other	—	.29***	.27***
	CES-D	—	.06*	.03*
CES-D alone	FSE self	—	.17***	.17***
	FSE other	—	.13***	.12***
NPT + CES-D	FSE self	.10***	.31***	.29***
	FSE other	.05***	.34***	.32***
CES-D + NPT	FSE self	.14***	.31***	.29***
	FSE other	.22***	.34***	.32***

Note. All variables were rank ordered due to skewed data.
[a]Neuropsychological variables included in the model: Trails A, Trails B, SRT, COWAT, PSI
*$p < .05$, **$p < .01$,***$p < .001$.

post-injury, 40 percent of our sample was reporting significant symptoms of depression, similar to previous research on rates of depression in TBI. As seen in Table 3, the two groups were of similar age and years of education. With Bonferroni correction, injury severity (as measured by TFC and GCS) and performance on most of the neuropsychological tests were not significantly different, with the exception of the SRT. However, there was a trend for the depressed group to have poorer neuropsychological performances. There was a significant difference between the two groups on the self-FSE (Mann-Whitney U test, $Z = -4.427$, $p < .001$), and significant other FSE ($Z = -3.07$, $p < .005$), with the depressed group having more functional limitations than the non-depressed group.

To examine the impact of depressive symptoms on the ecological validity of the neuropsychological tests, the individual neuropsychological measures were correlated (Spearman's rho) with the self and other FSE scores for each group separately (Table 4). Again, there was a high degree of relatedness between the self and significant other ratings of everyday functioning for both the depressed ($r = .81$, $p < .001$) and non-depressed groups ($r = .82$, $p < .001$). As expected, in the non-depressed group, even with Bonferroni correction for multiple comparisons ($p = .01$), the neuropsychological measures were significantly related to both the self and other ratings of everyday functioning. In the depressed group, whereas the correlations between the neuropsychological tests and the other FSE were similar to those in the non-depressed group, the correlations with the self-FSE were somewhat lower, with only the PSI and TMT A reaching significance when using the Bonferroni correction. None of the correlations between the neuropsychological tests and the self or other FSE in the depressed and non-depressed groups were significantly different using Fisher's z transformation procedure.

To determine if the ecological validity for the neuropsychological tests as a group is poorer in the depressed sample than the non-depressed sample, the self and other FSE scores were each regressed on the ranked TMT A, TMT B, SRT, PSI, and COWAT, for the depressed and non-depressed groups separately (Table 5). The neuropsychological measures accounted for a significant amount of variance in significant other-rated everyday functioning in the non-depressed and depressed participants and in self-rated everyday functioning in the non-depressed group, whereas in the depressed group, the neuropsychological measures showed a trend toward predicting everyday functioning. The multiple correlations between the neuropsychological tests and

Table 3. Demographic characteristics, neuropsychological test performance, and everyday functioning in depressed and non-depressed groups

	Depressed[1] ($N = 87$)		Non-Depressed ($N = 129$)		
	Median	IQR[2]	Median	IQR	p*
Age	25.52	19.61–31.20	23.09	19.20–36.01	.320
Education	11.00	10–12	12.00	10–12.5	.228
TFC (days)	4.80	.57–12.43	2.94	.56–10.44	.375
GCS	7.20	5.8–7.2	7.20	5.8–8.6	.197
Self-FSE	16.33	9.75–22.25	11.06	4–17	<.001*
Other-FSE	17.00	9.75–22	12.00	4.22–18.67	.002*
Neuropsychological Tests					
Trails A (seconds)	31.00	23–44	27.00	20–41	.038
Trails B (seconds)	87.50	60–122.25	72.00	52–105	.052
SRT (number correct)	70.00	59.5–79.5	75.00	67.5–85.5	.004*
PSI (standard score)	84.00	76–91	88.00	76.75–99	.051
COWAT (raw score)	30.00	24–38.75	34.00	26–42	.065

*P-value for significance using the Bonferroni correction: $p \leq .05/10 = .005$
[1]"Depressed" refers to a score ≧ 16 on the CES-D.
[2]IQR = Inter quartile range

https://doi.org/10.1017/S1355617707070592 Published online by Cambridge University Press

Table 4. Spearman correlations between everyday functioning and neuropsychological test performance in the depressed *versus* non-depressed groups

	FSE self		FSE other	
	Depressed[1] (N = 83–86)	Non-Depressed (N = 124–128)	Depressed[1] (N = 69–70)	Non-Depressed (N = 114–117)
Trails A	.30**	.43***	.44***	.39***
Trails B	.20	.34***	.37**	.35***
SRT	.24*	.34***	.39***	.36***
PSI	.32**	.43***	.45***	.50***
COWAT	.16	.34***	.24*	.40***

Note. Scores reflected where appropriate to make correlations positive.
*$p < .05$, **$p < .01$(Significant with Bonferroni correction: $p \leq .05/5$), ***$p < .001$
[1]"Depressed" refers to a score ≧ 16 on the CES-D.

the self or other FSE in the depressed and non-depressed groups were not significantly different using Fisher's z transformation procedure.

DISCUSSION

The neuropsychological measures employed in this study (TMT A & B, SRT, PSI, and COWAT) appear to have adequate ecological validity, individually and as a group, when predicting self and significant other reports of everyday functioning in TBI. Consistent with previous research (Chaytor et al., 2006), the ecological validity of the neuropsychological measures was found to be moderate in the TBI sample as a whole, accounting for 21% to 30% of the variance in functional status. A recent review of the literature on the ecological validity of neuropsychological tests identified several variables that could possibly affect ecological validity research (Chaytor & Schmitter-Edgecombe, 2003). One such variable is depression, such that neuropsychological tests may have less ecological validity in individuals who are depressed.

Interestingly, depression severity and neuropsychological test performance were only weakly related. This finding runs contrary to some previous research and clinical lore suggesting a strong link between depression and poor neuropsychological test performance (Rapoport et al., 2005; Veiel, 1997), although our findings are consistent with several other studies (Heaton et al., 2004; Grant et al., 1993; Rohling et al., 2002). Although the magnitude of the relationships was relatively weak, depression severity was significantly correlated with 4 of 5 neuropsychological measures. The direction of this relationship, however, remains unclear, as it is possible that the presence of cognitive dysfunction results in increased risk of depression (i.e., a reactive depression), or that they are both related to the same underlying neuropathological and neurophysiologic processes. Since correlation does not imply causality, perhaps one way to study this problem is to investigate whether improvement in depressive symptoms leads to improvement in cognitive functions. Whereas improvement in neuropsychological test performance (over a 9 week period) has been documented after being successfully treated for depression in a mild TBI sample, this study did not have a non-treated control group, making it possible that this improvement was because of practice effects or post-injury cognitive recovery (Fann et al., 2001).

Table 5. Rank regression of the neuropsychological measures[a] on everyday functioning in depressed *versus* non-depressed groups

Block	Dependent variable	Multiple R	R^2	Adjusted R^2	p
Depressed[1]	FSE self	.35	.12	.06	.079
	FSE other	.51	.26	.21	.001
Non-depressed	FSE self	.50	.25	.21	<.001
	FSE other	.55	.30	.27	<.001

[a]Ranked variables included in the model: Trails A, Trails B, SRT, PSI, COWAT
[1]"Depressed" refers to a score ≧ 16 on the CES-D.

The neuropsychological measures were significant predictors of functional status, over and above depression scores. Likewise, depressive symptoms were significant predictors of functional status independent of neuropsychological test performance, although to a lesser degree. Thus, emotional functioning and neuropsychological functioning are both unique and largely independent predictors of functional status after TBI. Of note, however, the neuropsychological measures were more closely related to functional status than depressive symptoms were.

Whereas it was hypothesized that the presence of depressive symptoms would reduce the ecological validity of the neuropsychological tests, the data did not support this. Other-reported everyday functioning was significantly related to the neuropsychological measures, regardless of the presence of depressive symptoms. This study suggested, however, that the presence of depressive symptoms after TBI

https://doi.org/10.1017/S1355617707070592 Published online by Cambridge University Press

may be associated with reduced relationships between neuropsychological tests and self-reported everyday functional ability. The neuropsychological measures accounted for 25% of the variance in self-reported everyday functional ability in the sample of TBI patients without significant depressive symptoms and each of the tests was significantly related to self-reported functioning. In contrast, although none of the differences are statistically significant, neuropsychological measures accounted for 12% of the variance in self-reported everyday functioning for the patients with depression and 2 of the 5 measures were significantly related to self-reported everyday functioning. If other studies confirm a lesser degree of ecological validity of the neuropsychological measures for self-reported everyday functioning in depressed people after TBI, the reasons are difficult to determine. Whereas one might think that impaired self-awareness in participants with TBI could account for this finding, in the current study, no evidence for systematic under or over reporting of problems compared to significant other report was observed. The level of endorsement of everyday problems was equivalent between self and other report. Furthermore, the self FSE was highly correlated with the other FSE, in the entire sample, as well as for the depressed and non-depressed groups separately.

Several unique features of the FSE are worth noting. First, the FSE assesses injury-related change in what the individual does in several domains of everyday functioning. Thus, this measure assesses what the person *does*, not what the person is *able to do*, with particular emphasis on the need for extra help performing daily activities. This type of assessment is influenced by not only cognitive deficits, but also by emotional changes that may lead to reduced motivation to perform everyday activities. The domains assessed by the FSE are not specific to cognition, although changes in cognition could certainly affect functioning in the majority of domains, but also assesses changes in employment, financial independence, mobility, self-care, and so forth. In short, whereas cognition is clearly an important predictor of everyday functioning, one would not expect cognition to account for all or even most of the variance in the FSE. In order to fully predict everyday functioning across multiple domains, a combination of variables, including neuropsychological, emotional, premorbid, physical injury sustained in the same accident, economic status, pain severity, and substance abuse (among others) would be necessary. The fact that this study was able to account for approximately one third of the variance in everyday functioning with cognitive and emotional variables alone appears quite reasonable.

The choice of criterion measure is critical in research on the ecological validity of neuropsychological tests. In this study, ecological validity was measured against self and significant other ratings of functional ability in everyday life, using the FSE. There are advantages and limitations of this approach. The literature suggests that self-report is a weaker measure of everyday cognitive performance than clinician and informant ratings in neurologically impaired individuals (Burgess et al., 1998; Chaytor & Schmitter-Edgecombe, 2003; Evans et al., 1997; Goldstein & McCue, 1995; Kaitaro et al., 1995; Sunderland et al., 1983), therefore we used significant other ratings, in addition to self-report. As this study demonstrates, self-report is appropriate in some contexts, although it may be less related to neuropsychological functioning in individuals who are depressed.

Of particular relevance to the current results, a recent study found that self-report of perceived change in cognitive functioning was related to documented changes in test performance, although self-report was not related to absolute levels of cognitive deficit (Christodoulou et al., 2005). Our study used absolute level of neuropsychological performance, which represents the combined effects of premorbid ability level and the effects of the injury, to study ecological validity. Perhaps use of different methods of inference to determine loss of function (Reitan & Wolfson, 1988) and addressing the implications of such losses in combination with the demands imposed on the individual in performing their everyday activities, as done in clinical interpretations, may be more appropriate for examining the ecological validity of neuropsychological measures.

It is important to emphasize that the current sample included a representative group of patients with documented moderate to severe TBI and participants were not assessed in a clinical context. Further, measures were taken to ensure that the participant's performance was minimally influenced by external incentives that may be operative in some clinical samples. In clinical contexts, patients may present with undocumented injuries and may be involved in compensation-related disability claims or may have other financial incentives that could impact their neuropsychological performance, their self-report of depressive symptoms, and/or their self-report of everyday functioning. In those situations, measures of symptom validity and corroboration from more objective sources may be very important when determining the accuracy of self-report data. Thus, our results may not generalize to populations that are dissimilar to ours in these important respects (i.e., involved in litigation). Whereas self-report of everyday functioning was closely related to significant other reports in our sample, this may not be the case in some clinical settings. The patients that are seen in a clinical context are likely very different from the larger more representative population of individuals with TBI. Whereas it is of course possible that individual participants in our sample put forth suboptimal effort or exaggerated their symptomatology, we believe that this is of less concern in our sample for the reasons outlined earlier.

Further, although we did not use measures of symptoms validity in our study, we nonetheless found minimal differences in neuropsychological performance between those reporting significant depressive symptoms and those who did not. In some clinical samples, particularly those with external incentives to perform poorly, symptom validity testing may be necessary in order to accurately assess the effect of depression on cognition. Thus, those who report depressive symptoms may perform poorly on neuropsychological measures because of poor effort rather than because of

depression *per se*. Prior research has shown that there are no significant cognitive effects of depression on test performance in individuals putting forth adequate effort (Rohling et al., 2002).

Although self and informant report is a practical, accepted and the most commonly used method of assessment of everyday functioning, more direct or performance based assessments may circumvent some of the problems described earlier. For example, Heaton et al. (2004) have employed functional simulations in their work, such as standardized assessment of medication management, cooking, and financial skills to examine the ecological validity of neuropsychological measures. Examined by such methods, neuropsychological measures appear to have very good ecological validity (accounting for up to 44 % of the variance in functional simulation performance; Heaton et al., 2004). Neuropsychological measures (and to some extent simulation tasks) perhaps represent what the individual *can do*, but not what they *will actually do*. In contrast, self and informant report measures such as the FSE may assess what the individual actually does. Such measures also have the advantage of assessing functioning in areas not easily assessed in a simulation format, such as social and recreational activities.

Whereas the CES-D has been shown to be an appropriate measure of depression symptom severity in TBI samples (Bush et al., 2004; McCauley et al., 2006) and has shown adequate sensitivity and specificity in identifying individuals with DSM diagnosed depression in neurological patient groups (Jones et al., 2005; Parikh et al., 1988), it is not equivalent to a formal diagnosis of depression. Therefore, the "depressed" patients in our sample may not meet all the diagnostic criteria for major depressive disorder. As such, the differences (or lack there of, in some cases) between the depressed and non-depressed groups in the current study may not generalize to traumatic brain injured individuals with DSM-IV diagnosed major depressive disorder. Further, whereas the CES-D has been validated in neurological samples, this measure has been used more extensively in populations without traumatic brain injuries. Therefore, its use in the current study may represent a potential limitation. It is important to note that the average CES-D score in the depressed sample was 27.64 ($SD = 8.33$), compared to 7.32 ($SD = 4.84$) in the non depressed sample, suggesting that the two groups differed substantially in the magnitude of reported depressive symptoms.

Because the current sample was part of a larger intervention study, it is possible that the results were affected by the presence of magnesium sulfate. However, the magnesium sulfate was only administered for 5 days after study enrollment; whereas the cognitive, emotional, and functional data was gathered 6 months post-injury, making it unlikely that the drug had any effect on the current results. No relationships were found between the presence of study drug and depressive symptoms, functional status, or cognitive performance. Thus, the current results appear unaffected by the larger intervention study.

CONCLUSION

In summary, the current study adds to the literature on the ecological validity of neuropsychological assessment by exploring the role of depressive symptoms when trying to understand the complex relationship between cognitive testing and real world functioning. The ecological validity of the neuropsychological measures was equal in the depressed and non-depressed groups when everyday functioning was reported by a significant other. Self-report of everyday functioning obtained from individuals experiencing symptoms of depression may be less closely related to neuropsychological performance than significant other report is, although the neuropsychological data obtained from these individuals does appear ecologically valid and the self-ratings appear accurate. Further, the presence of depressive symptoms was not strongly related to neuropsychological performance, contrary to some prior research, suggesting that cognitive deficits on neuropsychological testing may not be attributable to depressive symptoms in patients with TBI.

ACKNOWLEDGMENTS

This research was supported by an NINDS/NIH grant (R01 NS19643).

REFERENCES

Burgess, P.W., Alderman, N., Evans, J., Emslie, H., & Wilson, B.A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, *4*, 547–558.

Burt, D.B., Zembar, M.J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin*, *117*, 285–305.

Buschke, H. (1973). Selective reminding for analysis of memory and learning. *Journal of Verbal Learning & Verbal Behavior*, *12*, 543–550.

Bush, B.A., Novak, T.A., Schneider, J.J., & Madan, A. (2004). Depression following traumatic brain injury: The validity of the CES-D as a brief screening device. *Journal of Clinical Psychology in Medical Settings*, *11*, 195–201.

Chaytor, N. & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*, 181–197.

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*, *21*, 217–227.

Choi, S.C. (1977). Tests of equality of dependent correlation coefficients. *Biometrika*, *64*, 645–647.

Christodoulou, C., Melville, P., Scheri, W.F., Morgan, T., MacAllister, W.S., Canfora, D.M., Berry, S.A., & Krupp, L.B. (2005). Perceived cognitive dysfunction and observed neuropsychological performance: Longitudinal relation in persons with multiple sclerosis. *Journal of the International Neuropsychological Society*, *11*, 614–619.

Dikmen, S., Bombardier, C., Machamer, J., Fann, J., & Temkin, N. (2004). Natural history of depression in traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *85*, 1457–1464.

Dikmen, S., Machamer, J., Miller, B., Doctor, J., & Temkin, N. (2001). Functional status examination: A new instrument for assessing outcome in traumatic brain injury. *Journal of Neurotrauma*, *18*, 127–140.

Evans, J.J., Chua, S.E., McKenna, P.J., & Wilson, B.A. (1997). Assessment of the dysexecutive syndrome in schizophrenia. *Psychological Medicine*, *27*, 1–12.

Fann, J.R., Uomoto, J.M., & Katon, W.J. (2001). Cognitive improvement with treatment of depression following mild traumatic brain injury. *Psychosomatics*, *42*, 48–54.

Franzen, M.D. & Arnett, P.A. (1997). The validity of neuropsychological assessment procedures. In H.W. Reese & M.D. Franzen (Eds.), *Biological and neuropsychological mechanisms: Life-span developmental psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Goldstein, G. & McCue, M. (1995). Differences between patient and informant functional outcome ratings in head-injured individuals. *International Journal of Rehabilitation and Health*, *1*, 25–35.

Grant, I., Olshen, R.A., Atkinson, J.H., Heaton, R.K., Nelson, J., McCutchan, J.A., & Weinrich, J.D. (1993). Depressed mood does not explain neuropsychological deficits in HIV-infected persons. *Neuropsychology*, *7*, 53–61.

Heaton, R.K., Marcotte, T.D., Mindt, M.R., Sadek, J., Moore, D.J., Bentley, H., McCutchan, J.A., Reicks, C., Grant, I., & the HNRC Group (2004). The impact of HIV-associated neuropsychological impairment on everyday functioning. *Journal of the International Neuropsychological Society*, *10*, 317–331.

Heinrichs, R.W. (1990). Current and emergent applications of neuropsychological assessment: Problems with validity and utility. *Professional Psychology: Research and Practice*, *21*, 171–176.

Jones, J.E., Hermann, B.P., Woodard, J.L., Barry, J.J., Gilliam, F., Kanner, A.M., & Meador, K.J. (2005). Screening for major depression in epilepsy with common self-report depression inventories. *Epilepsia*, *46*, 731–735.

Jorge, R.E., Robinson, R.G., Moser, D., Tateno, A., Crespo-Facorro, B., & Arndt, S. (2004). Major depression following traumatic brain injury. *Archives of General Psychiatry*, *61*, 42–50.

Kaitaro, T., Koskinen, S., & Kaipio, M. (1995). Neuropsychological problems in everyday life: A 5-year follow-up study of young severely closed-head-injured patients. *Brain Injury*, *9*, 713–727.

Kreutzer, J.S., Seel, R.T., & Gourley, E. (2001). The prevalence and symptom rates of depression after traumatic brain injury: A comprehensive examination. *Brain Injury*, *15*, 563–576.

Long, C.J. & Collins, L.F. (1997). Ecological validity and forensic neuropsychological assessment. In R.J. McCaffrey, A.D. Williams, J.M. Fisher & L.C. Laing (Eds.), *The Practice of Forensic Neuropsychology: Meeting Challenges in the Courtroom*. New York: Plenum Press.

Long, C.J. & Kibby, M.Y. (1995). Ecological validity of neuropsychological tests: A look at neuropsychology's past and the impact that ecological issues may have on its future. *Advances in Medical Psychotherapy*, *8*, 59–78.

McCauley, S.R., Pedroza, C., Brown, S., Boake, C., Levin, H.S., Goodman, H.S., & Merritt, S.G. (2006). Confirmatory factor structure of the center for epidemiologic studies-depression scales (CES-D) in mild-to-moderate traumatic brain injury. *Brain Injury*, *20*, 519–527.

Parikh, R.M., Eden, D.T., Price, T.R., & Robinson, R.G. (1988). The sensitivity and specificity of the Center for Epidemiologic Studies Depression Scale in screening for post-stroke depression. *International Journal of Psychiatry and Medicine*, *18*, 169–181.

Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Harcourt Brace & Company.

Radloff, L.S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.

Rapoport, M.J., McCullagh, S., Shammi, P., & Feinstein, A. (2005). Cognitive impairment associated with major depression following mild and moderate traumatic brain injury. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *17*, 61–65.

Reitan, R.M. & Wolfson, D. (1988). *Traumatic brain injury: Recovery and rehabilitation*, Vol. 2. Tucson, AZ: Neuropsychology Press.

Reitan, R.M. & Wolfson, D. (1995). Category Test and Trail Making Test as measures of frontal lobe functions. *The Clinical Neuropsychologist*, *9*, 50–56.

Rohling, M.L., Green, P., Allen, L.M., & Iverson, G.L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology*, *17*, 205–222

Satz, P., Forney, D.L., Zaucha, K., Asarnow, R.R., Light, R., McCleary, C., Levin, H., Kelly, D., Bergsneider, M., Hovda, D., Martin, N., Namerow, N., & Becker, D. (1998). Depression, cognition, and functional correlates of recovery outcome after traumatic brain injury. *Brain Injury*, *12*, 537–553.

Sbordone, R.J. (1996). Ecological validity: Some critical issues for the neuropsychologist. In R.J. Sbordone & C.J. Long (Eds.), *Ecological Validity of Neuropsychological Testing*, pp. 15–41. Delray Beach, FL: GR Press/St. Lucie Press.

Sbordone, R.J. (1997). The ecological validity of neuropsychological testing. In A.M. Horton, D. Wedding, & J. Webster (Eds.), *The Neuropsychology Handbook*, Vol. 1, *Foundations and Assessment* (2nd ed.). New York: Springer Publishing Company.

Sbordone, R.J. & Guilmette, T.J. (1999). Ecological validity: Prediction of everyday and vocational functioning from neuropsychological test data. In J.J. Sweet (Ed.), *Forensic neuropsychology: Fundamentals and practice*. Lisse, The Netherlands: Swets and Zeitlinger.

Silver, C.H. (2000). Ecological validity of neuropsychological assessment in childhood traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *15*, 973–988.

Spreen, O. & Benton, A. (1977). *Neurosensory Center Comprehensive Examination for Aphasia*. Victoria, B.C., Canada: University of Victoria Neuropsychology Laboratory.

Sunderland, A., Harris, J.E., & Baddeley, A.D. (1983). Do laboratory tests predict everyday memory? A neuropsychological study. *Journal of Verbal Learning and Verbal Behavior*, *22*, 341–357.

Temkin, N.R., Anderson, G.D., Winn, H.R., Ellenbogen, R.G., Britz, G.W., Schuster, J., Lucas, T., Newell, D.W., Mansfield, P.N., Machamer, J.E., Barber, J., & Dikmen, S.S. (2007). Magnesium sulfate for neuroprotection after traumatic brain injury: A randomized trial. *Lancet Neurology*, *1*, 29–38.

Temkin, N., Machamer, J., & Dikmen, S. (2003). Correlates of functional status 3–5 years after traumatic brain injury with CT abnormalities. *Journal of Neurotrauma*, *20*, 229–241.

Veiel, H.O. (1997). A preliminary profile of neuropsychological deficits associated with major depression. *Journal of Clinical and Experimental Neuropsychology*, *19*, 587–603.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scales—III*. San Antonio, TX: Psychological Corporation.