Journal of the International Neuropsychological Society (2004), **10**, 807–817.

DOI: 10.1017/S1355617704106012

Prediction of global outcome with acute neuropsychological testing following closed-head injury

NICHOLAS J. PASTOREK,[1] H. JULIA HANNAY,[1,2] AND CHARLES S. CONTANT[2]

[1]Department of Psychology, University of Houston, Houston, Texas
[2]Department of Neurosurgery, Baylor College of Medicine, Houston, Texas

(Received May 5, 2002; Revised February 26, 2004; Accepted March 17, 2004)

Abstract

Delaying assessment until emergence from post-traumatic amnesia increases completion rates, but this practice causes variable time delays from the date of injury to testing, which can complicate the interpretation of research findings. In the current study, the performance of 105 head injury survivors on simple tests of language comprehension and attention was used to predict global outcome. It was hypothesized that 1 month performance on these measures would aid in the prediction of Disability Rating Scale (DRS) and Glasgow Outcome Scale (GOS) scores collected at 6 months post injury. Only raw scores on the modified Test of Complex Ideational Material accounted for a significant amount of the variance in DRS scores (4.4%) above that accounted for by age, education, Glasgow Coma Scale score, and pupil response. However, testability at 1 month post injury on all four tests consistently accounted for a larger portion of the variance in DRS scores (10.1–13.2%) and significantly improved prediction of GOS scores. Galveston Orientation and Amnesia Test scores collected at 1 month post injury accounted for substantially less variance in DRS scores (7.7–8.4%). Neuropsychological data, including the testability of patients, collected uniformly at 1 month following injury can contribute to the prediction of global outcome. (*JINS*, 2004, *10*, 807–817.)

Keywords: Closed-head injury, Assessment, Outcome

INTRODUCTION

Traditionally, neuropsychological assessment of head-injured patients does not begin until the patient is out of post-traumatic amnesia (PTA) and has continuous memory for events (Boake et al., 2001). This resolution of PTA can occur on the same day as the injury or several months post injury, as in the case of initially comatose or otherwise severely head-injured patients. Patients who have significant memory deficits may not reach the criteria for termination of PTA on standardized tests before discharge from hospital care. Early testing of patients while they are in PTA is possible even with severely head-injured patients, given that the tests are relatively simple. This has been demonstrated by researchers in the past (Ewert et al., 1989; Fodor, 1972; Levin et al., 1988b; Meyers & Levin, 1992; Ruesch & Moore, 1943) and more recently by Hannay et al. (1994). Such assessments can be important for planning individualized rehabilitation regimens and for the projection of outcome data for head-injured patients discharged from hospital care before clearing PTA. There is, however, a need to demonstrate the efficacy of the ability of early testing to predict later outcome. The present study represents an analysis of the prognostic utility of several neuropsychological tests used specifically for the purpose of testing head-injured patients during the early stages of recovery.

Proper exploration of the relationship between early testing and global outcome requires consideration of several variables. Ideally, a study examining this relationship would consist of a large number of consecutively admitted or otherwise unselected patients. Only in this way can a prognostic indicator provide enough ecological validity to be generalized to the general clinical population. All neuropsychological tests must be relatively simple in terms of administration procedures, processing demands, and response requirements to make early testing possible for patients with a range of severities. Further, empirically validated measures of severity and outcome should be employed in an attempt

Reprint requests to: H. Julia Hannay, Ph.D., Department of Psychology, University of Houston, 4800 Calhoun, Houston, TX 77204-5022. E-mail: jhannay@uh.edu

to avoid ambiguity and facilitate replication. Perhaps most important is that assessment takes place within small time intervals around specified time periods post injury, and that the inability of a patient to complete testing be systematically examined as a meaningful point of data (Ruesch & Moore, 1943). A systematic study of the relationship between early testing and global outcome in a severely head-injured population has yet to be reported. Support for the ability of these patients to participate in testing during the early stages of recovery does exist but few, if any, studies meet the criteria outlined above.

Ruesch and Moore (1943) conducted one of the earliest attempts to assess the cognitive status of head-injured patients during the early stages of recovery. Descriptive statistics were provided on 190 consecutively admitted head-injured patients detailing their neurological condition and how it related to their ability to complete testing within the first 24 hr following injury. Patients in this study appeared to represent a range of injury severity, although currently accepted indices of severity, such as the Glasgow Coma Scale (GCS; Teasdale & Jennett, 1974), were not available at the time of their study. Surprisingly, only 9% of the 70 patients who were unable to complete testing failed to do so because of reasons directly related to the head trauma, such as coma and confusion. The relative success of Ruesch and Moore (1943) in examining patients representing a range of severity within the first 24 hr following head injury is promising. Undoubtedly, this success was largely attributable to their choice of five measures representing a broad range of functions, which would now be considered neuropsychological in nature, that required only simple verbal responses.

Since this pioneering effort, relatively little has been done to study the validity and utility of neuropsychological testing during the early stages of recovery following head injury (Ewert et al., 1989; Fodor, 1972; Meyers & Levin, 1992). This paucity of research may stem from the commonly held belief that neuropsychological testing during early recovery from head injury, especially during PTA, is not possible because of pervasive arousal and attentional problems in the population (Stuss & Buckle, 1992). It has been common practice in the United States not to test patients clinically until they achieve a score of 76 or more on the Galveston Orientation and Amnesia Test (GOAT; Levin et al., 1979b) and are thus fully oriented (Boake et al., 2001). However, Hannay et al. (1994) found that a high percentage of severely head-injured patients with a GOAT score of 40 or more regain remote memory for information such as name and date of birth and are able to understand directions for completion of easy tasks. Five of the seven neuropsychological tests administered to these patients had completion rates exceeding 70%. These tests included measures of visual acuity, short-term memory, attention, and language comprehension. The success of administering these tests to patients at 1 month post injury is undoubtedly attributable to their employing simple directions and response requirements, characteristics that are shared with those tests administered by Reusch and Moore (1943). These findings further support the hypothesis that early testing of even severely head-injured individuals is possible.

The current study evaluated the prognostic utility of language comprehension and attention measures taken from a small group of neuropsychological tests used for longitudinal assessment beginning at 1 month following severe head injury (Hannay & Sherer, 1996). These measures were chosen, in part, because their relatively simple directions and response requirements would be most likely to produce favorable completion rates in those recovering from very severe injuries. Directing early assessment efforts towards the measurement of simple language and attentional functioning seemed logical, since the integrity of these functions is necessary for the successful completion of most cognitive measures. Furthermore, deficits in language and attention are well documented in patients recovering from traumatic brain injuries. Individuals with traumatic brain injuries typically demonstrate language deficits in visual naming, word fluency, and auditory comprehension (Groher, 1977; Heilman et al., 1971; Hinchliffe et al., 1998; Levin et al., 1976, 1979a; Sarno, 1980, 1984; Sarno et al., 1986; Thomsen, 1975), as well as deficits in focused and divided attention (Binder et al., 1997; Chan, 2000; Dikmen et al., 1983; Stuss et al., 1985).

We specifically hypothesized that measures of language and attention collected at 1 month post injury would be predictive of global outcome at 6 months post injury, as assessed by the Disability Rating Scale (DRS; Rappaport et al., 1982) and the Glasgow Outcome Scale (GOS; Jennett & Bond, 1975), after taking into account previously established predictors of outcome following head injury. It was expected that poorer performance on the neuropsychological measures at 1 month post injury would be predictive of poorer global outcome at 6 months post injury.

METHODS

Research Participants

Patients were drawn from an ongoing database of consecutive surviving brain-injured patients admitted to the Neurosurgery Intensive Care Unit at Ben Taub General Hospital in Houston. Patients with severe traumatic brain injury had a Best Day 1 GCS of 3–8, an abnormal neurological evaluation, and typically abnormal CT findings. Moderate traumatic brain injuries were associated with a Best Day 1 GCS of 9–12. Complicated mild traumatic brain injury was associated with Best Day 1 GCS of 13–15 and abnormalities on CT scan. Best Day 1 GCS was uniformly defined as the highest GCS score obtained during the first 24 hr post injury. GCS scores were collected after admission to the hospital. Patients were at least 15 years of age. Those with a history of drug and/or alcohol abuse were not excluded in an attempt to provide a more representative sample of the brain-injured population. Patients were excluded if they had a history or current diagnosis of the following conditions: a previous head injury that required medical attention, a disorder of the central nervous system that affected cognitive functioning; or a major psychiatric diagnosis. These individuals were not followed because their premorbid conditions could interfere with administration of the neuropsychological tests and/or interpretation of the results. At the time this study was initiated, 6-month outcome data, which was necessary for the evaluation of the prognostic utility of one-month neuropsychological data, were available on 105 patients. The demographic characteristics of these 105 patients are included in Table 1.

Procedure

Included in a group of tests given at 1 month post injury ± 1 week were two tests of language comprehension and two tests of attention. Language comprehension was assessed with eight yes–no questions from the Test of Complex Ideational Material (CIM) from the Boston Diagnostic Aphasia Exam (Goodglass & Kaplan, 1983), and an 18 question Mini Token Test (MTT; H.J. Hannay, personal communication, 2001). The MTT requires the patient to perform a simple set of actions on colored tokens (e.g., pointing) in response to verbal commands administered by an examiner. The items had little memory load (e.g., put a circle on a square) unlike those of the standard token tests (De Renzi & Vignolo, 1962). Attention was assessed with an Auditory Number Search Test (ANS; Levin et al., 1988a), and a Visual Number Search Test (VNS; Levin et al., 1988a). During the ANS task, patients are instructed to respond by raising a finger after hearing either of two target numbers (i.e., *2*s and *5*s) interspersed among 30 single-digit distractors. Practice items are read by the examiner who corrects performance. The actual test items are then delivered by audio tape. The VNS task requires patients to identify two target numbers (i.e., *2*s and *5*s) that are repeated among 96 other single digit distracters. The 32 target stimuli are randomly placed on a sheet of paper with the restriction that eight appear in each of four quadrants that are not noticeable to the patient. Patients are awarded 1 point for each target number cancelled out with a pencil stroke. The MTT, VNS, and ANS are experimental measures so low in difficulty that errors are rarely made by intact individuals, thus rendering invalid any reliability coefficients obtained through examination of normal individuals. The internal consistency of the full version of the test of CIM has been estimated to be .89 (Goodglass & Kaplan, 1983).

Test completion codes detailing the success of test administration and possible reasons for incomplete testing were recorded for each test individually (see Table 2). Orientation and amnesia were measured with the GOAT at 1 month post injury. GOAT scores were used to provide a comparison standard for interpretation of outcome prediction. Outcome measures were collected at 6 months post injury ±2 weeks. These measures included both the DRS and GOS. The GOS has been criticized for being an insensitive measure of change during recovery from head trauma relative to the DRS, which does not suffer from the restricted range of the GOS (Rappaport et al., 1982). This relative sensitivity to change makes the DRS more attractive for complex statistical modeling, such as associating rate of change on this measure with significant others' perception of cognitive, affective, and physical recovery following brain injury (McCauley et al., 2001). Data collection and scoring were all quality controlled by the second author (H.J.H.) and her staff after administration.

Table 1. Demographic and clinical characteristics of the sample ($n = 105$)

Characteristic	*N* (%)	*M*	*SD*	Range
Sex				
Female	13 (12.4)			
Male	92 (87.6)			
Race				
White	44 (41.9)			
Hispanic	38 (36.2)			
Black	20 (19.0)			
Asian/PI	3 (2.9)			
Age		31.5	12.7	15–72
Education		10.8	3.4	0–20
Best Day 1 GCS	105 (100)	10.3	3.4	3–15
[a]1-month GOAT				
Complete sample	105 (100)	64.1	37.1	−8–100
Complicated mild	33 (31.4)	92.8	8.1	64–100
Moderate	27 (25.7)	69.1	33.2	−8–99
Severe	45 (42.9)	40.1	36.5	−8–100
6-month DRS	105 (100)	2.97	4.20	0–30
6-month GOS				
Good recovery	48 (45.7)			
Moderate disability	43 (41)			
Severe disability	13 (12.4)			
Death	1 (1)			

[a]1-month GOAT scores for those with varying severities of injury.

Statistical Analyses

Regression analyses were used to determine the relation between performance on neuropsychological tests at 1 month post injury and outcome measures collected at 6 months post injury (SPSS, 1999). Linear regression analyses were used to determine the relationship between performance on individual tests and the DRS outcome measure. GOS scores were dichotomized since this procedure frequently produces the best prediction of outcome (Clifton et al., 2001; Goodman et al., 1995; Koura et al., 1988). Good recovery and moderate disability were considered favorable outcomes, while severe disability, vegetative state, and death were considered poor outcomes. Logistic regression analyses were used to determine the relationship between performance on individual tests and the dichotomized GOS scores.

Table 2. Description of test completion codes and the distribution of patients within each code by test

Code	Description	Number of patients ($N = 105$)			
		ANS	VNS	CIM	MTT
0	This test was completely administered.	72	80	89	85
1	Patient with acute confusional state, unable to follow commands, or unable to arouse. This applies to any situation where the testing is not fully administered due to level of arousal problems that are not due to specific medical complications.	5	6	6	8
2	Patient with medical complications. This includes situations such as high fever, respiratory problems, vomiting, etc.	5	4	3	3
3	Patient refused testing or not responsive (not due to 1 or 2). If patient refuses all or any part of the testing, code it here.	10	10	5	7
4	Other (e.g., examiner not available, testing materials not available in Spanish, etc.).	13	5	2	2

Note. In the current study, only patients with codes of 1–3 were considered untestable for reasons directly attributable to the patient, and thus were considered as untestable for the purposes of the dichotomous test completion variable. ANS = Auditory Number Search Test (Levin et al., 1988a); VNS = Visual Number Search Test (Levin et al., 1988a); CIM = Test of Complex Ideational Material (Goodglass & Kaplan, 1983); MTT = Mini Token Test (H.J. Hannay, personal communication, 2001).

Four covariates were chosen based on established predictive validity for outcome following head injury (Dikmen et al., 1994; Levin et al., 1990; Mamelak et al., 1996; Zafonte et al., 1997). Demographic covariates included age and education at the time of injury. Acute care covariates included Best Day 1 GCS and pupil response immediately following injury. Best Day 1 GCS was chosen since post-resuscitation GCS scores are more likely to be confounded by other variables such as the presence of alcohol or sedatives, and because of its relation to outcome following head injury (Cifu et al., 1997). For the current study, pupil response was calculated as the presence or absence of at least one fixed and unresponsive pupil. These four covariates were entered into the linear and logistic models as a block prior to addition of the individual neuropsychological test variables to assess the utility of early neuropsychological testing in the prediction of outcome above and beyond the level of prediction attainable with demographic and acute care information. Because individual patients were sometimes unable to complete all four of the neuropsychological tests, separate restricted models were constructed that included only those patients who were able to complete the neuropsychological test being added to the model.

Separate models were constructed including the four aforementioned covariates and GOAT scores to provide a comparison standard for prediction of global outcome, as the GOAT has repeatedly been shown to account for variance in global outcome (Ellenberg et al., 1996; Katz & Alexander, 1994; Zafonte et al., 1997). Ideally, GOAT scores would have been included as covariates in order to directly assess the ability of neuropsychological test data to account for outcome variance above that accounted for GOAT scores. However, the inclusion of neuropsychological test data and GOAT scores in the same models was not possible due to multicollinearity between these two variables.

RESULTS

Prediction of 6-Month DRS Scores From 1-Month Test Scores

Multiple linear regression was used to explore the relation between performance on neuropsychological tests at 1 month post injury and DRS scores at 6 months post injury. Natural log transformations were performed on DRS scores to produce model residuals that were normally distributed. An R^2_{change} approach was utilized so that the variance attributable to demographic and clinical characteristics could be removed from the outcome variable before entering each neuropsychological test variable. This analysis was repeated for each of the four neuropsychological tests administered at 1 month post injury. Models containing two or more of the four tests could not be constructed due to problems of multicollinearity and because of differential completion rates among the tests. Thus, the number of patients in each model varied slightly since the administration of each test was attempted independently of each patient's ability to complete other tests.

In general, tests of the restricted models revealed that each was significant and accounted for 13.5% to 21.5% of the variance in the outcome measure (Figure 1). The effectiveness of individual covariates in predicting outcome scores was evaluated with tests of individual beta weights (not shown). In the four restricted models, beta weights for the Best Day 1 GCS scores were all significant ($ps < .05$) and consistently the highest of the covariates, indicating that the Best Day 1 GCS score contributed more than the other covariates to the prediction of 6-month DRS status. The only other variable to attain significance was pupil response ($p < .05$), and this occurred only in the restricted model for the MTT.

https://doi.org/10.1017/S1355617704106012 Published online by Cambridge University Press

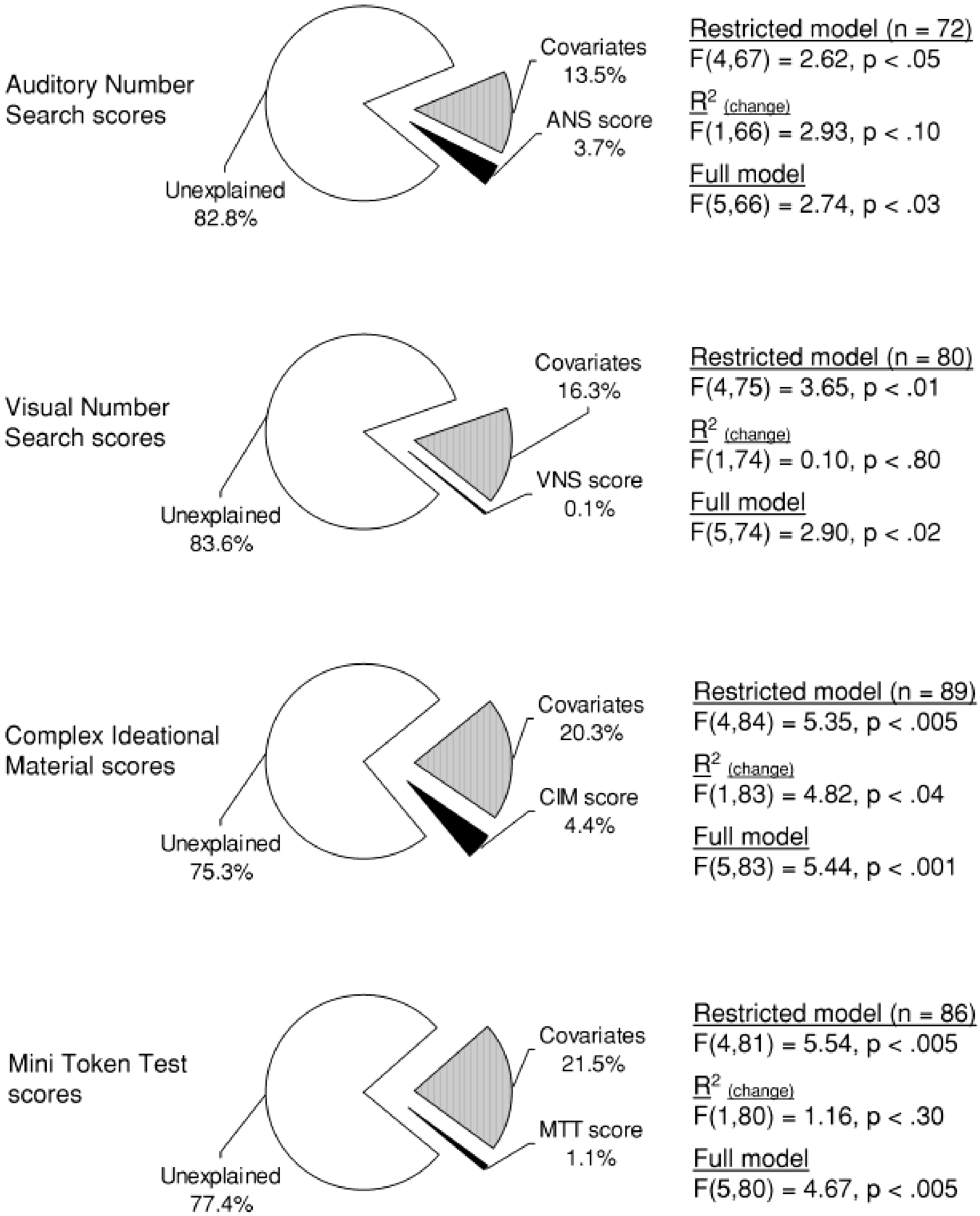

Fig. 1. Percentage of variance in 6-month DRS scores accounted for by covariates and individual neuropsychological test scores. Restricted models included the covariates of age, education, Best Day 1 GCS, and pupil response. Full models included all four covariates and 1-month neuropsychological test scores. R^2_{change} was used to assess the utility of adding neuropsychological test scores to the restricted models that contained known predictors of outcome following head trauma.

Only for CIM scores did the R^2_{change} statistics account for a significant amount of the variance within the outcome measure after removal of variance attributable to the covariates (Figure 1). The addition of the CIM score to its restricted model increased the amount of variance accounted for in DRS scores from 20.3% to 24.7%. Addition of the VNS score did the least to improve on its restricted model, accounting only for an additional 0.1% of variance in DRS scores. Incidentally, the beta weights associated with the Best Day 1 GCS score decreased considerably from the restricted to the full models, indicating that the Best Day 1 GCS score and neuropsychological test scores were not entirely independent. Bivariate correlations confirmed that GCS and test scores were significantly related (*p*s $< .01$), with correlations ranging from .281 with MTT scores to .355 with CIM scores. These correlations suggest that addition of neuropsychological test scores to the restricted models resulted in an attenuation of the strength off the relationship between the GCS scores and the outcome measures, as represented by reductions in beta weights from

https://doi.org/10.1017/S1355617704106012 Published online by Cambridge University Press

restricted to full models. Substituting GOAT scores for each of the raw test scores in the four models resulted in nonsignificant R^2_{change} statistics.

Prediction of 6-Month GOS Scores From 1-Month Test Scores

Logistic regression was used to explore the relationship between performance on neuropsychological tests at 1 month post injury and GOS scores at 6 months post injury. The extremely small number of patients in the severe disability group precluded the use of analytic techniques allowing for a categorical outcome variable with greater than two groups, such as multinomial logistic regression. Given the limited variability of this sample of long-term survivors on GOS scores, this outcome measure was dichotomized into two groups of patients, those with moderate disability and those with good recovery values. Patients with a classification of severe disability on the GOS were removed from the sample, resulting in a loss of 2 to 5 patients, depending on the neuropsychological test being analyzed. As with the multiple linear regression analyses, the demographic and clinical covariates were entered as a block prior to the inclusion of each neuropsychological test score in the model. The chi-square statistic was used to evaluate the improvement in each model's predictive utility after addition of the four covariates, and again after addition of each neuropsychological test score. The usefulness of individual variables, including neuropsychological test scores, for prediction of GOS status was assessed with tests of individual coefficients and their associated odds ratios.

Nonsignificant chi-square values indicated that all restricted and full models failed to improve prediction of 6-month GOS status. Given the lack of significance found during model testing, statistical tests associated with individual predictors were not examined. As with prediction of DRS outcome, substituting GOAT scores for raw test scores failed to improve prediction of GOS outcome.

Development of the Test Completion Variable

The previously reviewed analyses of the relation between acute neuropsychological test scores and 6-month outcome measures failed to take into account data from those patients who were unable to complete testing at 1 month post injury, but who still had valid global outcome measures at 6 months post injury. It was possible that this method of sample selection restricted ranges on the global outcome measures. Indeed, only 5 patients in the above sample had worse outcomes than the moderately disabled category on the GOS (the second most favorable of five total categories). Analyses were performed to evaluate the effectiveness of information regarding patients' abilities to complete testing, rather than raw test scores, for prediction of global outcome. Completion codes were documented for each test for all 105 patients (Table 2). Completion codes for each individual test were used to dichotomize the sample into those who completed testing, and those who could not complete testing for reasons directly attributable for the patient (i.e., confusional state, medical complications, or unresponsiveness). Those who were unable to complete testing for other reasons (e.g., patient not available, examiner not available, etc.) were removed from the analyses, as these reasons did not reflect medically or neuropsychologically meaningful reasons for being unable to complete testing.

The dichotomized test completion variable served as the predictor of global outcome, rather than actual test scores. The dichotomized test completion variables could not be included in the same models that contained raw test scores due to lack of independence between these two variables. Test completion status on ANS for 5 patients could not be determined because the testing material was not available in Spanish. As expected, the use of completion status as a predictor of outcome effectively increased sample size relative to the analyses using raw test scores (Figures 1 and 2).

Prediction of 6-Month DRS Scores From the 1-Month Test Completion Variable

Four linear regressions were used to determine the ability of the test completion status on the four neuropsychological tests to predict 6-month DRS scores. This set of analyses did not otherwise differ from the previous set of analyses concerning prediction of DRS scores. Regarding the prediction of DRS scores, the covariates alone accounted for 27.1% to 28.2% of the variance in the outcome measure (Figure 2). Tests of individual beta weights within the restricted model revealed that Best Day 1 GCS, pupil response, and age accounted for a significant amount of the variance in the outcome measure (ps $<$ .05), with Best Day 1 GCS consistently having the highest beta weight. Addition of the test completion variable accounted for an additional 10.1% to 13.2% of the variance in the outcome measure with completion status on VNS accounting for the most variance, followed by ANS, CIM, and MTT (Figure 2). Substituting GOAT scores for the four test completion variables explained 7.7% to 8.4% of the variance above that accounted for by the covariates (ps $<$ .05 for all restricted and full models, and R^2_{change} statistics).

Prediction of 6-Month GOS Scores From the 1-Month Test Completion Variable

Four logistic regressions were performed with 6-month GOS status as the outcome measure. The addition of patients who were unable to complete testing at 1 month post injury resulted in a distribution of GOS scores that allowed for a more traditional dichotomization of this outcome variable. Those patients scoring in the top two levels (moderate disability and good recovery) represented one of the dichotomous categories in the analyses including test completion

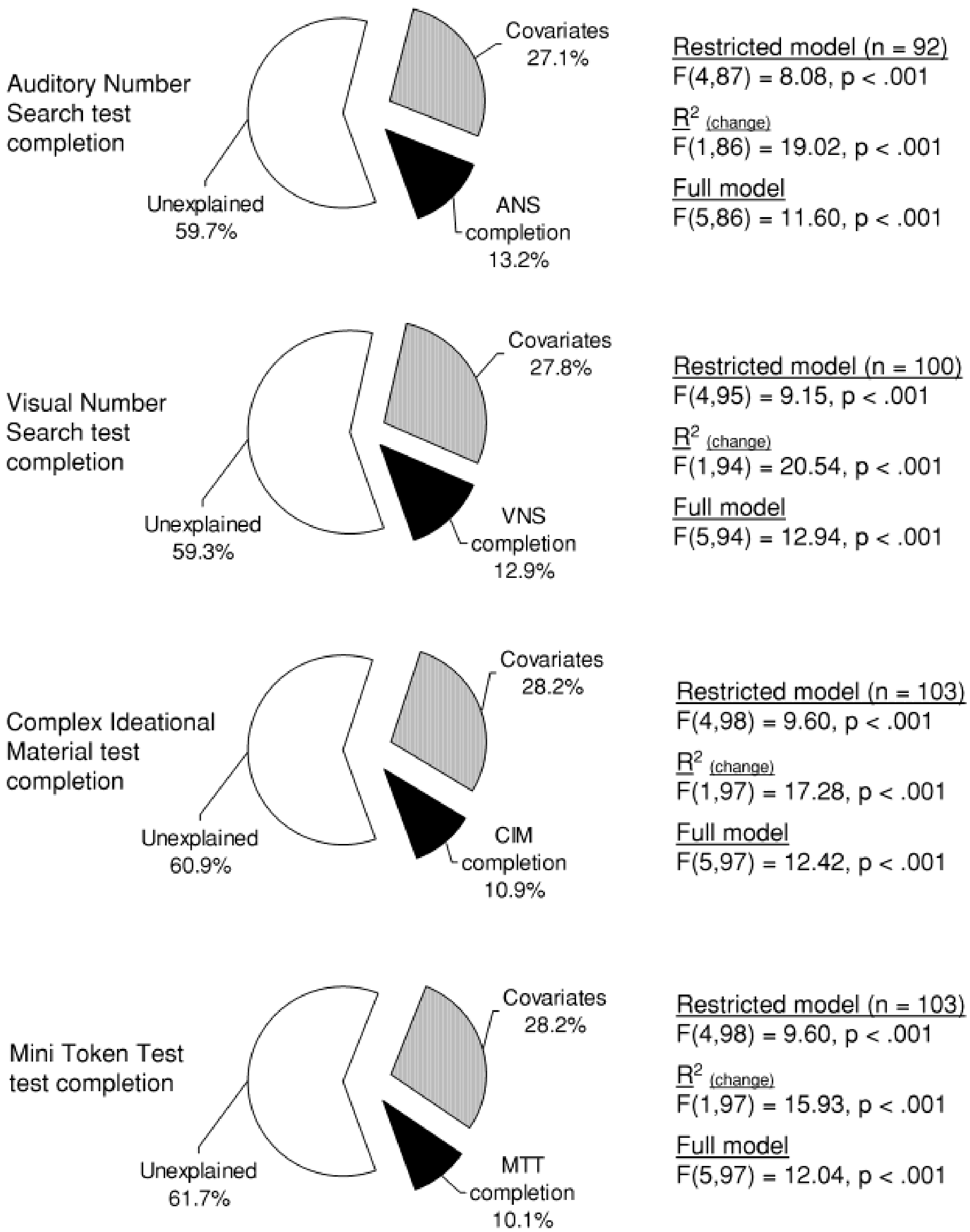

Fig. 2. Percentage of variance in 6-month DRS scores accounted for by covariates and test completion information. Restricted models included the covariates of age, education, Best Day 1 GCS, and pupil response. Full models included all four covariates and 1-month test completion information. R^2_{change} was used to assess the utility of adding the neuropsychological test completion status (complete or incomplete) variable to the restricted models that contained known predictors of outcome following head trauma.

variables, whereas these two GOS levels separately represented both dichotomized categories in the analyses including raw neuropsychological test scores. In order to facilitate comparison between results of the linear and logistic regression analyses, R^2 statistics were calculated for the logistic regression models by squaring the bivariate correlations between observed group scores and predicted group scores (Tabachnick & Fidell, 2001).

The four logistic regression analyses yielded results similar to those found with linear regression (see Table 3). The restricted model including the four covariates significantly increased accuracy of outcome prediction on the GOS at 6 months post injury. Odds ratios revealed that only age and Best Day 1 GCS score significantly increased the accuracy of outcome prediction in both restricted and full models ($ps < .05$). Addition of the test completion variable for each of the four neuropsychological tests produced significant increases in accuracy of outcome prediction. Significant odds ratios ($ps < .005$) associated with the test completion variables indicated that the ability to complete

https://doi.org/10.1017/S1355617704106012 Published online by Cambridge University Press

Table 3. Logistic regression analyses for prediction of GOS outcome scores from 1-month neuropsychological test completion variables

Model	n	$^{c}R^2$	$^{d}R^2_{change}$	χ^2_{change}	χ^2_{model}
Auditory Number Search	92				
[a]Covariates		.296	.296	$\chi^2(4) = 25.79$**	$\chi^2(4) = 25.79$**
[b]Auditory Number Search		.491	.195	$\chi^2(1) = 12.54$**	$\chi^2(5) = 38.32$**
Visual Number Search	100				
[a]Covariates		.300	.300	$\chi^2(4) = 28.66$**	$\chi^2(4) = 28.66$**
[b]Visual Number Search		.513	.213	$\chi^2(1) = 13.97$**	$\chi^2(5) = 42.63$**
Complex Ideational Material	103				
[a]Covariates		.274	.274	$\chi^2(4) = 27.36$**	$\chi^2(4) = 27.36$**
[b]Complex Ideational Material		.489	.215	$\chi^2(1) = 10.95$*	$\chi^2(5) = 38.31$**
Mini Token Test	103				
[a]Covariates		.274	.274	$\chi^2(4) = 27.36$**	$\chi^2(4) = 27.36$**
[b]Mini Token Test		.473	.199	$\chi^2(1) = 10.11$*	$\chi^2(5) = 37.48$**

[a]Covariate models included information regarding age, education, Best Day 1 GCS, and pupil response. [b]Full models were formed by adding the indicated neuropsychological test completion variable to the previously listed covariate model. $^{c}R^2$ was calculated as the squared bivariate correlation between observed group scores and the predicted group scores, as recommended by Tabachnick and Fidell (2001). $^{d}R^2_{change}$ was calculated by subtracting $R^2_{restricted}$ from R^2_{full}.
*$p < .005$. **$p < .001$.

testing (relative to the inability to complete testing) at 1 month post injury increased the odds of a favorable GOS outcome at 6 months post injury by a factor or 28.2 for VNS, 26.7 for ANS, 17.1 for CIM, and 12.9 for MTT (see Table 3). Approximations of R^2 statistics for these logistic regression models were comparable to the linear regression models presented in Figure 2, ranging from 19.5% for the addition of ANS test completion variables to 21.5% for the addition of CIM test completion variables. Substituting GOAT scores for the four test completion variables resulted in significant improvement in prediction over the restricted models ($ps < .05$ for all restricted and full models, and R^2_{change} statistics), but consistently accounted for less variance than test completion scores, with R^2_{change} statistics ranging only from 7.6% to 11.3%.

DISCUSSION

In cases of closed-head injury, complicating factors such as confusion, agitation, and immobility often preclude the use of neuropsychological testing for several months to 1 year. Even relatively small neuropsychological batteries yield only 52% completion rates when administered to severely head-injured patients at 6 months post injury (Scheibel et al., 1998). Unfortunately, many severely head-injured patients are discharged before a formal neuropsychological assessment can be conducted, leaving clinicians unable to utilize neuropsychological data to better predict eventual global outcome. The simplicity of the brief neuropsychological tests being examined in this study could allow for a more accurate approximation of the global outcome of head-injured patients before they are discharged from medical settings.

This study examined the prognostic validity of four brief neuropsychological tests of attention and language comprehension administered at 1 month post injury in a sample of mildly to severely head-injured patients. Neuropsychological data were used to predict measures of global outcome at 6 months post injury. The effects of several known predictors of outcome following head injury were statistically controlled to confirm or deny the possibility that neuropsychological test scores would increase the accuracy of outcome prediction above the level of prediction already attainable through analyses of clinical and demographic information. Due to varying completion rates for each of the four neuropsychological tests, four separate restricted models were constructed, each one including only those patients with valid raw scores for each test in question. Each restricted model contained information regarding age, education, Best Day 1 GCS, and pupil response. The four restricted models associated with the four neuropsychological tests proved to be statistically viable for the prediction of 6-month DRS status, accounting for approximately 13.5% to 21.5% of the variance in the outcome measure. With regards to individual covariates, the Best Day 1 GCS scores contributed most to the prediction of outcome, with the other covariates generally failing to attain significance. This finding was inconsistent with an abundance of literature that suggests age is robustly related to outcome in head-injured populations (Fearnside et al., 1993; Gomez et al., 2000; Mamelak et al., 1996; Signorini et al., 1999; Zafonte et al., 1997). Of course, the tests used in the current study were so simple that age might not be related to performance, only severity of injury.

Only the addition of CIM scores collected at 1 month post injury to the restricted models significantly increased the amount of variance accounted for in the DRS outcome scores. Given these results, further application of raw scores on the VNS, ANS, and MTT for prediction of outcome is not warranted at this time. The increased utility of CIM

https://doi.org/10.1017/S1355617704106012 Published online by Cambridge University Press

performance over the other neuropsychological tests likely resulted from several related factors. CIM has few instructions, individual questions require only a simple yes–no verbal response, and total administration time is relatively short. This allows for more severely impaired patients to complete this task, as reflected in the higher number of patients in PTA completing this test relative to the three other tests. Inclusion of the more severely impaired patients increased sample size and increased variability across subjects. This was apparent in the distribution of CIM scores, which more closely approximated a normal distribution than other test score distributions. Prominent ceiling effects on the other three tests resulted in diminished ability to discriminate between patients. Although promising, there is still considerable room for improvement in the prediction of DRS scores given the rather modest effect size associated with the CIM test ($R^2_{\text{change}} = 4.4\%$). An effect size of this magnitude may have limited practical importance (Maxwell & Delaney, 2000), although it might be argued that few variables will have relatively large effect sizes in relation to something as complex as outcome from closed-head injury. Despite the success of CIM scores in improving prediction of DRS scores, raw scores on all four neuropsychological tests failed to improve prediction of dichotomized GOS scores, again suggesting the limited utility of using raw test scores alone.

Statistical limitations related to sample size and power may have resulted in reduced ability to detect significant change in prediction of outcome related to the addition of neuropsychological test variables to their respective restricted models. For example, R^2_{change} statistics associated with addition of the ANS variable (3.7%) and the CIM variable (4.4%) appear similar, although only the addition of the CIM variable proved to significantly increase outcome prediction. Power analyses were conducted to explore the possibility that varying sample sizes may have altered the power to detect significant effects related to the addition of neuropsychological test variables to the restricted models (Borenstein & Cohen, 1988). There was 68% power to detect an R^2_{change} of 5.0% in the CIM analysis, and only 54% power to detect this level of R^2_{change} in the analysis of ANS data. Larger sample size (89 in the CIM analysis *vs.* 72 in the ANS analysis) and better fit of covariates (20.3% variance accounted for in the CIM restricted model *vs.* 13.5% variance accounted for by the ANS restricted model) resulted in more power to detect the effect associated with addition of the CIM variable relative to the ANS variable. Thus, the current study did not provide adequate power to detect the small to moderate effect sizes associated with the addition of some neuropsychological test scores to models predicting global outcome.

The conclusions drawn above resulted from analyses including only those patients who were able to complete testing at 1 month post injury. However, Ruesch and Moore (1943) noted in their study of acute testing in a head-injured sample that ability to complete testing might itself provide meaningful information about the severity of injury, and ultimately prediction of outcome. Further analyses were conducted in order to test the hypothesis that ability to complete testing, as opposed to the raw test scores themselves, would contribute significantly to global outcome prediction. This procedure increased sample size and variability in the outcome measures by allowing for the inclusion of patients who were unable to complete testing at 1 month post injury, but who still had their global outcome assessed at 6 months post injury. As with the previous analyses involving raw test scores, each of the four restricted models accounted for a significant portion of the DRS scores, ranging from 27.1% to 28.2%. With regard to individual covariates, age, education, and Best Day 1 GCS all contributed significantly to the prediction of six-month DRS scores. The addition of test completion variables to the restricted models accounted for a significant amount of the variance in the DRS outcome measure above that accounted for by the four covariates. R^2_{change} statistics associated with the addition of the test completion variables ranged from 10.1% to 13.2%, resulting in full models for each test that accounted for 38.3% to 40.8% of the variance in the DRS outcome measure. Similarly, the test completion variables significantly improved prediction of dichotomized GOS scores.

These findings add to a growing literature that suggests the testability of patients, rather than their raw test scores, during early recovery can significantly contribute to the prediction of outcome measures (Boake et al., 2001; Dikmen et al., 1994). In addition, the capacity to undergo testing at 1 year post injury is related to several indices of injury severity, such as GCS score, duration of coma, pupillary reactivity, and the presence of intracranial lesions with mass effect (Levin et al., 1990). When developing models to predict the long-term outcome of head injured individuals, it is critically necessary to include information from those patients who could not complete testing due to continuing medical complications. Excluding this subset of the population may lead to a bias in the data (i.e., an underestimation of the degree and rate of less favorable long-term outcome in more severely head-injured patients). This problem was apparent in the current study, in which only 5 patients who had valid scores on neuropsychological tests administered at 1 month post injury had GOS scores below the second most favorable category (moderately disabled) at 6 months post injury. Thus, including only those patients who were able to complete testing created a selection bias in favor of those patients with more positive outcomes. Selection biases similar to these are important factors to consider when conducting forensic work, as a review of the current literature might lead litigation teams to overestimate the likely outcome of more severely injured clients. Examining the test completion variable rather than the raw test scores produced models with higher predictive utilities (Figures 1 and 2). The fact that both covariates and neuropsychological measures were more predictive of outcome in the models including test completion variables should not be surprising, since these analyses led to an increase in the number of patients in each model and, more importantly, led to an increase in the range and variability in the measures included in these models.

Creating separate models to assess the contribution to outcome prediction of both raw test scores and testability elevated the likelihood of committing a Type I error in this study. In all, 16 regressions were performed. Similar outcome prediction studies have employed the Bonferroni correction to maintain an experiment-wise Type I error rate of .05 (Rao & Kilgore, 1992). Applying this method to the 16 regressions in the current study would result in a significance level of .003 for each regression analysis. Interpretation of the current data under this restriction would have resulted in a complete lack of significant effects associated with all models that used raw test scores. However, the findings associated with the test completion variables were robust in that patients' ability to complete testing would still be considered a significant predictor of outcome even after applying this conservative Bonferroni correction.

The current study provided evidence that neuropsychological data collected during early recovery could be used to predict eventual outcome. More importantly, patients who are unable to complete testing should not be systematically removed from analyses, but rather their inability to complete testing should be treated as a valuable piece of information. Further, ample evidence now exists that information regarding the testability of patients can improve prediction of later outcome. This simple dichotomization of patients' testability should give way to more informative analyses. Specifically, the relationship between particular reasons why testing could not be completed and later outcome should be determined. Also, our clinical observations have suggested that the testability of patients at 1 month may help to predict when they will be prepared to undergo testing in the future, an observation that may lead to more effective scheduling of comprehensive follow-up assessments. Despite our success in using early neuropsychological testing data to predict outcome, much room remains for improvement in outcome prediction. Even our best fitting models only accounted for 40.7% of the variance in the outcome measure. In the future, models that simultaneously include performance on a group of simple tests assessing distinct cognitive functions may serve to further increase accuracy of outcome prediction in the acute stages of recovery following closed-head injury. Lastly, prediction of global outcome measures from early test performance should give way to prediction of domain-specific outcome measures, such as measures of return to work or social functioning, to provide specific and meaningful predictions that may be used to allocate resources and develop individualized rehabilitation plans.

ACKNOWLEDGMENTS

This study was completed in partial fulfillment of the requirement for a M.A. in psychology at the University of Houston. The authors would like to express our appreciation to our colleagues who assisted in data collection: Margaret Struchen, Kathleen Fuchs, Stephen McCauley, Jeanie Ricketts, Steve Pluth, Mara Gittess, Lucia Romano, and Garland Stallings. Preparation of this paper was supported in part by National Institutes of Health Grants PO1-NS27616 and PO1-NS38660 as well as National Institute of Disability and Rehabilitation Research Grant H133B40002. Portions of this paper were previously presented at the annual conference of the Houston Neuropsychological Society in 2001 and at the 29th Annual INS Meeting in Chicago, Illinois, in February, 2001.

REFERENCES

Binder, L.M., Rohling, M.L., & Larrabee, J. (1997). A review of mild head trauma. Part I: Meta-analytic review of neuropsychological studies. *Journal of Clinical and Experimental Neuropsychology*, *19*, 421–431.

Boake, C., Millis, S.R., High, W.M., Jr., Delmonico, R.L., Kreutzer, J.S., Rosenthal, M., Sherer, M., & Ivanhoe, C.B. (2001). Using early neuropsychologic testing to predict long-term productivity outcome from traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *82*, 761–768.

Borenstein, M. & Cohen, J. (1988). *Statistical power analysis: A computer program*. Hillsdale, NJ: L. Erlbaum.

Chan, R.C. (2000). Attentional deficits in patients with closed head injury: A further study to the discriminative validity of the test of everyday attention. *Brain Injury*, *14*, 227–236.

Cifu, D.X., Keyser-Marcus, L., Lopez, E., Wehman, P., Kreutzer, J.S., Englander, J., & High, W. (1997). Acute predictors of successful return to work 1 year after traumatic brain injury: A multicenter analysis. *Archives of Physical Medicine and Rehabilitation*, *78*, 125–131.

Clifton, G.L., Miller, E.R., Choi, S.C., Levin, H.S., McCauley, S., Smith, K.R., Muizelaar, P., Wagner, F.C., Jr., Marion, D.W., Luerssen, T.G., Chesnut, R.M., & Schwartz, M. (2001). Lack of effect of induction of hypothermia after acute brain injury. *New England Journal of Medicine*, *344*, 556–563.

De Renzi, E. & Vignolo, L.A. (1962). The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain*, *85*, 665–678.

Dikmen, S., Reitan, R.M., & Temkin, N.R. (1983). Neuropsychological recovery in head injury. *Archives of Neurology*, *40*, 333–338.

Dikmen, S.S., Temkin, N.R., Machamer, J.E., Holubkov, A.L., Fraser, R.T., & Winn, H.R. (1994). Employment following traumatic head injuries. *Archives of Neurology*, *51*, 177–186.

Ellenberg, J.H., Levin, H.S., & Saydjari, C. (1996). Posttraumatic amnesia as a predictor of outcome after severe closed head injury: Prospective assessment. *Archives of Neurology*, *53*, 782–791.

Ewert, J., Levin, H.S., Watson, M.G., & Kalisky, Z. (1989). Procedural memory during posttraumatic amnesia in survivors of severe closed head injury. Implications for rehabilitation. *Archives of Neurology*, *46*, 911–916.

Fearnside, M.R., Cook, R.J., McDougall, P., & McNeil, R.J. (1993). The Westmead Head Injury Project outcome in severe head injury. A comparative analysis of pre-hospital, clinical and CT variables. *British Journal of Neurosurgery*, *7*, 267–279.

Fodor, I.E. (1972). Impairment of memory functions after acute head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *35*, 818–824.

Gomez, P.A., Lobato, R.D., Boto, G.R., De la Lama, A., Gonzalez, P.J., & de la Cruz, J. (2000). Age and outcome after severe head injury. *Acta Neurochirurgica, 142*, 373–380.

https://doi.org/10.1017/S1355617704106012 Published online by Cambridge University Press

Goodglass, H. & Kaplan, E. (1983). *Assessment of aphasia and other related disorders* (2nd ed.). Philadelphia: Lea and Febiger.

Goodman, J.C., Robertson, D.P., Gopinath, S.P., Narayan, R.K., Grossman, R.G., Simpson, R.K., Jr., & Robertson, C.S. (1995). Measurements of lactic acid and amino acids in the cerebral cortex of head-injured patients using microdialysis. In T. Tsubokawa, A. Marmarou, C. Robertson, & G. Teasdale (Eds.), *Neurochemical monitoring in the intensive care unit* (pp. 78–83). Tokyo: Springer.

Groher, M. (1977). Language and memory disorders following closed head trauma. *Journal of Speech Hearing Research*, *20*, 212–223.

Hannay, H.J., Struchen, M.A., Contant, C.F., Stallings, G.A., & Roberston, C.S. (1994, February). *Assessment of severely head injured patients and the prediction of outcome*. Paper presented at the 22nd Annual INS Meeting, Cincinnati, Ohio.

Hannay, H.J. & Sherer, M. (1996). Assessment of outcomes from head injury. In R.K. Narayan, J.E. Wilberger, & J.T. Povlishock (Eds.), *Neurotrauma: A complete textbook on head and spinal injury* (pp. 723–747). New York: McGraw Hill.

Heilman, K.M., Safran, A., & Geschwind, N. (1971). Closed head trauma and aphasia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *34*, 265–269.

Hinchliffe, F.J., Murdoch, B.E., & Chenery, H.J. (1998). Towards a conceptualization of language and cognitive impairment in closed-head injury: Use of clinical measures. *Brain Injury*, *12*, 109–132.

Jennett, B. & Bond, M. (1975). Assessment of outcome after severe brain damage. *Lancet*, *1*, 480–484.

Katz, D.I. & Alexander, M.P. (1994). Traumatic brain injury: Predicting course of recovery and outcome for patients admitted to rehabilitation. *Archives of Neurology*, *51*, 661–670.

Koura, S.S., Doppenberg, E.M.R., Marmarou, A., Choi, S., Young, H.F., & Bullock, R. (1988). Relationship between excitatory amino acid release and outcome after severe human head injury. *Acta Neurochirurgica*, *71*, 244–246.

Levin, H.S., Gary, H.E., Jr., Eisenberg, H.M., Ruff, R.M., Barth, J.T., Kreutzer, J., High, W.M., Jr., Portman, S., Foulkes, M.A., Jane, J.A., Marmarou, A., & Marshall, L.F. (1990). Neurobehavioral outcome 1 year after severe head injury. Experience of the Traumatic Coma Data Bank. *Journal of Neurosurgery*, *73*, 699–709.

Levin, H.S., Grossman, R.G., & Kelly, P.J. (1976). Aphasic disorder in patients with closed head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *39*, 1062–1070.

Levin, H.S., Grossman, R.G., Rose, J.E., & Teasdale, G. (1979a). Long-term neuropsychological outcome of closed head injury. *Journal of Neurosurgery*, *50*, 412–422.

Levin, H.S., Hannay, H.J., Martin, A., & Kreutzer, J. (1988a). *Acute battery tests for the Army Penetrating Head Injury Project.* Unpublished manuscript.

Levin, H.S., High, W.M., Jr., & Eisenberg, H.M. (1988b). Learning and forgetting during posttraumatic amnesia in head injured patients. *Journal of Neurology, Neurosurgery, and Psychiatry*, *51*, 14–20.

Levin, H.S., O'Donnell, V.M., & Grossman, R.G. (1979b). The Galveston Orientation and Amnesia Test: A practical scale to assess cognition after head injury. *Journal of Nervous and Mental Disease*, *167*, 675–684.

Mamelak, A.N., Pitts, L.H., & Damron, S. (1996). Predicting survival from head trauma 24 hours after injury: A practical method with therapeutic implications. *Journal of Trauma*, *41*, 91–99.

Maxwell, S.E. & Delaney, H.D. (2000). *Designing experiments and analyzing data: A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McCauley, S.R., Hannay, H.J., & Swank, P.R. (2001). Use of the Disability Rating Scale recovery curve as a predictor of psychosocial outcome following closed-head injury. *Journal of the International Neuropsychological Society*, *7*, 457–467.

Meyers, C.A. & Levin, H.S. (1992). Temporal perception following closed head injury: Relationship of orientation and attention span. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, *39*, 1001–1007.

Rao, N. & Kilgore, K.M. (1992). Predicting return to work in traumatic brain injury using assessment scales. *Archives of Physical Medicine and Rehabilitation*, *73*, 911–916.

Rappaport, M., Hall, K., Hopkins, K., Belleza, T., & Cope, D.N. (1982). Disability Rating Scale for severe head trauma: Coma to community. *Archives of Physical Medicine and Rehabilitation*, *63*, 118–123.

Ruesch, J. & Moore, B.E. (1943). Measurement of intellectual function in the acute stage of head injury. *Archives of Neurology and Psychiatry*, *50*, 165–170.

Sarno, M.T. (1980). The nature of verbal impairment after closed head injury. *Journal of Nervous and Mental Disease*, *168*, 685–692.

Sarno, M.T. (1984). Verbal impairment after closed head injury. Report of a replication study. *Journal of Nervous and Mental Disease*, *172*, 475–479.

Sarno, M.T., Buonaguro, A., & Levita, E. (1986). Characteristics of verbal impairment in closed head injured patients. *Archives of Physical Medicine and Rehabilitation*, *67*, 400–405.

Scheibel, R.S., Levin, H.S., & Clifton, G.L. (1998). Completion rates and feasibility of outcome measures: Experience in a multicenter clinical trial of systemic hypothermia for severe head injury. *Journal of Neurotrauma*, *15*, 685–692.

Signorini, D.F., Andrews, P.J., Jones, P.A., Wardlaw, J.M., & Miller, J.D. (1999). Predicting survival using simple clinical variables: A case study in traumatic brain injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *66*, 20–25.

SPSS for Windows (Rel. 10.0.1) [computer software]. (1999). Chicago: SPSS, Inc.

Stuss, D.T. & Buckle, L. (1992). Truamatic brain injury: Neuropsychological deficits and evaluation at different stages of recovery and in different pathologic subtypes. *Journal of Head Trauma Rehabilitation*, *7*, 40–49.

Stuss, D.T., Ely, P., Hugenholtz, H., Richard, M.T., LaRochelle, S., Poirier, C.A., & Bell, I. (1985). Subtle neuropsychological deficits in patients with good recovery after closed head injury. *Neurosurgery*, *17*, 41–47.

Tabachnick, B.G. & Fidell, L.S. (2001). *Using Multivariate Statistics* (4th ed.). Boston: Allyn & Bacon.

Teasdale, G. & Jennett, B. (1974). Assessment of coma and impaired consciousness. *Lancet*, *ii*, 81–84.

Thomsen, I.V. (1975). Evaluation and outcome of aphasia in patients with severe closed head trauma. *Journal of Neurology, Neurosurgery, and Psychiatry*, *38*, 713–718.

Zafonte, R.D., Mann, N.R., Millis, S.R., Black, K.L., Wood, D.L., & Hammond, F. (1997). Posttraumatic amnesia: Its relation to functional outcome. *Archives of Physical Medicine and Rehabilitation*, *78*, 1103–1106.