Validation of a New Technique to Detect Malingering of Cognitive Symptoms: The b Test

Kyle Brauer Boone
Harbor-UCLA Medical Center

Po Lu
Pepperdine University & Harbor-UCLA Medical Center

Dale Sherman
Harbor-UCLA Medical Center

Barton Palmer
Harbor-UCLA Medical Center

Carla Back
Harbor-UCLA Medical Center & California School of Professional Psychology

Elias Shamieh
California School of Professional Psychology

Kimberly Warner-Chacon
California School of Professional Psychology

Nancy G. Berman
Harbor-UCLA Medical Center

We administered the b Test, a new measure to identify malingering requiring recognition of overlearned information, to 34 suspected malingerers and to 161 subjects in various clinical groups (moderate to severe head injury, elderly depressed, learning disability, schizophrenia, right and left CVA, and elderly normals). Comparisons of groups revealed more commission...
Within the past 10 years, numerous publications have emerged on the issue of the detection of mali
ergery cognitive symptoms. This research has focused both on the identification of “malingering” patterns on standard cognitive measures, such as the Warrington Recognition Memory Test (Iverson & Franzen, 1994; Millis, 1992), Rey Auditory Verbal Learning Test (Barrash, Suhr, & Manzel, 1998; Bernard, Houston, & Natoli, 1993; Chouinard & Rouleau, 1997; Greiffenstein, Baker, & Gola, 1994; Suhr, Tranel, Wefel, & Barrash, 1997), Wechsler Memory Scale-Revised (Bernard, McGrath, & Houston, 1993; Iverson & Franzen, 1996; Martin, Franzen, & Orey, 1998; Mittenberg, Azrin, Millsaps, & Heilbronner, 1993), Wisconsin Card Sorting Test (Bernard, McGrath, & Houston, 1996), Stroop Test (Osimani, Alon, Berger, & Abarbanel, 1997), Digit Span (Binder & Willis, 1991; Greiffenstein, et al., 1994; Heaton, Smith, Lehman, & Vogt, 1978; Iverson & Franzen, 1994, 1996; Martin, Hayes, & Gouvier, 1996; Mittenberg, Theroux-Fichera, Zieliinski, & Heilbronner, 1995; Suhr et al., 1997; Trueblood, 1994; Trueblood & Schmidt, 1993; Youngjohn, Burrows, & Erdal, 1995), Bender Gestalt (Schretlen, Wilkins, Van Gorp, & Bobholz, 1992), Ravens Standard Progressive Matrices (Gudjonsson & Schackleton, 1986), K-ABC hand movements (Bowen & Littell, 1997), Seashore Rhythm Test (Gfeller & Craddock, 1998), and finger tapping, grip strength, and grooved pegboard (Greiffenstein, Baker, & Gola, 1996), and also on development of tests specifically designed to detect faking, such as the Rey 15-item Memorization Test (Lezak, 1995), Rey Word Recognition Test (Lezak, 1983), Rey Dot Counting Test (Lezak, 1995), Portland Digit Recognition Test (PDRT; Binder, 1993), Hiscock Digit Memory Test (Hiscock & Hiscock, 1989), and Test of Memory Malingering (Rees, Tombaugh, Gansler, & Moczynski, 1998).

These two approaches to the detection of malingering (Iverson & Franzen, 1996) have relied on the fact that the lay public as a group holds many inaccurate beliefs regarding the neuropsychological consequences of head injury (Gouvier, Prestholdt, & Warner, 1988; Willer, Johnson, Rempel, & Linn, 1993). In particular, the general population seems to assume that brain injury causes losses in recognition memory, basic attention span, overlearned information, and motor strength and dexterity, when in actuality, these domains are relatively preserved in all but the most severely brain injured patients (Baddeley & Warrington, 1970; Black, 1986; Heaton et al., 1978; Mittenberg, Rotholc, Russell, & Heilbronner, 1996; Rawling & Brooks, 1990; Rubinsky & Brandt, 1986; Wiggins & Brandt, 1988). This faulty knowledge base causes the malingerer to respond to tests measuring these skills in a manner at variance with that displayed by cooperative brain-injured patients; specifically, malingerers overplay deficits in these areas.

The tests specifically designed to identify faking of cognitive symptoms have primarily focused on documenting feigned impairments in short term memory (e.g., PDRT, Hiscock Digit Memory Test, 15-item Memorization Test, Rey Word Recognition Test), although some tests have been developed to capture other feigned cognitive symptoms, such as malingered losses in mental speed/calculation ability (Dot Counting Test). One understudied area ripe for the development of malingering tests involves measurement...
of recognition of overlearned and highly familiar information. This type of knowledge is particularly resistant to brain injury as reflected by the fact that highly practiced skills, such as word reading, are relatively spared in acquired brain impairment, and in fact this is the rationale behind the use of word-reading tasks to assess premorbid level of function (Nelson & McKenna, 1975). Sparing of sight reading is also the mechanism on which the “Stroop effect” is based, namely, that in brain injury the patient has difficulty overriding this highly automatic behavior. The presence of relatively spared sight reading skills has been confirmed in patients with head injury (Crawford, Parker, & Besson, 1988), dementia (Crawford, Parker, & Besson, 1988; Nelson & McKenna, 1975), heterogeneous cortical disease (Ruddle & Bradshaw, 1982), schizophrenia (Crawford, Besson, Bremner, Ebmeier, Cochrane, & Kirkwood, 1992), and depression (Crawford, Besson, Parker, Sutherland, & Keen, 1987). Thus, a malingering test constructed using this type of overlearned information should have a relatively low false positive rate in patients with actual cerebral dysfunction.

At the same time, because the general public does not seem to be cognizant of the relative preservation of overlearned information in brain injury, the prospect of true positive identification of malingering is enhanced. Anecdotally, the first author has assessed several malingerers who have stated that they had become dyslexic (i.e., now see letters upside down and backwards) after an equivocal head injury, a complaint not reported in cooperative head-injured patients. This experience led to the development of a measure of overlearned information involving “reading” of individual letters (i.e., visual letter discrimination). The purpose of the present study was to assess the validity, in terms of sensitivity and specificity, of this measure in a sample of suspected malingerers versus several clinical groups: moderate to severe head injury, schizophrenic patients, learning disabled adult college students, normal elderly, depressed elderly, and stroke patients. It was hypothesized that malingerers would make more errors of commission (especially involving foils that were the reversed or upside down image of the target), more errors of omission, and take longer to complete the task than the comparison groups.

**METHOD**

**Subjects**

The recruitment procedures for the subject groups and the criteria for inclusion and exclusion are described below. Demographic characteristics of the groups are shown in Table 1.

**Suspected malingerers.** This group is composed of 34 suspected malingerers who were referred for neuropsychological assessment at Harbor-UCLA Medical Center or the private practice of the first author. All subjects were in litigation or seeking to obtain or maintain disability benefits for their reported symptoms and impairments, except for one subject who was diagnosed with a factitious disorder. All subjects were fluent in English and most were native English-speakers. All were outpatients and fully independent in activities of daily living.

To be included in the suspected malingering group, each subject had to show at least two observations of non-credible cognitive symptoms drawn from at least two of the following six tests:

1. Rey 15-item:
   a) <9 items correct (Lezak, 1995),
b) <9 spatial score (Greiffenstein et al., 1996)

2. Rey Dot Counting:
   a) mean grouped dot counting time >130’’ or mean ungrouped time >180’’ (Paul, Franzen, Cohen, & Fremouw, 1992),
   b) <9 correct (Martin et al., 1996),
   c) errors in counting 7 or 8 dots, or
d) mean grouped time ≥ mean ungrouped time

3. Rey Word Recognition:
   a) score ≤6 (after subtracting false positives; Greiffenstein et al., 1996), or
   b) ≤ RAVLT trial 1 recall (Lezak, 1983)

4. Warrington Recognition Memory Test-Words:
   a) <33 (Iverson & Franzen, 1994)

5. Digit Span age-corrected scaled score:
   a) ≤4 (Iverson & Franzen, 1994; Suhr et al., 1997)

6. RAVLT 30’ recognition trial:
   a) ≤7 (Suhr et al., 1997), or
   b) recognition ≤30’ free recall

**Older depressed patients.** This group consisted of 38 fluent English-speaking individuals over age 45. All met the diagnostic criteria for Major Depression (in accordance with the revised third edition of the Diagnostic and Statistical Manual of Mental Disorders) (American Psychiatric Association, 1987) as assessed by the Structured Diagnostic Interview for DSM-III-R (SCID) by an experienced clinician formally trained in the administration of the SCID. At the time of evaluation, subjects were outpatients and were off medications. All subjects underwent complete physical and neurological examinations and subjects were excluded if they reported the presence of psychotic symptoms, a history of a manic or hypomanic episode, drug or alcohol abuse, stroke, epilepsy, Parkinson’s disease, or evidence of hemiparesis or hemisensory deficits on physical examination.

**Elderly normals.** The elderly normal sample consisted of 17 paid, fluent English-speaking community volunteers over age 45 who were recruited through newspaper advertisements, flyers, and personal contacts for research projects conducted at Harbor/UCLA Medical Center. The subjects underwent the same physical and neurological examinations as the depressed group described above and received the same psychiatric interviews. The subjects were free of current or past psychotic, major affective, neurologic, and substance or alcohol dependence disorders.

**Head injured patients.** Six subjects with traumatic head injuries were recruited from Scripps Memorial Hospital Outpatient Rehabilitation Program in Encinitas, California, and another 14 subjects were recruited from Mesa College Head Injury Program in San Diego, California. Subjects were excluded if they were involved in litigation, or reported the presence of psychosis or a significant history of alcohol or drug abuse. They had a minimum Rancho scale level of VI and were at least 6 months post injury. The subjects at Scripps Memorial Hospital were outpatients in a physical rehabilitation program while the subjects from Mesa College were enrolled in ongoing brain recovery classes. Subjects were considered eligible for the study if their hospital or school records indicated that they had a brain CT or MRI scan confirming the presence of a lesion in the brain due to trauma. Inclusion criteria also required fluency in the English language, enough limb use to make circles or write, and adequate comprehension skills to understand simple test instructions.
Patients with left- or right-hemisphere cerebrovascular accident (L- or R-CVA). Twelve subjects with CVA (4 L-CVA; 8 R-CVA) were recruited from Scripps Memorial Hospital Outpatient Rehabilitation Program in Encinitas, California and an additional 8 subjects (6 L-CVA, and 2 R-CVA) were recruited from Mesa College Head Injury Program in San Diego, California. Patients with subcortical stroke or diagnosis of multi-infarct dementia were excluded from the study. The subjects at Scripps Memorial Hospital were outpatients in a physical rehabilitation program while the subjects from Mesa College were enrolled in ongoing brain recovery classes. Subjects were considered eligible for the study if their hospital or school records indicated that they had a brain CT or MRI scan confirming the presence of a lesion in the brain due to stroke. Inclusion criteria also required fluency in the English language, enough limb use to make circles or write, and adequate comprehension skills to understand simple test instructions.

Learning disabled. This group was comprised of 38 fluent English-speaking university students over the age of 18 who were recruited from the Office of Disabled Student Services (ODSS) at the California State University, Northridge (CSUN). All subjects were evaluated by a screening process involving 6–7 hours of diagnostic testing. They were considered eligible for participation in the study if they demonstrated learning deficits and were actively receiving services through ODSS. Exclusion criteria included attention deficit disorder as a primary diagnosis. (Of some surprise, one LD subject had to be excluded due to evidence of noncooperation on two cognitive malingering tests [e.g., reproduced bottom row on the 15-item test in reverse order, miscounted 8 dots as 4 on the Dot Counting Test]; in retrospect, the possibility of secondary gain is an issue in this population given that students designated as learning disabled are provided with advantages [e.g., untimed tests] not available to other students.)

Schizophrenics. This sample included 28 patients from the population of psychiatric inpatients and outpatients at Harbor-UCLA Medical Center. Patients were considered eligible for the study if they met the DSM-III-R diagnostic criteria for schizophrenic disorders with a chronic course, were fluent in English, did not have a preexisting history of neurological disorder or positive findings on a neurological exam, and were able to provide informed consent. Specific diagnoses included paranoid, undifferentiated, disorganized, and catatonic, and most of the patients’ disorders were of a moderate severity (e.g., Brief Psychiatric Rating Scale, [BPRS; Overall & Gorham, 1962], mean score = 40.67, SD = 11.20). All but one of the patients were on neuroleptic medication at the time of testing.

Procedure. The b Test measure was administered to subjects in the form of a 15-page booklet. On the first, fourth, seventh, tenth, and thirteenth pages, 20 lower case “b’s” were interspersed with lower case “d’s” (19), “p’s” (16), and “q’s” (17). The pages were identical to each other with the exception that stimuli on successive pages became progressively smaller, with letters on the first page measuring 7/8" and the letters on page 13 measuring 1/8". The first line of these stimulus pages is reproduced in Figure 1.

On the second, fifth, eighth, eleventh, and fourteenth pages, 16 lower case “b’s” were interspersed with lower case “d’s” (8), “p’s” (8), and “q’s” (8), but also with b’s, d’s, p’s, and q’s with diagonal stems (32); again the pages were identical to each other with the exception that the stimuli on successive pages became smaller, as described above. The first line of these pages is reproduced in Figure 2.

On the third, sixth, ninth, twelfth, and fifteenth pages, 15 lower case “b’s” were again interspersed with lower case “d’s” (5), “p’s” (8), and “q’s” (5), but also with b’s, d’s, p’s
and q’s with additional stems (39); again, the pages were identical to each with the exception that the print became smaller on successive pages as described above. The first line of these pages appears in Figure 3.

The rationale for shrinking the letter size across successive pages was to make the task seem to become more difficult while actually maintaining a trivial level of difficulty to induce malingering patients to overplay fabricated deficits (Slick, Hopp, Strauss, & Spellacy, 1996).

The participants were provided with the following instructions:

I want you to circle all the letter “b’s” on these pages as quickly as you can. The letters you circle should look exactly like this (show example).

The scores derived from the test include the total number of commission errors (circling figures that were not b’s), omission errors (not circling b’s), and total time to complete the task. In addition, within the commission errors, the types of errors (d, p, and q letters; diagonal stems; additional stems) were tabulated.

RESULTS

In Table 1 are shown the gender distribution, and means and standard deviations for age and education, for the eight groups.

As expected, significant group differences in age were detected; \( F(7, 182) = 56.79, p < .00001 \). Scheffe tests indicated that the suspected malingerers were significantly younger \( (p < .05) \) than the normal elderly, depressed elderly, and patients with right hemisphere stroke. When all clinical groups were combined and compared with the suspected malingerers on age, no significant difference was found, \( t (192) = -.82, p = .42 \).

Level of education also varied significantly across groups, \( F(7, 184) = 3.16, p = .0036 \). However, Scheffe analyses revealed that no two groups were significantly different at the .05 level. When all clinical groups were combined and contrasted with suspected ma-
lingerers on education, a significant difference remained, with lower educational level in the suspected malingerers, $t(190) = -2.11, p = .036$.

No significant differences were documented in gender distribution across groups, $\chi^2(7, n = 195) = 6.54, p > .30$.

Although some significant group differences were documented in demographics as noted above, no significant correlations were documented between b Test scores and age (commissions, $r = .13, p = .47$; omissions, $r = .31, p = .07$; time, $r = -.01, p = .94$) or education (commissions, $r = .21, p = .26$; omissions, $r = .14, p = .84$; time, $r = .11, p = .56$) in the suspected malingerers. No significant correlations were documented between education and b Test scores in the comparison groups combined (commissions, $r = -.14, p = .09$; omissions, $r = -.08, p = .32$; time, $r = -.11, p = .20$), and no significant relationships were found between age and omissions ($r = -.04, p = .64$) and additions ($r = .16, p = .09$), although a significant correlation was observed between age and time ($r = .39, p = .0001$). Comparison of males and females on b Test scores, both within the suspected malingerer group and all comparisons groups combined, failed to reveal any significant differences. Because of the general lack of association between b Test scores and demographic variables, no attempt was made to adjust for demographic variables in comparisons on b Test performance.

Because not all test scores were normally distributed across groups, nonparametric group comparisons (Kruskal–Wallis ANOVAs) were computed. Bonferroni corrected significance levels were set at .01. When significant, these were followed by posthoc group comparisons using Mann-Whitney U tests at a significance level of .01.

A Kruskal–Wallis ANOVA comparing all groups on errors of commission was significant, $\chi^2(7) = 65.48, p = .00001$. As shown in Table 2, Mann Whitney U comparisons revealed that the suspected malingerers obtained significantly more commission errors than all other comparison groups except the right stroke patients.

A Kruskal–Wallis ANOVA comparing all groups on errors of omission was also significant, $\chi^2(7) = 32.87, p = .00001$. As shown in Table 2, Mann Whitney U comparisons revealed that the suspected malingerers omitted significantly more targets than all comparison groups except the right stroke patients.

Similarly, a Kruskal-Wallis ANOVA comparing all groups on time was significant, $\chi^2(7) = 61.56, p = .00001$. Results of Mann–Whitney U comparisons, shown in Table 2, revealed significantly slower performance in the suspected malingerers versus all groups except the left stroke and right stroke patients and normal elderly.

In Figure 4 are shown the medians and ranges for the b Test scores across groups.

In the suspected malinger group, a modest but significant correlation was documented between b Test time scores and commission errors ($r = .36, p = .04$); no significant relationship was found between time and omission errors ($r = .05, p = .78$), or between omission and commission errors ($r = .24, p = .17$). In the comparison groups
<table>
<thead>
<tr>
<th>Demographic Characteristics for Each Group of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspected Malingerers</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Mean Age Years</td>
</tr>
<tr>
<td>Mean Education Years</td>
</tr>
<tr>
<td>Gender</td>
</tr>
</tbody>
</table>
combined, a significant relationship was documented between commission errors and time \((r = .39, p = .0001)\); no significant relationship was found between omission errors and time \((r = -.07, p = .40)\) or commission errors \((r = .14, p = .09)\).

To examine b Test score classification accuracy, logistic regression analyses were performed using the three b Test scores as predictors. The subjects were divided into two groups: suspected malingerers versus all other groups combined. 57.58% of suspected malingerers and 98.14% of the combined non-malingering comparison group were correctly classified, resulting in an overall classification accuracy of 91.24%. A second logistic regression analysis was computed, again using the three b Test variables as predictors, and with group membership limited to suspected malingerers and head injured subjects. The rationale for this circumscribed computation was that alleged head injury is the typical scenario in a population feigning cognitive impairment, and thus it is most important to be able to discriminate malingerers from head trauma patients. In this analysis, 87.88% of suspected malingerers and 90.00% of brain injured subjects were correctly classified, with an overall classification accuracy of 88.68%. Results from both logistic regression analyses are summarized in Table 3.

Cut-off scores for the three b Test variables, based on visual inspection of the data to maximize sensitivity and specificity, are shown in Table 4.

Analysis of qualitative errors revealed that nearly half (47%) of the commission errors displayed by the suspected malingerers involved circling the letter d; circled letters with extra stems accounted for one quarter of the errors (23%), circled diagonal stems constituted 13% of the errors, and circled p’s and q’s accounted for 11% and 5% of the errors, respectively. In contrast, the actual distribution of potential errors within the non-b foils was as follows: 23.6% extra stems, 19.4% diagonals, 19.4% d’s, 19.4% p’s, and 18.2% q’s. Thus, the suspected malingerers did not respond randomly when generating commission errors, but in fact over selected d’s. A Kruskal–Wallis analysis comparing groups on number of d errors was significant, \(\chi^2(7) = 91.71, p = .00001\). As shown in Table 2, subsequent Mann–Whitney U analyses revealed significantly more d errors in the suspected malingerers relative to all comparison groups. Sensitivity and specificity data for cut-off scores on d commission errors across all groups are also shown in Table 4.

**DISCUSSION**

The present investigation was an initial attempt to explore the utility of a test involving overlearned and familiar information, namely, letter discrimination, in the detection
FIGURE 4. b Test performance: medians and ranges for number of commissions and omissions, and time in minutes.
Detection of Malingering with the b Test

We are not aware of any other study that has empirically investigated recognition of overlearned information in the identification of malingering.

It is important to develop and validate a wide variety of cognitive malingering detection tasks for several reasons. First, use of several tests allows for convergent identification of malingering and results in a more confident and accurate conclusion than that afforded by a single measure. Second, the availability of numerous malingering tasks helps to safeguard the malingering detection process; it is less likely that a patient can be coached to pass malingering tests if numerous tasks are in common use. Finally, malingerers are often selective in which tasks they choose to feign impairment (memory versus mental speed versus reading skill, etc.), and it is important to have tests available that can capture these differing approaches to symptom magnification.

Performance of suspected malingerers on the b Test was compared with that of a wide variety of comparison groups including elderly normals, head injured and stroke

TABLE 4
Cutoffs and Accuracy of b-Test Variables for Individual Groups

<table>
<thead>
<tr>
<th>Cutoffs</th>
<th>Suspected Malingerer</th>
<th>Head Injured</th>
<th>Learning Disabled</th>
<th>Older</th>
<th>Left CVA</th>
<th>Right CVA</th>
<th>Schizophrenia</th>
<th>Elderly Normal</th>
<th>All Non-Malingerers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commission Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;2</td>
<td>76.5</td>
<td>100</td>
<td>94.7</td>
<td>81.6</td>
<td>70.0</td>
<td>40.0</td>
<td>67.9</td>
<td>94.1</td>
<td>82.6</td>
</tr>
<tr>
<td>&gt;3</td>
<td>76.5</td>
<td>100</td>
<td>97.4</td>
<td>86.8</td>
<td>90.0</td>
<td>40.0</td>
<td>75.0</td>
<td>94.1</td>
<td>87.0</td>
</tr>
<tr>
<td>&gt;4</td>
<td>70.6</td>
<td>100</td>
<td>97.4</td>
<td>89.5</td>
<td>90.0</td>
<td>60.0</td>
<td>78.6</td>
<td>100</td>
<td>90.1</td>
</tr>
<tr>
<td>&gt;5</td>
<td>52.9</td>
<td>100</td>
<td>100.0</td>
<td>100</td>
<td>90.0</td>
<td>80.0</td>
<td>89.3</td>
<td>100</td>
<td>96.3</td>
</tr>
<tr>
<td>Omission Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;0</td>
<td>79.0</td>
<td>100</td>
<td>90.0</td>
<td>92.0</td>
<td>90.0</td>
<td>80.0</td>
<td>82.0</td>
<td>88.0</td>
<td>89.0</td>
</tr>
<tr>
<td>&gt;1</td>
<td>74.0</td>
<td>100</td>
<td>92.0</td>
<td>95.0</td>
<td>100</td>
<td>80.0</td>
<td>89.0</td>
<td>100</td>
<td>94.0</td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;720 seconds</td>
<td>57.6</td>
<td>85.0</td>
<td>97.4</td>
<td>97.4</td>
<td>50.0</td>
<td>30.0</td>
<td>78.6</td>
<td>82.4</td>
<td>83.9</td>
</tr>
<tr>
<td>&gt;750 seconds</td>
<td>57.6</td>
<td>85.0</td>
<td>100</td>
<td>100</td>
<td>50.0</td>
<td>50.0</td>
<td>78.6</td>
<td>88.2</td>
<td>87.0</td>
</tr>
<tr>
<td>&gt;850 seconds</td>
<td>48.5</td>
<td>95.0</td>
<td>100</td>
<td>100</td>
<td>80.0</td>
<td>50.0</td>
<td>85.7</td>
<td>94.1</td>
<td>92.0</td>
</tr>
</tbody>
</table>
patients in a rehabilitation setting, outpatients with major depression, inpatient and outpatients diagnosed with schizophrenia, and learning disabled adults. In the validation process of the b Test it was important not only to assess false positive rates in head injury patients, but also to obtain data on whether the presence of other neurological disturbance (i.e., stroke) and psychiatric disturbance (i.e., depression, schizophrenia) impacted test performance. In addition, given that the b Test involves discrimination/identification of visually-presented letters, it was also critical to determine if the presence of documented learning disability significantly compromised b Test performance.

Suspected malingerers displayed significantly worse b Test performance than the head injured patients and most of the other comparison groups, a finding consistent with many previous investigations of malingering that have found that malingerers overplay their deficits relative to cooperative patients (e.g., Bernard et al., 1996; Martin, Bolter, Todd, Gouvier, & Nicolls, 1993). Specifically, the suspected malingerers committed significantly more commission errors (circling of d’s, p’s, q’s, letters with extra or diagonal stems) and omission errors than all clinical comparison groups except for the right stroke patients, and required significantly more time to complete the task than all clinical groups except for the right and left stroke patients and normal elderly.

The b Test scores, especially number of commission errors, differentiated the suspected malingerers from other groups with good classification accuracy. When a cut-off score of $>2$ commission errors was used, 76.5% of the suspected malingerers were correctly classified, with specificity for the head injury subjects at 100.0, and at 82.6% for all comparison groups combined. Less than 1/5 of non-malingering subjects obtained three or more commission errors, in contrast to nearly 4/5 of suspected malingerers. Lower sensitivity rates were documented for omission errors and time, but specificity remained high. A cut-off of $>40$ omission errors correctly classified more than half of suspected malingerers (58.8) with a specificity of 95.0 for the head injured subjects and 85.1 for all comparison groups combined, while use of a cut-off of $>12$ minutes also correctly classified more than half of suspected malingerers (57.6%), with specificity of 85.0 for the head injury subjects and 83.9 for all comparison groups combined.

Qualitative analysis of commission error type revealed that tabulation of letter d errors resulted in consistently high classification rates for all clinical groups, including stroke patients and patients with schizophrenia, groups that had obtained somewhat lowered specificity rates on other b Test scores. Using a cut-off of $>0$ circled d’s, 79% of suspected malingerers were correctly identified, with specificity for all comparison groups at 80% or higher. Thus, analysis of d commission errors in particular seems to show considerable potential for the accurate detection of malingering.

The fact that time scores were sensitive to group membership corroborates emerging evidence obtained on other cognitive malingering tests that subjects feigning cognitive impairment respond significantly slower than cooperative subjects when completing tasks. Specifically, significantly slowed performance has been documented in malingerers on a computerized priming test (Davis, King, Klebe, Bajszar, Bloodworth, & Wallick, 1997), a computer-administered driving inventory (Ray, Engum, Lambert, Bane, Nash, & Bracy, 1997), on forced choice recognition (Beetar & Williams, 1994; Rees et al. 1998; Rose, Hall, & Szalda-Petree, 1995; Slick, et al. 1996), and learning and memory tasks (Beetar & Williams, 1994).

In addition to the finding that b Test performance is minimally impacted by the presence of neurologic, psychiatric, and learning disabilities, another strength of the test is its relative lack of association with demographic variables. Data from the current study indicate that performance on the b Test is unrelated to education and gender. In addition, although b Test time scores are related to age, omission and commission errors are uncorrelated with age. In contrast, performance on the Rey 15-item Memorization Test has
been found to be significantly inversely correlated with age (Arnett et al., 1995; Griffin, Glassmire, Henderson, & McCann, 1997; Hays, Emmons, & Lawson, 1993; Philpott, 1993; Schretlen, Brandt, Krafth, & Van Gorp, 1991) and significantly positively correlated with education (Back, Boone, Edwards, Parks, Burgoyne, & Silver 1996; Griffin et al., 1997; Hays et al., 1993) with equivocal results for gender (Arnett et al., 1995; Hays et al., 1993). Dot Counting Test performance seems to be adversely effected by increasing age (Arnett & Franzen, 1997; Back et al., 1996), although negative findings have also been reported (Philpott, 1993). Lowered educational level has been found to be related to more Dot Counting errors (Back, 1994) and trend reversals (Arnett & Franzen, 1997). Dot Counting performance has also been reported to be affected by gender, with women scoring below men (Martin et al., 1996). The PDRT, although not correlated with age or gender (Binder & Willis, 1991), seems to be significantly impacted by educational level (Binder & Kelly, 1996). To the extent that malingering test performance is explained by demographic factors, it becomes a less effective measure in documenting motivation/cooperation.

In conclusion, the data from this preliminary study suggest that the b Test seems to be an effective addition to the current complement of cognitive malingering tests. Additional research is needed to cross validate these preliminary findings in independent populations.

REFERENCES


