Journal of Clinical and Experimental Neuropsychology
2002, Vol. 24, No. 5, pp. 561–573
1380-3395/02/2405-561$16.00


# The Rey 15-Item Recognition Trial: A Technique to Enhance Sensitivity of the Rey 15-Item Memorization Test

Kyle Brauer Boone[1], Xavier Salazar[1], Po Lu[1], Kimberly Warner-Chacon[2], and Jill Razani[3]
[1]Harbor-UCLA Medical Center, Torrance, CA, USA, [2]Patton State Hospital, Patton, CA, USA, and
[3]California State University – Northridge, Northridge, CA, USA

## ABSTRACT

Numerous publications on the Rey 15-item Memorization Test have cited limitations primarily in test sensitivity, as well as to some extent in specificity. In the current study, 49 patients with suspect effort, 36 neuropsychology clinic patients not in litigation or attempting to secure disability, 33 learning disabled college students, and 60 normal controls were administered the Rey Test in standard format followed by a recognition trial. A free recall score $<9$ was found to have excellent specificity (97–100%), although sensitivity was modest (47%). However, use of a combined recall and recognition score (i.e., free recall + [recognition − false positives] < 20) substantially increased sensitivity (71%) while maintaining high specificity ($\geq 92\%$).

The Rey 15-item Memorization Test (Lezak, 1995) has been one of the most widely used techniques to detect suspect effort (Slick, Sherman, & Iverson, 1999). Advantages of the test include administration brevity (under 5 min), simplicity and portability of test stimuli (1 stimulus page), and ease of scoring (i.e., tabulation of correct responses out of a possible of 15).

However, numerous publications have cited major limitations primarily in test sensitivity (Arnett, Hammeke, & Schwartz, 1995; Davidson, Suffield, Orenczuk, Nantau, & Mandel, 1991; Frederick, Sarfaty, Johnston, & Powel, 1994; Griffin, Glassmire, Henderson, & McCann, 1997; Guilmette, Hart, Giuliano, & Leininger, 1994; Millis & Kler, 1995; Rogers, Harrell, & Liff, 1993; Schretlen, Brandt, Krafft, & van Gorp, 1991; Vickery, Berry, Inman, Harris, & Orey, 2001), and to some extent in specificity as well (Davidson et al., 1991; Griffin et al., 1997; Griffin, Normington, & Glassmire, 1996; Guilmette et al., 1994; Rogers et al., 1993; Schretlen et al., 1991). In Tables 1 and 2 sensitivity and specificity data are reproduced for studies using a cut-off of $<9$. Sensitivity in samples of volunteer simulators has ranged from 7 to 72%, while sensitivity among samples containing "real world" patients in litigation and/or suspected of malingering has ranged from 5% to 86%.

In the light of the above limitations, several authors have questioned whether the 15-item test should be employed for the determination of malingering (Davidson et al., 1991; Rogers et al., 1993). However, given the above advantages of the test, as well as the fact that it is so well known among psychologists, the question arises as to whether it might be possible to "rehabilitate" the test through increasing sensitivity and specificity.

One potential approach to improve the effectiveness of the test is to attach a recognition trial to the standard administration. Recognition paradigms have been found to be particularly sensitive

Address correspondence to: Kyle Brauer Boone, Ph.D., Harbor-UCLA Medical Center, Department of Psychiatry, Box 498, 1000 W. Carson Street, F-9, Torrance, CA 90509-2910, USA. E-mail: kboone@rei.edu
Accepted for publication: March 6, 2002.

Table 1. Summary of Previously Published Sensitivity Values for the Rey 15-Item Memorization Test Using a Cut-Off of < 9.

| Sensitivity (%) | $n$ | Sample | Author |
|---|---|---|---|
| 12 | 69 | Simulators (normals and inpatient substance abusers instructed to fake) | Schretlen et al. (1991) |
| 43 | 7 | Civil litigation/forensic suspected fakers | Schretlen et al. (1991) |
| 43 | 14 | Malingerers in forensic inpatient setting | King (1992) |
| 38 | 16 | Neurological outpatients in litigation | Lee, Loring, & Martin (1992) |
| 16 | 56 | Neuropsychology clinic referrals with motive to exaggerate symptoms | Frederick et al. (1994) |
| 70 | 18 | Suspected forensic malingerers | Frederick et al. (1994) |
| 62 | 43 | Probable malingerers | Greiffenstein, Baker, & Gola (1994) |
| 15 | 20 | Simulators (college students) | Guillmette et al. (1994) |
| 86 | 14 | Forensic/criminal malingerers | Simon (1994) |
| 61 | 49 | Simulators (college students) | Arnett et al. (1995) |
| 72 | 25 | Simulators (1st year medical students) | Arnett et al. (1995) |
| 5 | 154 | Workers compensation "stress" claimants | Boone et al. (1995) |
| 71 | 7 | Suspected malingerers | Millis & Kler (1995) |
| 7 or 57 | 32 (approx) | Simulators (college students) | Griffin et al. (1997) |

to suspect effort due to the misconception of the general public that recognition is equally or more difficult than free recall; malingerers suppress recognition performance much more than actual patients (Chouinard & Rouleau, 1997; Greiffenstein, Baker, & Gola, 1996; Iverson & Franzen, 1994; Millis, 1992; Suhr, Tranel, Wefel, & Barrash, 1997).

The purpose of the present study was to develop and evaluate the validity of a recognition trial for the Rey 15-item Test. To this end, the 15-item test was administered in standard fashion followed by a recognition trial to four groups of subjects: (1) patients with independent evidence of suspect effort, (2) a general neuropsychological referral population not attempting to secure disability or in personal injury litigation, (3) learning disabled college students (who due to their disability might be expected to have difficulty in processing and recall of the Rey 15-item stimuli), and (4) normal controls.

## METHOD

### Subjects

The recruitment procedures for the subject groups and the criteria for inclusion and exclusion are described below.

#### *Patients With Suspect Effort*

This group was composed of 49 patients with noncredible symptoms 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 cognitive symptoms associated with alleged medical or psychiatric disorders, except for 3 subjects diagnosed with a factitious/somatoform disorder, and 1 person who was incarcerated and had a motive to feign cognitive impairment to evade prosecution for his crimes. All subjects were fluent in English and most were native English-speakers. All were independent in activities of daily living. Presenting diagnoses included: head trauma ($n=19$), "stress" or depression ($n=6$), learning disability ($n=4$), toxic exposure ($n=4$), psychosis or bipolar disorder ($n=3$), stroke ($n=3$), somatoform or factitious disorder ($n=3$), dementia ($n=2$), chronic fatigue syndrome ($n=2$), anoxia ($n=1$), narcolepsy ($n=1$), and decreased memory from ECT ($n=1$).

The criteria for inclusion in the suspect effort group are shown in Table 3.

#### *Neuropsychology Clinic Referrals*

This group included 36 fluent English-speaking patients over the age of 18 evaluated in the Neuropsychology Service at Harbor-UCLA Medical Center. Patients were excluded from the study if they were in personal injury litigation or were attempting to secure disability benefits, if they met DSM-IV criteria for dementia, or if their Full Scale IQ, as measured

Table 2. Summary of Previously Published Specificity Values for the Rey 15-Item Memorization Test Using a Cut-Off of < 9.

| Specificity (%) | *n* | Sample | Authors |
|---|---|---|---|
| 100 | 50 | Psychiatric inpatients | Goldberg & Miller (1986) |
| 62 | 16 | Mentally retarded patients | Goldberg & Miller (1986) |
| 89 | 18 | Brain damaged inpatients in rehabilitation | Bernard & Fowler (1990) |
| 100 | 16 | Controls | Bernard & Fowler (1990) |
| 80 | 60 | Referrals for neuropsychological evaluation (primary neurological diagnosis with at least mild memory impairment, not in litigation or disability proceedings) | Morgan (1991) |
| 100 | 80 | Controls | Schretlen et al. (1991) |
| 73 | 148 | Mixed patient group (neuropsychiatric, brain injured, mixed dementia, severe psychiatric, genuine amnesics) | Schretlen et al. (1991) |
| 56 | 16 | Forensic inpatients with moderate to severe organic mental disorders | King (1992) |
| 93 | 100 | Temporal lobe seizure patients | Lee et al. (1992) |
| 92 | 40 | Nonlitigating neurologic outpatients | Lee et al. (1992) |
| 13 | 15 | Mild dementia | Philpott (1992) |
| 0 | 22 | Moderate dementia | |
| 0 | 12 | Moderate/severe dementia | |
| 98 | 47 | Normal middle-aged/older controls | Philpott (1992) |
| 97 | 49 | Neuropsychology clinic referrals without motive to malinger | Frederick et al. (1994) |
| 90 | 18 | Forensic patients referred for psychiatric treatment | Frederick et al. (1994) |
| 88 | 33 | Traumatic brain injury | Greiffenstein et al. (1994) |
| 55 | 20 | Moderate to severe brain damaged inpatients in rehab (CVA, HI, hypoxia, dementia with HI, alcohol dementia with HI, PSP) | Guilmette et al. (1994) |
| 90 | 20 | Psychiatric inpatients with affective disorder (19 major depression, 4 with psychotic features; 1 bipolar) | Guilmette et al. (1994) |
| 86 | 14 | Severe psychiatric disturbance (schizophrenia, schizoaffective disorder, delusional disorder, dementia) but not mentally retarded or acutely psychotic | Simon (1994) |
| 100 | 40 | Learning disabled college students | Warner-Chacon (1994) |
| 76 | 34 | Neurologically impaired inpatients and outpatients referred for neuropsychological evaluation | Arnett et al. (1995) |
| 84 | 25 | Neurologically impaired inpatients and outpatients referred for neuropsychological evaluation | Arnett et al. (1995) |
| 100 | 7 | Traumatic brain injury | Millis & Kler (1995) |
| 87 | 30 | Schizophrenic outpatients (20 outpatients, 10 inpatients) | Back, Boone, Edwards, & Parks (1996) |
| 90 | 20 | Severe head trauma patients | Shamieh (1996) |
| 60 | 20 | Stroke patients | Shamieh (1996) |
| 34 | 37 (approx) | Board and care residents due to schizophrenia or other major mental illness, developmental disability, or advanced age | Griffin et al. (1997) |
| 100 | 32 (approx) | College students | Griffin et al. (1997) |
| 95 | 64 | Older outpatients with major depression | Lee et al. (2000) |

Table 3. Criteria for Inclusion in the Nonforensic Suspect Effort Group.

To be included in the nonforensic suspect effort group, each subject had to show observations of noncredible cognitive symptoms drawn from at least two of the following six tests:

**(1)** ***Rey Dot Counting Test:***
(a) combination score (mean ungrouped dot counting time + mean grouped dot counting time + number of errors) $\geq 17$ (Boone et al., in press)

**(2)** ***Rey Word Recognition***:
(a) score $\leq 6$ after subtracting false positives (Greiffenstein et al., 1996), or
(b) $\leq$ Rey Auditory Verbal Learning Test trial 1 recall (Lezak, 1983)

**(3)** ***Harbor-UCLA b Test***:
(a) $> 2$ commission errors,
(b) $> 0$ "d" commission errors,
(c) $> 40$ omission errors, or
(d) $> 12$ min completion time (Boone et al., 2000)

**(4)** ***Warrington Recognition Memory Test – Words:***
(a) $< 33$ (Iverson & Franzen, 1994)

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

**(6)** ***Rey Auditory Verbal Learning Test 30′ recognition trial***:
(a) $\leq 7$ (Suhr et al., 1997), or
(b) recognition $\leq$ 30-min free recall

In addition, all nonforensic suspect effort subjects met at least one of the following behavioral criteria:

(1) Pattern of neuropsychological scores not consistent with medical or psychiatric condition

(2) Severity of neuropsychological impairment not consistent with medical and psychiatric condition

(3) Markedly inconsistent responses during testing session (e.g., marked discrepancy in scores on tests measuring the same skill [average versus < 1st percentile), or patient observed executing a behavior he/she claimed he/she could not do [e.g., claims cannot read but repeatedly "reads" words on Stroop interference task])

(4) Marked inconsistency between neuropsychological scores and activities of daily living (e.g., severely impaired scores but lives independently, handles own finances, drives, etc.)

(5) Marked inconsistency in postinjury/illness neuropsychological scores across separate testing evaluations (e.g., marked drop from average to impaired [< 2nd percentile] on identical tests)

(6) Implausible self-reported symptoms (e.g., claims postinjury/illness that now sees letters upside down and backwards, cannot see through glass, etc.)

(7) Major contradictions between self-report of symptoms, medical records, and observed behavior including surveillance

(8) Noncredible findings on neurologic or other medical exam.

by the WAIS–R/III, fell below 70. Patients presented with complaints of cognitive dysfunction related to HIV infection ($n=7$), head trauma ($n=6$), depression ($n=6$), attention deficit disorder ($n=1$), psychosis ($n=4$), stroke ($n=3$), alcohol or other substance abuse ($n=3$), hydrocephalus ($n=1$), brain tumor ($n=1$), PTSD ($n=1$), bipolar disorder ($n=1$), anoxia ($n=1$), and cerebrovascular disease ($n=1$).

### *Learning Disabled*

This group was comprised of 33 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 (Warner-Chacon, 1994). All subjects were evaluated by a screening process involving 6–7 hr 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, current or past substance abuse, and history of head trauma or significant neurologic disorder (e.g., cerebral palsy). (Of some surprise, one subject had to be excluded due to evidence of noncooperation on a separate effort test [i.e., miscounted 8 dots as 4 on the Rey Dot Counting Test]; the possibility of secondary gain is an issue in this population given that students designated as learning disabled are provided with advantages, such as untimed tests, not available to other students).

*Controls*

The normal sample consisted of 60 paid, fluent English-speaking community volunteers who were recruited through newspaper advertisements, flyers, and personal contacts for research projects conducted at Harbor-UCLA Medical Center. The subjects were free of current or past psychotic, major affective, neurologic, and substance or alcohol dependence disorders.

## Procedure

The Rey 15-item Memorization Test and its recognition component were administered as part of a neuropsychological battery. The Rey 15-item Memorization Test, developed by Andre Rey and described by Lezak (1995), consists of a stimulus page with 15 items arranged in 5 rows (3 items per row). The recognition task consists of a page containing the 15 items from the original stimulus interspersed with 15 foils similar to the target items. The recognition stimulus page is reproduced in Figure 1.

Administration instructions for the original recall task were as follows:

> "I'm going to show you a page with *15* different things on it for just a short period of time, and I want you to learn as many of the things as you can. When I take away the page, I'll want you to draw as many of the things as you can remember. Keep in mind, there are *15* different things so you will have to learn them very quickly."

The word "15" was emphasized when giving the above instructions. Subjects were then handed a blank sheet of paper and shown the stimulus page for 10 s. The page was then removed from sight, with the reminder:

> "Now, draw what you can remember."

When the subjects were finished, their reproductions were removed from view and the recognition form was placed in front of them with the instructions:

> "On this page are the 15 things I showed you as well as 15 items that were not on the page. I want you to circle the things you remember from the page I showed you."

Scores used for analysis included: (1) recall correct, (2) recognition correct, (3) false positive recognitions, and (4) a combination score derived from the following equation:

$$\text{recall correct} + (\text{recognition correct} - \text{false positives})$$

# RESULTS

In Table 4 the means, ranges, and standard deviations for age and education, and the gender distribution, for the four groups are shown.

Significant group differences in age were detected [$F(3, 172) = 30.64$, $p = .0001$]. Scheffe tests indicated that the suspect effort group was significantly younger than the normal controls ($p = .0001$), and significantly older than the learning disabled group ($p = .001$); the suspect effort and clinic patient group did not differ in age ($p = .66$). When the control, learning disabled, and clinic patient groups were combined and compared with the suspect effort group on age, no significant difference was found [$t(174) = -.86$, $p = .39$].

Significant group differences were also found in education [$F(3, 172) = 7.57$, $p = .0001$], with Scheffe tests revealing that the suspect effort group had less education than the controls ($p = .04$) and the learning disabled subjects ($p = .0001$); noncredible subjects and clinical patients did not differ in education ($p = .19$). When noncredible subjects were compared with the learning disabled, controls, and clinic patients combined, a significant difference remained in educational level [$t(174) = -3.91$, $p = .0001$].

No significant correlations were observed between age and the various 15-item scores in the suspect effort group (recall correct, $r = -.16$; recognition correct, $r = -.001$; recognition false positives, $r = .22$; combination score, $r = -.14$). However, significant relationships ($p < .05$) were found between education and all test scores

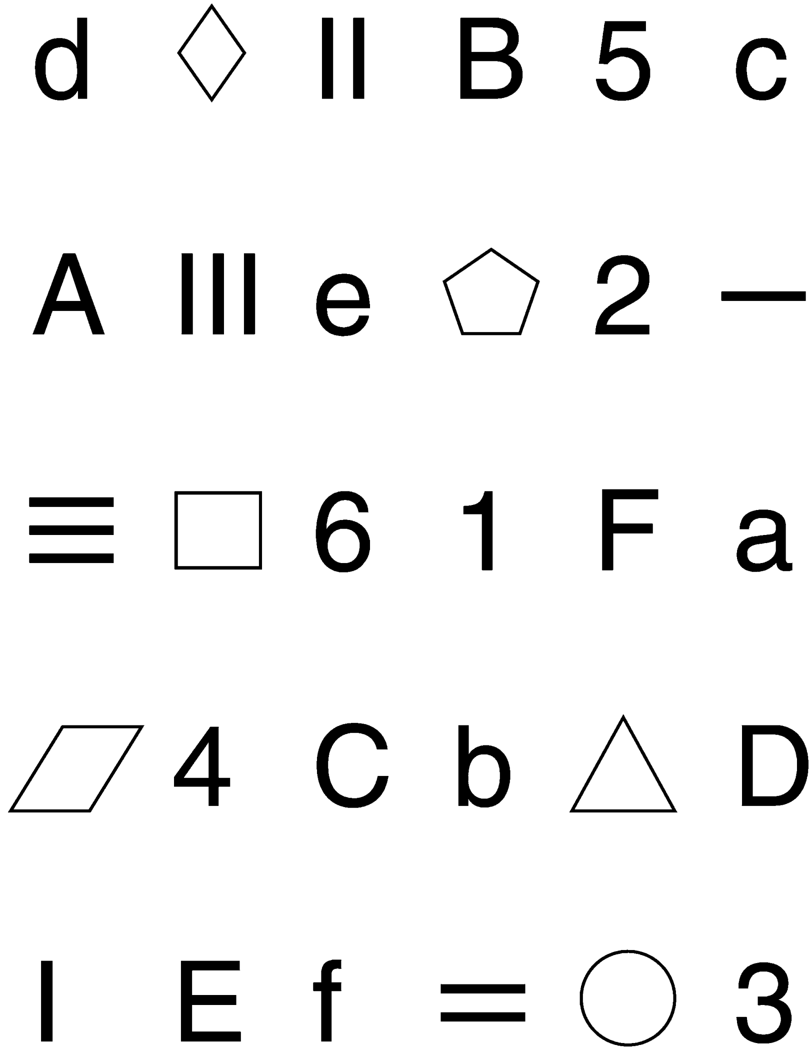



Fig. 1. Rey 15-Item Recognition Stimulus Page.

except recognition correct ($r=.21$; recall correct, $r=.38$; recognition false positives, $r=-.28$; combination score, $r=.37$), although the amount of shared variance did not exceed 14%. In the control, learning disabled, and clinical patient groups combined, significant relationships were found between age and all scores except recognition correct ($r=-.15$; recall correct, $r=-.20$; recognition false positives, $r=.27$; combination score, $r=-.23$), and between education and all scores except recognition false positives ($r=0.15$; total correct, $r=.22$; recognition correct, $r=.23$;

Table 4. Means, Standard Deviations, Ranges, and/or Frequencies for Demographic Variables.

| | Suspect effort | Clinic patients | Learning disabled | Controls |
|---|---|---|---|---|
| *n* | 49 | 36 | 33 | 60 |
| Age | 45.3 ± 12.6 (9–77) | 41.1 ± 13.9 (18–64) | 30.7 ± 12.7 (19–72) | 61.3 ± 19.1 (17–80) |
| Education | 12.9 ± 3.1 (4–20) | 14.1 ± 3.3 (9–22) | 15.7 ± 0.8 (14–18) | 14.4 ± 2.2 (7–18) |
| Gender (m/f) | 24/25 | 25/11 | 15/18 | 16/44 |

Table 5. Mann–Whitney *U* Comparisons Between Suspect Effort Versus Comparison Groups on Rey 15-Item Memorization Test Scores.

| | Recall correct | | Recognition correct | | False positives | | Combination score | |
|---|---|---|---|---|---|---|---|---|
| | *U* | *p* | *U* | *p* | *U* | *p* | *U* | *p* |
| Clinic patients | 389.0 | .0001 | 278.0 | .0001 | 711.0 | .09 | 302.0 | .0001 |
| Learning disabled | 245.0 | .0001 | 205.0 | .0001 | 612.5 | .04 | 186.5 | .0001 |
| Controls | 529.0 | .0001 | 400.5 | .0001 | 1263.0 | .17 | 434.0 | .0001 |

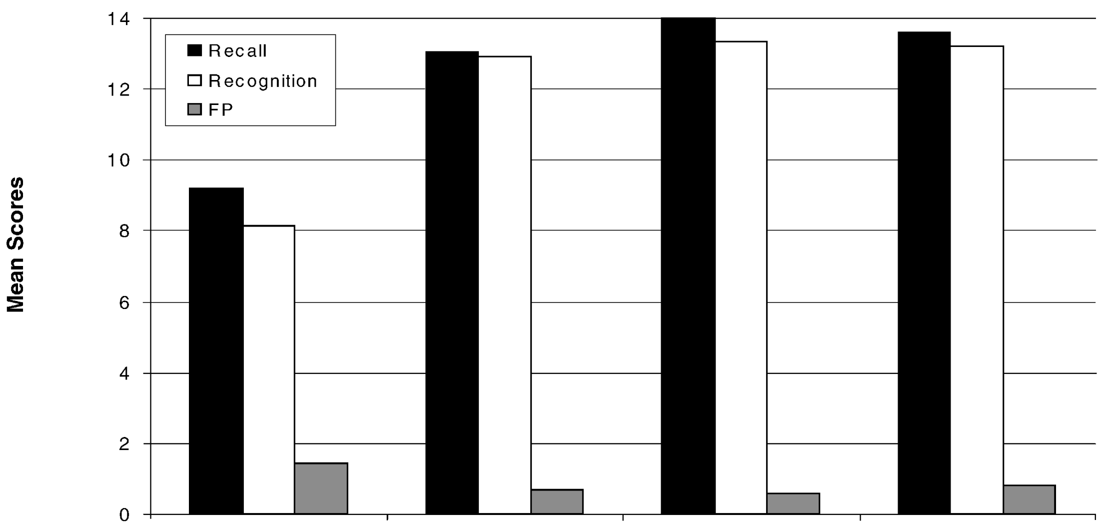
combination score, $r = .26$). However, again, the amount of shared variance was minimal ($< 8\%$).

Because of the modest associations between Rey 15-item Test scores and demographic variables, no attempt was made to adjust demographic differences in group comparisons on test scores.

The Rey 15-item Test scores were not normally distributed, necessitating Kruskal–Wallis analyses. Significant group differences were documented on recall correct ($X^2 = 49.04$, $p = .0001$), recognition correct ($X^2 = 57.20$, $p = .0001$) and combination score ($X^2 = 54.84$, $p = .0001$), but groups did not differ on recognition false positives ($X^2 = 5.41$, $p = .14$). As shown in Table 5, subsequent Mann–Whitney *U* comparisons revealed that the suspect effort group obtained significantly lower recall correct, recognition correct, and combination scores than all three comparison groups. The three comparison groups did not differ from each other on any test variable.

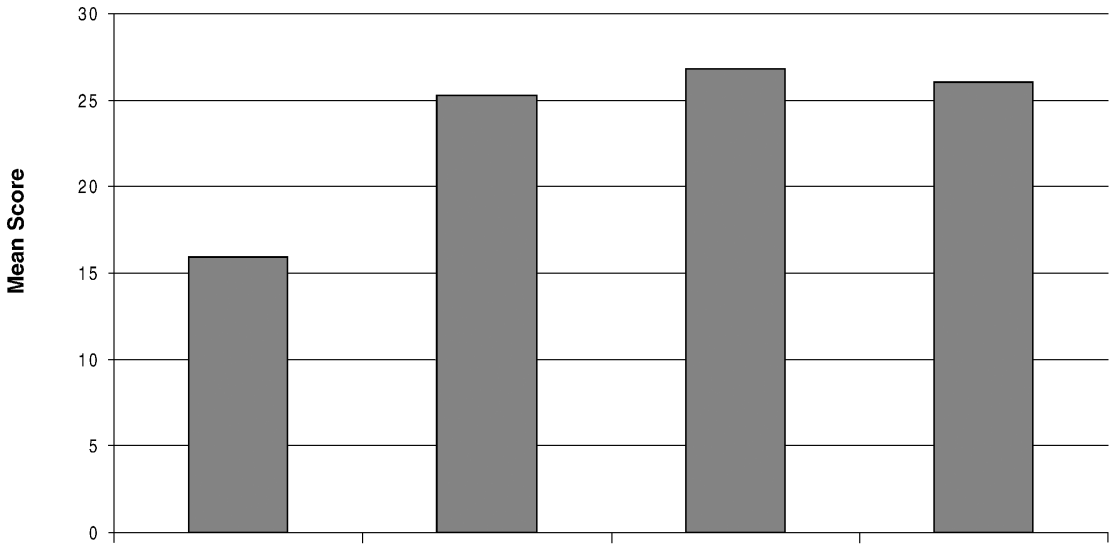
The means, standard deviations, and ranges for each score across groups are depicted in Figures 2 and 3.

In the suspect effort group, a significant correlation was documented between recall correct and recognition correct ($r = .73$), although no significant relationship was found between false positive recognitions and recall correct




| | Suspect Effort | Clinic Patients | Learning Disabled | Controls |
|---|---|---|---|---|
| Recall | 9.18 ± 4.03 (2-15) | 13.06 ± 2.37 (6-15) | 14.00 ± 1.52 (10-15) | 13.62 ± 2.08 (5-15) |
| Recognition | 8.14 ± 4.00 (0-15) | 12.92 ± 2.36 (5-15) | 13.36 ± 1.69 (9-15) | 13.23 ± 2.16 (6-15) |
| False Positive Recognition | 1.43 ± 2.06 (0-10) | 0.69 ± 1.17 (0-4) | 0.58 ± 1.03 (0-3) | 0.83 ± 1.22 (0-4) |

Fig. 2. Means and ranges of Rey 15-item scores for patients with suspect effort and clinical groups.




| | Suspect Effort | Clinic Patients | Learning Disabled | Controls |
|---|---|---|---|---|
| Mean | 15.90 | 25.28 | 26.79 | 26.02 |
| S.D. | 7.95 | 4.84 | 3.33 | 4.30 |
| Range | (2-30) | (8-30) | (19-30) | (12-30) |

Fig. 3. Means and ranges of Rey 15-item combination score for patients with suspect effort and clinical groups.

($r = -.19$) or recognition correct ($r = .01$). In the three comparison groups combined, recall correct was significantly associated with recognition correct ($r = .63$), but not recognition false positives ($r = -.21$), while recognition correct and false positive recognitions were significantly correlated ($r = -.34$).

Cut-off scores, based on visual inspection of the data to maximize sensitivity while maintaining specificity of at least 90% (Baker, Donders, & Thompson, 2000), and confirmed through subsequent ROC analyses, are shown in Table 6.

As can be seen in the table, recognition false positives $\geq 4$ had specificity greater than 93% but only detected 10% of noncredible patients. Recall correct scores and recognition correct scores of $< 9$ identified approximately half of the suspect effort patients with specificity maintained at greater than 94%. However, the best sensitivity was obtained with a combination score $< 20$, with nearly three fourth of the suspect effort subjects correctly assigned and with a specificity of $\geq 92\%$.

Positive and negative predictive power of the various test scores at varying base rates of suspect effort are shown in Table 7. Using the study base rate of 28% noncredible subjects, a cut-off score of $< 20$ for the combination score resulted in a positive predictive value (i.e., proportion of subjects below the cut-off who were actually in the suspect effort group) of 78%, with the negative predictive value (i.e., proportion of subjects above the cut-off who are actually in the nonsuspect effort group) at 90%.

While groups did not differ in total false positive errors, qualitative examination of the types of false positive errors revealed some differences in performance pattern between the suspect effort and comparison groups. As shown in Table 8, 6% of suspect effort subjects circled "f" or "6," while no subject in any of the 3 comparison groups made these errors. In addition, 6% of the suspect effort subjects circled the number "5," while no clinic patients or learning disabled subjects, and only 1 control, committed this error. Similarly, 6% of the suspect effort patients circled a pentagon, while none of the controls and learning disabled, and only 1 of the clinic patients, marked this item.

Seven suspect effort subjects (14%) committed at least one of these four false positive errors,

Table 6. Sensitivity and Specificity Values for 15-Item Memorization Scores.

| Cut-off scores | Sensitivity | Specificity | | |
|---|---|---|---|---|
| | Suspect effort (%) | Clinic patients (%) | Learning disabled (%) | Controls (%) |
| Recall correct | | | | |
| < 3 | 75.5 | 61.1 | 72.7 | 68.3 |
| < 12 | 63.3 | 80.6 | 93.9 | 86.7 |
| < 11 | 31.2 | 86.1 | 97.0 | 93.3 |
| < 10 | 59.2 | 86.1 | 100 | 95.0 |
| < 9 | 46.9 | 97.2 | 100 | 98.3 |
| < 8 | 40.8 | 97.2 | 100 | 98.3 |
| Recognition correct | | | | |
| <13 | 81.6 | 69.4 | 75.8 | 66.7 |
| <12 | 77.6 | 83.3 | 84.8 | 83.3 |
| <11 | 69.4 | 88.9 | 90.9 | 90.0 |
| <10 | 59.2 | 91.7 | 97.0 | 91.7 |
| <9 | 49.0 | 94.4 | 100 | 95.0 |
| <8 | 46.9 | 94.4 | 100 | 98.3 |
| False positive recognitions | | | | |
| ≥ 1 | 51.0 | 63.9 | 69.7 | 60.7 |
| ≥ 2 | 34.7 | 83.3 | 84.8 | 76.7 |
| ≥ 3 | 24.5 | 88.9 | 87.9 | 83.3 |
| ≥ 4 | 10.2 | 94.4 | 100 | 96.7 |
| Combination score | | | | |
| < 25 | 77.6 | 61.1 | 78.8 | 71.7 |
| < 24 | 75.5 | 77.8 | 81.8 | 78.3 |
| < 23 | 75.5 | 80.6 | 84.8 | 78.3 |
| < 22 | 73.5 | 83.3 | 87.9 | 81.7 |
| < 21 | 73.5 | 88.9 | 90.9 | 88.3 |
| < 20 | 71.4 | 91.7 | 93.9 | 91.7 |
| < 19 | 67.3 | 91.7 | 100 | 93.3 |
| < 18 | 61.2 | 91.7 | 100 | 95.0 |

while only two members of the three normal effort groups combined (.02%) made these errors, a highly significant difference ($X^2 = 12.00$, $p =$ .001). Thus, while these errors are not frequent in the suspect effort group, they may serve as virtual pathognomonic signs of noncredible performance when present.

## DISCUSSION

In the current study, standard administration of the Rey 15-item Memorization Test, using a cut-off of < 9 as indicative of suspect effort, revealed a false positive rate of ≤ 5% in clinical and control comparison groups; specificity values were 97% for patients with various neurologic and psychiatric diagnoses referred for neuropsychological testing, 100% for learning disabled college students, and 98% for controls. In contrast, sensitivity was modest, with only 47% of patients with independent evidence of noncredible cognitive symptoms correctly identified.

However, the addition of a recognition trial and use of a combined recall and recognition score (i.e., free recall + [recognition − false positives] < 20) raised sensitivity by 50% while maintaining high specificity across comparison groups

Table 7. Positive Predictive Values (PPV) and Negative Predictive Values (NPV) for Rey 15-Item Memorization Scores at Base Rates of 15%, 28%, and 40%.

| | 15% Base rate | | 28% Base rate | | 40% Base rate | |
|---|---|---|---|---|---|---|
| | PPV (%) | NPV (%) | PPV (%) | NPV (%) | PPV (%) | NPV (%) |
| Recall correct | | | | | | |
| < 13 | 29.0 | 93.6 | 46.8 | 87.9 | 60.7 | 80.9 |
| < 12 | 45.9 | 92.9 | 64.6 | 86.3 | 76.3 | 78.2 |
| < 11 | 57.1 | 92.7 | 75.0 | 86.2 | 83.0 | 78.4 |
| < 10 | 64.0 | 92.8 | 78.4 | 85.8 | 85.7 | 77.5 |
| < 9 | 86.7 | 91.4 | 92.0 | 83.0 | 94.3 | 73.4 |
| < 8 | 84.6 | 90.3 | 90.9 | 81.4 | 93.5 | 71.4 |
| Recognition correct | | | | | | |
| < 13 | 32.8 | 95.5 | 50.6 | 90.9 | 64.4 | 85.2 |
| < 12 | 46.7 | 95.5 | 64.4 | 90.8 | 75.3 | 84.8 |
| < 11 | 55.9 | 94.4 | 72.3 | 88.5 | 81.7 | 81.4 |
| < 10 | 61.5 | 92.8 | 76.3 | 85.7 | 84.0 | 77.3 |
| < 9 | 68.4 | 91.2 | 82.8 | 83.2 | 89.7 | 74.1 |
| < 8 | 81.3 | 91.4 | 88.5 | 82.9 | 91.7 | 73.2 |
| False positive recognition | | | | | | |
| ≥ 1 | 20.3 | 88.1 | 34.7 | 77.4 | 48.0 | 66.0 |
| ≥ 2 | 23.7 | 87.1 | 40.5 | 76.5 | 54.3 | 65.2 |
| ≥ 3 | 25.0 | 86.7 | 40.0 | 75.0 | 53.1 | 63.0 |
| ≥ 4 | 42.9 | 86.0 | 55.6 | 74.0 | 63.6 | 61.7 |
| Combination score* | | | | | | |
| < 25 | 32.3 | 94.7 | 50.0 | 89.2 | 63.2 | 82.4 |
| < 24 | 39.2 | 94.5 | 57.8 | 89.5 | 70.1 | 83.2 |
| < 23 | 40.8 | 94.6 | 59.7 | 89.7 | 72.0 | 83.5 |
| < 22 | 45.5 | 94.8 | 63.2 | 89.3 | 74.3 | 82.4 |
| < 21 | 55.6 | 95.1 | 72.0 | 89.8 | 81.3 | 83.3 |
| < 20 | 61.3 | 94.6 | 77.8 | 89.5 | 85.0 | 83.1 |
| < 19 | 69.2 | 94.1 | 82.5 | 88.4 | 88.9 | 81.5 |
| < 18 | 69.6 | 92.9 | 83.3 | 86.6 | 89.8 | 79.1 |

*Note*. *Combination score = total recall + (true positive recognition − false positive recognition).

(sensitivity = 71%; specificity ≥ 92%). The suspect effort patients significantly differed from comparison groups on recall and recognition scores, but the combination score had substantially higher sensitivity at comparable levels of specificity than any of the individual scores. Individual test scores generally only shared a modest amount of score variance, which predicted that they might be more effective in detecting suspect effort when used in combination, a hypothesis which was confirmed. We have previously found that an index employing several Rey Dot Counting Test variables was also consistently more sensitive in identifying noncredible patients than use of individual scores (Boone et al., in press). Thus, the results from these two studies would suggest that use of combinations of scores rather than single scores may be the most effective approach to detection of suspect effort.

Most previous studies have found generally adequate specificity on the Rey 15-item Memorization Test using a cut-off of < 9 (see Table 2). The minority of reports documenting low specificity have generally included patients with mental retardation (Goldberg & Miller, 1986) or dementia (Guilmette et al., 1994; Philpott, 1992;

Table 8. Tabulation of Individual False Positive Recognition Errors Across Groups.

| | Suspect effort | Clinic patients | Learning disabled | Controls |
|---|---|---|---|---|
| D | 8% (4) | 6% (2) | 9% (3) | 5% (3) |
| E | 6% (3) | 0% (0) | 6% (2) | 3% (2) |
| F | 4% (2) | 0% (0) | 0% (0) | 0% (0) |
| d | 20% (10) | 14% (5) | 12% (4) | 18% (11) |
| e | 8% (4) | 11% (4) | 0% (0) | 3% (2) |
| f | 6% (3) | 0% (0) | 0% (0) | 0% (0) |
| 4 | 6% (3) | 6% (2) | 0% (0) | 3% (2) |
| 5 | 6% (3) | 0% (0) | 0% (0) | 2% (1) |
| 6 | 6% (3) | 0% (0) | 0% (0) | 0% (0) |
|  | 6% (3) | 3% (1) | 9% (3) | 10% (6) |
|  | 18% (9) | 6% (2) | 9% (3) | 17% (10) |
|  | 20% (10) | 6% (2) | 12% (4) | 15% (9) |
|  | 10% (5) | 11% (4) | 0% (0) | 5% (3) |
|  | 6% (3) | 3% (1) | 0% (0) | 0% (0) |
|  | 10% (5) | 6% (2) | 0% (0) | 0% (0) |

Schretlen et al., 1991). Patients with mental retardation and dementia were excluded from the clinic patient group in the current study, which no doubt enhanced specificity values. As a result of this exclusion, it is unknown whether the 15-item recall/recognition paradigm has utility for use in the determination of suspect versus actual mental retardation and dementia.

Control subjects in the present study were on average 15 years older than the suspect effort subjects, yet consistently outperformed the noncredible group. This finding, in conjunction with the modest correlations between age and test scores, and the high specificity values observed in the controls, suggests that the 15-item recall/recognition paradigm is applicable for use across an unrestricted age range.

Similarly, the finding that the learning disabled college students consistently scored better than the suspect effort subjects and had high specificity values across scores, indicates that Rey 15-item recall/recognition performance is not contaminated by the language processing deficits found in learning disability.

It was anticipated that recognition scores would exceed free recall scores in the nonsuspect effort patients given consistent evidence of superior recognition performance in cooperative subjects (Binder, Villanueva, Howieson, & Moore, 1993). However, of some surprise, the comparison groups exhibited mean recognition scores which were either equal to, or slightly lower than, mean free recall scores. In hindsight, the uncoupling of the overlearned associations (i.e., a-b-c, 1-2-3, etc.) in the creation of the recognition stimuli inadvertently made the recognition trial slightly more difficult than the free recall task; specifically, instead of having to recall just a few concepts, subjects had to scan the recognition page for 15 separate items. Fortunately, despite this reversal of the recall-recognition gradient of difficulty, the Rey 15-item recall + recognition paradigm was still effective in detecting suspect effort.

While sensitivity and specificity values are unrelated to base rates of suspect effort, positive and negative predictive power values are contingent on base rates. The base rate of suspect effort in the current study was 28%, a level which is consistent with other reports (Binder, 1990; Binder et al., 1993). However, positive and negative predictive values for the Rey 15-item recall/recognition scores will differ in those settings in which the base rate of noncredible symptoms substantially deviates from 30%. Because of this consideration, we provide positive and negative predictive power values for the various test scores for suspect effort base rates of 15%, 28%, and 40%.

Etcoff and Kampfer (1996) concluded that the Rey 15-item Test in its original format possessed only "tentative" utility in forensic testing, as defined by Rogers (1997) as showing "statistical significance in the expected direction, but (having) little or no practical value in classifying subjects" (p. 14). However, the data from the current study suggest that the addition of a recognition trial to the standard administration format may enable the test to meet standards for "probable" certainty in identifying suspect effort (defined as 75% correct classification of individual subjects).

Similarly, Vallabhajosula and van Gorp (2001) asserted that the Rey 15-item Test did not meet the Daubert standard for admissibility of scientific evidence based on its failure to achieve a criterion of positive predictive value $\geq 80\%$ using a base rate of suspect effort of 30%. However, as shown in the present study, it may be possible to modify the test through brief additions to existing test administration format, thereby enabling it to approach this standard.

## ACKNOWLEDGMENT

Christopher Barr, Ph.D., is acknowledged for his adaptation of the recognition stimuli.

## REFERENCES

Arnett, P.A., Hammeke, T.A., & Schwartz, L. (1995). Quantitative and qualitative performance on Rey's 15-Item Test in neurological patients and dissimulators. *The Clinical Neuropsychologist*, *9*, 17–26.

Back, C.L., Boone, K.B., Edwards, C.T., & Parks, C. (1996). The performance of schizophrenics on three cognitive tests of malingering: Rey 15-Item Memory Test, Rey Dot Counting, and Hiscock Forced-Choice Method. *Assessment*, *3*, 449–457.

Baker, R., Donders, J., & Thompson E. (2000). Assessment of incomplete effort with the California Verbal Learning Test. *Applied Neuropsychology*, *7*, 111–114.

Bernard, L.C., & Fowler, W. (1990). Assessing the validity of memory complaints: Performance of brain-damaged and normal individuals on Rey's task to detect malingering. *Journal of Clinical Psychology*, *46,* 432–436.

Binder, L.M. (1990). Assessment of malingering after mild head trauma and the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology*, *15*, 170–182.

Binder, L.M., Villanueva, M.R., Howieson, D., & Moore, R.T. (1993). The Rey AVLT Recognition Memory Task measures motivational impairment after mild head trauma. *Archives of Clinical Neuropsychology*, *8*, 137–147.

Boone, K.B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., Shamieh, E., & Warner- Chacon, K. (in press). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology.*

Boone, K.B., Lu, P., Sherman, D., Palmer, B., Back, C., Shamieh, E., Warner-Chacon, K., & Berman, N. (2000). Validation of a new technique to detect malingering of cognitive symptoms: The b Test. *Archives of Clinical Neuropsychology*, *15*, 227–241.

Boone, K.B., Savodnik, I., Ghaffarian, S., Lee, A., Freeman, D., & Berman, N. (1995). Rey 15-item Memorization and Dot Counting Scores in a "stress" claim worker's compensation population: Relationship to personality (MCMI) scores. *Journal of Clinical Psychology, 51*, 457–463.

Chouinard, M.J., & Rouleau, I. (1997). The 48-pictures Test: A two-alternative forced-choice recognition test for the detection of malingering. *Journal of the International Neuropsychological Society*, *3*, 545–552.

Davidson, H., Suffield, B., Orenczuk, S., Nantau, K., & Mandel, A. (1991, February). *Screening for malingering using the Memory for Fifteen Items Test (MFIT)*. Poster presented at the International Neuropsychological Society 19th Annual Meeting, San Antonio.

Etcoff, L.M., & Kampfer, K.M. (1996). Practical guidelines in the use of symptom validity and other psychological tests to measure malingering and symptom exaggeration in traumatic brain injury cases. *Neuropsychology Review*, *6*, 171–201.

Frederick, R., Sarfaty, S., Johnston, J.D., & Powel, J. (1994). Validation of a detector of response bias on a forced-choice test of nonverbal ability. *Neuropsychology*, *8*, 118–125.

Goldberg, T.O., & Miller, H.R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology*, *42*, 792–795.

Greiffenstein, M.F., Baker, W.J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*, 218–224.

Greiffenstein, M.F., Baker, W.J., & Gola, T. (1996). Comparison of multiple scoring methods for Rey's malingered amnesia measures. *Archives of Clinical Neuropsychology*, *11*, 283–293.

Griffin, G.A.E., Glassmire, D.M., Henderson, E.A., & McCann, C. (1997). Rey II: Redesigning the Rey Screening Test of malingering. *Journal of Clinical Psychology*, *53*, 757–766.

Griffin, G.A.E., Normington, J., & Glassmire, D. (1996). Qualitative dimensions in scoring the Rey Visual Memory Test of malingering. *Psychological Assessment*, *8*, 383–387.

Guilmette, T.J., Hart, K.J., Giuliano, A.J., & Leininger, B.E. (1994). Detecting simulated memory impairment: Comparison of the Rey Fifteen-Item Test and the Hiscock Forced-Choice Procedure. *The Clinical Neuropsychologist*, *8*, 283–294.

Iverson, G.L., & Franzen, M.D. (1994). The Recognition Memory Test, Digit Span, and Knox Cube Test as markers of malingered memory impairment. *Assessment*, *1*, 323–334.

King, C. (1992). *The detection of malingering of cognitive deficits in a forensic population.* Unpublished doctoral dissertation, California School of Professional Psychology, Los Angeles.

Lee, A., Boone, K.B., Lesser, I., Wohl, M., Wilkins, S., & Parks, C. (2000). Performance of older depressed patients on two cognitive malingering tests: False positive rates for the Rey 15-Item Memorization and Dot Counting Tests. *The Clinical Neuropsychologist*, *14*, 303–308.

Lee, G., Loring, D., & Martin, R. (1992). Rey's 15-item visual memory test for the detection of malingering: Normative observations on patients with neurological disorders. *Psychological Assessment*, *4*, 43–46.

Lezak, M.D. (1983). *Neuropsychological Assessment* (2nd ed.). New York: Oxford University Press.

Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Millis, S.R. (1992). The Recognition Memory Test in the detection of malingered and exaggerated memory deficits. *The Clinical Neuropsychologist*, *6*, 406–141.

Millis, S.R., & Kler, S. (1995). Limitations of the Rey Fifteen-Item Test in the detection of malingering. *The Clinical Neuropsychologist*, *9*, 421–244.

Morgan, S.F. (1991). Effect of true memory impairment on a test of memory complaint validity. *Archives of Clinical Neuropsychology*, *6*, 327–334.

Philpott, L.M. (1992). *The effects of severity of cognitive impairment and age on two malingering tests: An investigation of the Rey Memory Test and Rey Dot Counting Test in Alzheimer's patients and normal middle aged/older adults*. Unpublished doctoral dissertation, California School of Professional Psychology, Los Angeles.

Rogers, R. (1997). Introduction. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 1–19). New York: Guilford Press.

Rogers, R., Harrell, E.H., & Liff, C.D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, *13*, 255–274.

Schretlen, D., Brandt, J., Krafft, L., & van Gorp, W. (1991). Some caveats in using the Rey 15-item Memory Test to detect malingered amnesia. *Psychological Assessment*, *3*, 667–672.

Shamieh, E.W. (1996). *The effects of severe brain damage on three malingering tests: An investigation of the Rey Memory Test, Rey Dot Counting Test, and b Test in head trauma and stroke patients*. Unpublished doctoral dissertation, California School of Professional Psychology, Los Angeles.

Simon, M.J. (1994). The use of the Rey Memory Test to assess malingering in criminal defendants. *Journal of Clinical Psychology*, *50*, 913–917.

Slick, D.J., Sherman, E.M.S., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545–561.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology*, *19*, 500–514.

Vallabhajosula, B., & van Gorp, W. (2001). Post-Daubert admissibility of scientific evidence on malingering of cognitive deficits. *Journal of the American Academy of Psychiatry and the Law*, *29*, 207–215.

Vickery, C.D., Berry, D.T.R., Inman, T.H., Harris, M.J., & Orey, S.A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, *16*, 45–73.

Warner-Chacon, K.R. (1994). *The performance of adult students with learning disabilities on cognitive tests of malingering*. Unpublished doctoral dissertation, California School of Professional Psychology, Los Angeles.