



PII S0887-6177(97)00005-X

Detection of Malingering on the Halstead-Reitan Battery: A Cross-Validation

R. K. McKinzey

Center For Psychological Studies, Oakland, CA

Elbert W. Russell

Kolitz & Russell, South Miami, FL

Using the Halstead-Reitan Battery profiles of 796 people, a formula for the detection of malingering was cross-validated to assess the false positive rate. Subjects included normals, psychiatric cases, and all major types of organics. The formula incorrectly designated 27% of the sample as fakers (i.e., as false positives). Of the 120 head trauma cases, 27 (22.5%) obtained Fake scores, while 93 (77.5%) were correctly assessed as not malingering. © 1997 National Academy of Neuropsychology. Published by Elsevier Science Ltd

The need for the detection of malingering on neuropsychological testing has been well documented (Faust, Ziskin, & Hiers, 1991; Rogers, 1988; Rogers, Harrell & Liff, 1993; Ziskin, 1995). Studies have demonstrated that the Halstead-Reitan Battery (HRB) can easily be faked by adults (Heaton, Smith, Lehman, & Vogt, 1978), adolescents (Faust, Hart, Guilmette, & Arkes, 1988) and children (Faust, Hart, & Guilmette, 1988).

To solve the problem, Mittenberg and colleagues validated a formula to detect such invalid profiles (Mittenberg, Rotholz, Russell, & Heilbronner, 1996). They argued that the use of such a formula adds to “more confidence in the diagnosis” by adding to the convergent evidence “from a variety of sources including the history, behavioral observations, and other test results.” They caution that “the use of any one method in isolation can result in diagnostic errors.” “The method presented here should not be the only basis for evaluating the validity of observed neuropsychological impairment in head trauma patients” (p. 279). Even so, a discriminative function (Table 1) was derived that yielded an overall hit rate of 88.75%, with 16.2% false positives and 6.2% false negatives, using a matched sample of 160 head trauma victims and normal fakers.

TABLE 1
Faking Formula: SCORE

If SCORE is greater than 0, then the profile should be considered malingered.

$$\text{SCORE} = (\text{Categories} \times .01335924) + (\text{TPT Time} \times -.04932242) + (\text{TPT Memory} \times -.1911619) + (\text{Seashore Rhythm correct} \times -.02631231) + (\text{Speech Perception errors} \times .03914169) + (\text{Trails A} \times .01072021) + (\text{Trails B} \times -.01152765) + (\text{Tapping} \times .004032426) + (\text{Sensory Suppressions} \times .02293813) + (\text{FTNW} \times -.02050771) + 1.80943$$

Note. Raw scores are used. Tapping is the sum of the average performance of the right and left hand trials. Sensory suppressions is the sum of all errors on tactile, auditory, and visual trials. Finger Tip Number Writing (FTNW) is the sum of right and left side errors.

However, Mittenberg et al.'s (1996) relatively small sample included no non-faking normals, psychiatric cases, or organics with etiologies other than head trauma. A cross-validation using a much larger and varied population was therefore done.

METHOD

Participants

As part of several large research projects (largely done at the Miami VA Medical Center), such as the development of the HRNES (Russell & Starkey, 1993), 796 subjects had been given the full HRB. All were evaluated by neurologists, who utilized all neurological methods deemed necessary. The cases were followed for at least 1 year to verify that there was no change in diagnosis. The sample's demographics are: age $M = 45$ (range 14–78; $SD = 14$), education $M = 11.96$ (range 0–20; $SD = 3$), 95% male, 90% white, Average Impairment Rating (AIR) (Russell, 1984) $M = 2.15$ (range .08–5.11; $SD = .88$), WAIS-R FSIQ $M = 91.6$ (range 36–137; $SD = 15.2$). The sample was divided into 13 diagnostic categories (Table 2).

None of the Ss had any reason to malingering. They were all evaluated for treatment purposes, not involved in litigation, and not seeking compensation. Anyone suspected of faking had been removed from the sample.

Materials

The formula uses the raw scores of the Category Test, Tactual Performance Test (TPT) (Total Time and Memory), Seashore Rhythm Test, Speech-Sounds Perception Test, Trail Making Test, Finger Tapping Test, and Perceptual Disorders Examination (Suppressions). Each test was given under the supervision of the junior author according to standard procedure (Reitan & Wolfson, 1985; Russell, Neuringer, & Goldstein, 1970). Score ceilings were used for the TPT (10 min per trial), Trails A (180"), and Trails B (300") (Russell & Starkey, 1993).

Procedure

The formula in Table 1 was applied to each subject's raw test scores. The pattern of false positives was explored with a combination of cross-tabulations, correlations, and chi-squares. Cross-tabulations were used to determine the accuracy of the formula results. Correlations were used to examine the relationship of the formula results with the S's age, education, and severity of impairment as measured by the AIR. Chi-squares were used to compare the relationship between the formula result and the severity of impairment for each diagnosis.

TABLE 2
Distribution of SCORE Category, Diagnosis, and AIR Category

AIR Category (AIRC)	True Negatives														Totals	% of N	% of AIRC
	Diagnosis (DX)																
	Neg	Schiz	Ukn	Cong	Tum	Hd tx	Cbv	MS	Neu	Tox	Alc	Inf	Misc				
Normal	97	22	6	0	8	19	7	1	1	0	7	2	2	172	22	81	
Borderline	32	19	6	3	9	31	17	1	5	2	14	2	4	145	18	72	
Mild	4	29	10	4	4	23	21	2	9	4	17	1	4	132	17	70	
Mod	0	6	2	3	6	15	46	3	9	1	10	0	4	105	13	81	
Mod-Sev	0	1	0	2	2	5	6	0	3	1	4	0	1	25	3	38	
Totals ^a	133	77	24	12	29	93	97	7	27	8	52	5	15	579	73		
DX % ^b	84	69	71	75	64	78	67	78	66	57	72	63	68				
N % ^c	17	10	3	2	4	12	12	.9	3	1	7	.6	2				

AIR Category (ARC)	False Positives														Totals	% of N	% of AIRC
	Diagnosis (DX)																
	Neg	Schiz	Ukn	Cong	Tum	Hd tx	Cbv	MS	Neu	Tox	Alc	Inf	Misc				
Normal	16	8	0	0	1	6	4	0	1	0	4	0	1	41	5	19	
Borderline	8	11	5	2	5	4	8	0	0	2	7	2	2	56	7	28	
Mild	1	11	1	0	1	10	14	0	4	2	9	0	3	56	7	30	
Mod	0	3	3	0	3	5	7	0	1	2	0	0	0	24	3	19	
Mod-Sev	0	2	1	2	6	2	15	2	8	0	0	1	1	40	5	62	
Totals ^d	25	35	10	4	16	27	48	2	14	6	20	3	7	217	27		
DX % ^e	16	31	29	25	36	22	33	22	34	43	28	37	32				
N % ^f	3	4	1	1	2	3	6	.3	1.8	.8	2.5	.4	.9				

Note. N = 796. All percentages are rounded. Diagnostic criterion was final neurological decision. Neg = normal; Schiz = schizophrenics; Ukn = unknown diagnosis; Cong = congenital; Tum = tumor; Hd tx = head trauma; Cbv = cerebrovascular (including TIA, CVA, AVM); MS = multiple sclerosis; Neu = neuronal; Tox = toxicity; Alc = Alcoholism; Inf = infectious; Misc = miscellaneous, including degenerative. Mod = Moderate; Mod-sev = Moderate-severe. There were no Severe AIRs in the sample. ^aTotal = sum of Ss with the specific diagnosis (i.e., column) getting a SCORE of Organic (i.e., True Negative). ^bDX % = percentage of Total^a/Total^a + Total^d (i.e., n of specific diagnosis). ^cN % = percentage of Total^a/N. ^dTotal = sum of Ss with the specific diagnosis (i.e., column) getting a SCORE of Fake (i.e., False Positive). ^eDX % = percentage of Total^d/Total^a + Total^d (i.e., n of specific diagnosis). ^fN % = percentage of Total^d/N.

RESULTS

The results of the formula (SCORE) were correlated with the demographics. The correlations were: SCORE and age .073 (*p* < .04), education -.081 (*p* < .03), Average Impairment Rating (AIR) .117 (*p* < .0009). Table 2 presents the distribution of the scores subdivided by AIR Category, Diagnostic category, and SCORE Category.

The overall false positive rate is 27%. Of the 158 Ss with normal neurological diagnoses, 25 (16%) had positive SCOREs. Of the five normal Ss with clearly abnormal HRB profiles, one (20%) produced a positive SCORE. Of the 120 victims of head trauma, 27 (22.5%) had positive SCOREs. Of the 65 Ss with moderate-severe AIRs, 40 (62%) had positive SCOREs, a much higher false positive rate than the lower AIR levels. The cerebrovascular diagnosis (CBV) contributed the most to the overall false positive rate.

The chi-square of diagnosis and SCORE category was not significant, χ^2 (12, N = 796) = 20.35, *p* < .07, although the chi-square of SCORE category and AIR category was: χ^2 (4, N = 796) = 50.926, *p* < .0001. Table 3 presents the chi-squares of SCORE category and AIR

TABLE 3
Chi-squares of SCORE Category by AIR Category Presented by Diagnosis

	Diagnosis												
	Neg	Schiz	Ukn	Cong	Tum	Hd tx	Cbv	MS	Neu	Tox	Alc	Inf	Misc
<i>df</i>	2	4	4	3	4	4	4	4	4	3	4	3	4
<i>n</i>	158	112	34	16	45	120	145	9	41	14	72	8	22
χ^2	.82	2.735	10.705	4.27	8.327	3.865	24.221	9	12.756	1.75	6.718	3.73	2.57
<i>p</i> <	ns	ns	.04	ns	ns	ns	.0001	ns	.02	ns	ns	ns	ns

Note. Total $N = 796$. ns = not significant, i.e., $p < .05$. Diagnostic criterion was final neurological decision. Neg = normal; Schiz = schizophrenics; Ukn = unknown diagnosis; Cong = congenital; Tum = tumor; Hd tx = head trauma; Cbv = cerebrovascular (including TIA, CVA, AVM); MS = multiple sclerosis; Neu = neuronal; Tox = toxicity; Alc = Alcoholism; Inf = infectious; Misc = miscellaneous, including degenerative.

category by each diagnosis: note the highly significant chi-square of the cerebrovascular diagnosis. When the *Ss* with moderate-severe AIR and/or a cerebrovascular diagnosis were removed from the sample, the overall false positive rate fell to 23.7%, and the correlations of SCORE with age, education, and AIR dropped to nonsignificant levels. Mittenberg et al. (1996) comment that no organic got a SCORE $>.39$. In contrast, 13% of our sample had SCOREs $>.39$. Of the 120 head trauma victims, 15 (12.5%) had a SCORE $>.39$.

DISCUSSION

Clearly, Mittenberg et al.'s (1996) formula for the detection of faking on the HRB needs improvement. However, it is currently the only such formula available, and perhaps can be used for screening. Those needing a more accurate method of detecting malingering may wish to consider the formula validated and cross-validated for the Luria-Nebraska Neuropsychological Battery (McKinzey, Podd, Krehbiel, Mensch, & Trombka, in press).

If a person is referred for evaluation after head trauma, that person would, a priori, be most like the 278 people comprising the combined Negative (i.e., normal) and Head Trauma groups. If that person received a completely normal HRB AIR (i.e., 1), there is a 16% chance of a spurious Fake SCORE. If that person receives a more impaired AIR (2–5), the chances rise to 21%. This 79% true negative rate is not that different from the HRB's hit rates (Russell, 1995).

The results of the formula obviously cannot be used mechanically. The diagnosis of malingering is a clinical one, and a positive SCORE should be only one element in such a decision.

A person with a highly elevated HRB profile ought to also have corresponding etiology, course, symptoms, and neurological findings. In such a profile, a positive SCORE would mean little. Without such a consistent history, the consideration of a positive SCORE would raise some questions about the validity of the profile, but a malingering diagnosis would be counterbalanced by the high false positive rate in such populations. Likewise, a profile of a CBV diagnosis ought to be accompanied by consistent collateral information.

REFERENCES

- Faust, D., Hart, K. J., & Guilmette, T. J. (1988). Pediatric malingering: The capacity of children to fake believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology*, *56*, 578–582.
- Faust, D., Hart, K. J., Guilmette, T. J., & Arkes, H. R. (1988). Neuropsychologists' capacity to detect adolescent malingerers. *Professional Psychology: Research & Practice*, *19*, 508–515.

- Faust, D., Ziskin, J., & Hiers, J. (1991). *Brain damage claims: Coping with neuropsychological evidence*. Los Angeles, CA: Law and Psychology Press.
- Heaton, R. K., Smith, H. H., Lehman, R. A., & Vogt, A. T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology*, **46**, 892–900.
- McKinzey, R. K., Podd, M. H., Krehbiel, M. A., Mensch, A. J., & Trombka, C. C. (in press). Detection of malingering on the Luria-Nebraska Neuropsychological Battery: An initial and cross-validation. *Archives of Clinical Neuropsychology*.
- Mittenberg, W., Rotholz, A., Russell, E., & Heilbronner, R. (1996). Identification of malingered head trauma on the Halstead-Reitan Battery. *Archives of Clinical Neuropsychology*, **11**, 271–281.
- Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.
- Rogers, R. (1988). *Clinical assessment of malingering and deception*. New York: The 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.
- Russell, E. (1984). Theory and development of pattern analysis methods related to the Halstead-Reitan Battery. In P. E. Logue & J. M. Schear (Eds.), *Clinical neuropsychology: A multidisciplinary approach* (pp. 50–98). Springfield, IL: Charles C Thomas.
- Russell, E. (1995). The accuracy of automated and clinical detection of brain damage and lateralization in neuropsychology. *Neuropsychology Review*, **5**, 1–68.
- Russell, E., Neuringer, C., & Goldstein, G. (1970). *Assessment of brain damage: A neuropsychological approach*. New York: John Wiley & Sons.
- Russell, E. W., & Starkey, R. (1993). *Halstead Russell Neuropsychological Evaluation System (HRNES)*. Los Angeles, CA: Western Psychological Services.
- Ziskin, J. (1995). *Coping with psychiatric and psychological testimony* (5th ed.). Los Angeles, CA: Law and Psychology Press.