

Norming subjects for the Halstead Reitan battery

Elbert W. Russell*, Miami Florida

S. Dixie Highway, Suite PH-III, Miami, FL 33156, USA

Accepted 12 November 2004

Abstract

The adequacy of neuropsychological norms is dependent on the characteristics of the subjects whose data are used. Volunteer participants and neurologically normal participants in major normative studies show markedly different patterns of test performance that reflects on the inadequacy of using volunteer participants to develop norms. In this study when all of the Halstead Reitan Battery norming studies with an *N* of 200 or more were examined, Wechsler FSIQ score differences of approximately 1 standard deviation above average for the volunteer normative participants was found. The norms from the norming study using neurologically normal patients were essentially average. The exclusion of neurologically suspect participants from volunteer normative studies lead to a bias reflected in artificially inflated levels of performance, and a restricted range of variation. This sets inappropriately rigid assessment cutoffs for defective performance. Furthermore, data collection from neurologically normal subjects follows the same format as that used in the assessment of neurologically compromised subjects, while the more research oriented protocol used to collect normative data from volunteer participants does not.

© 2004 National Academy of Neuropsychology. Published by Elsevier Ltd. All rights reserved.

Keywords: Halstead Reitan battery; Neuropsychological test; Volunteers

Neuropsychological test norms typically come from one of two sources, volunteer participants and clinical patients found to be neurologically normal. There are arguments criticizing subjects that are not volunteers. In part Mitrushina, Boone, and D'Elia (1999) selectively eliminated a major set of norms from their Handbook (Russell, 2003), which were not volunteers. On the other hand, Reitan (Reitan & Wolfson, 1993 (pp. 33–35), 1997) contended that neurologically normal subjects represent the ideal group for neuropsychological assessment controls.

* Corresponding author.

E-mail address: ewr@bellsouth.net (E.W. Russell).

Nevertheless, this volunteer population presents a number of confounds that limit the usefulness of their data in interpreting neuropsychological test scores collected from persons with known brain pathology. These confounds include differences in data collection methods, screening out pathological cases from the volunteers, self-selection biases among the volunteers, and a restriction of range in volunteer-based norms that leads to inappropriately low cut off scores for identifying defective performances.

The main argument in favor of using volunteer subjects is their screened and confirmed “normalcy” in contrast to neurologically normal subjects, who had enough problems to warrant referral for neuropsychological evaluation. Regrettably, this screening ensures that the volunteer normative populations are never truly random, and the deviation from random selection is in the direction of raising the bar for normalcy. In *Dodrill's (1987)* control group, nearly one-third of the original sample was excluded when screened for evidence of organic or psychiatric dysfunction. Similarly, *Fastenau (1998)* excluded nearly a quarter of randomly selected older participants who had no known neurological history because many had possible histories of undiagnosed cerebrovascular problems.

Volunteer normative participants select themselves for studies, and are often encouraged to do so with compensation for participation. They may not always have the motivation to perform optimally, and effort measures to ensure their effortful performance have never been done (*Green, Rohling, Lees-Haley, & Allen, 2001*). Lacking evidence of poor effort, some normative studies have reported unexpected results showing greater impairment among normal than among patient groups in the domains of memory (*Haaland, Linn, Hunt, & Goodwin, 1983*), psychomotor functions (*Pauker, 1981*), complex acoustic functions (*Bornstein, 1985*), or speeded information processing (*Gronwall, 1977*).

Thus, one set of variables causes volunteers to provide data that raises the ceiling for normalcy, while another set of variables may cause the performance of volunteers to under represent “normal” performance. What is the empirically driven neuropsychologist to do? One solution might be to use a well-established method for judging the general ability level of the normative sample, a procedure outlined by *Anastasi and Urbina (1997)* under the rubric “national anchor norms” (p. 70). Of all the tests in our armamentarium, none is a better candidate for universal standard than the Wechsler tests.

In the meta-analysis reported by *Stanczak, Stanczak, and Templer (2000)*, significant differences were observed in the Wechsler scores of volunteer participants when compared with psychiatric, neurological, and normal neurological comparison groups. These differences came about because of the superior intellectual performance of the volunteer participants, and not because of defective performance among the other groups. The mean WAIS-R FSIQ score for the self-report volunteer group was almost 112 (111.6) (D.E. Stanczak, personal communication, March 27, 2003). Not surprisingly, the volunteer participants averaged nearly 3 years more education than the referral participants.

The higher than average FSIQ for volunteers was also supported by a reexamination of a review of smaller studies (*Steinmeyer, 1986*). For the norms with IQ scores, the nine volunteer groups had a mean FSIQ of 116.9, while the five neurologically normal groups had a FSIQ of 103.6. The observation that the volunteer participants in these studies had considerably higher IQ scores than the referral participants raises the possibility that a similar bias exists in other norming studies that used volunteer participants.

To test this possibility, the participants in the five large scale normative studies of the Halstead Reitan battery, with an $N > 200$ and which had Wechsler (1955, 1981) IQ measures for the sample, were examined. The specific hypothesis tested was that volunteer normative participants have considerably higher IQ scores than normal neuropsychological subjects who are referred as part of a neurological examination.

1. Methods and results

Subject demographics reported in the normative studies of Fromm-Auch and Yeudall (1983), Heaton, Grant, and Matthews (1991), Pauker (1981), Russell and Starkey (2001) and Yeudall, Reddon, Gill, and Stefanyk (1987). Of these, only the norms of Russell and Starkey (2001) are composed entirely of neurologically normal subjects; the remaining four studies used at least a large portion of volunteers in their norming. The norms for Heaton et al. (1991) were used in this examination since the revision by Heaton, Miller, Taylor, and Grant (2004) did not provide IQ levels for its Caucasian subjects. Since the Heaton et al. (1991) data were used for the Heaton et al. (2004) norms the IQ levels of the Heaton et al. (2004) norms is probably approximately the same. These data are presented in Table 1.

This demonstrates that the mean FSIQ of the volunteer subjects was 115, a full standard deviation above average. Clearly, then, these “normal” volunteer participants were not normal, but represent the upper one-sixth of the population. Only the norms of Russell and Starkey (2001), using referred but neurologically normal participants, showed a more average IQ of 102.

2. Discussion

Much of the controversy concerning the appropriate groups to use for normative studies comes from a failure to differentiate between the demands of the research laboratory versus those of clinical practice. Rigorous exclusion and scrubbing of participants in research projects helps to ensure that a rigorous comparison between brain damaged and control subjects is possible. When applied to norms development, however, the clinical assessment question becomes one of differentiating between persons with a neurological condition and those without. For this, the most appropriate reference group is persons who are neurologically evaluated and found to be neurologically normal.

In addition, the use of referred, but neurologically normal participants in normative studies is mandated by the Standards of Educational and Psychological Testing (AERA, 1999). It specifies that norms should represent the same population as their intended use (p. 55), and norms should be collected under the same testing conditions as will be used in the intended assessment procedure (p. 63).

Some might argue that referred but neurologically normal participants may actually be abnormal with some undiagnosed pathology. This issue was addressed by Russell (1990), who described a series of 200 referred but neurologically normal subjects followed for over a year after their initial presentation and evaluation. Over half were initially diagnosed with

Table 1
Descriptive statistics for the co-normed HRB Studies

Characteristics	Sets of norms				
	Pauker	Fromm-Auch Yeudall	Yeudall et al.	Heaton et al. ^a	Russell and Starkey ^b
Date Published	1981	1983	1987	1991	2001
<i>N</i>	363	193	225	378	200
Age, \bar{x}	39.0	25.4	24.7	41.8	44.6
Age range	19–76	15–32	15–40	20–81	18–89
Education, \bar{x}	12.5	14.8	14.6	13.6	12.7
FSIQ, \bar{x}	115, WAIS	119, WAIS	112, WAIS-R	114, WAIS	102, WAIS-R
Norming method	Stratified	Stratified	Stratified	Regression	Regression
Demographic categories	IQ and age 4 × 3	Age and sex 4 × 2	Age and sex 4 × 2	Ed., age, sex 6 × 9 × 2	Ed., age, sex ^c 4 × 6 × 2
Location ^d	Canada	Canada	Canada	CO, WI, CA ^e	FL, OH
Subjects type	Volunteer	Volunteer	Volunteer	Unclear ^f	Negative neurological
Tests covered ^g	HRB(8)	HRB(13)	HRB(13)	HRB(12)+W+9	HRB(12)+W+12

^a The Heaton et al. (2004) norms were not used in this table since the FSIQ for the Caucasian subjects was not provided.

^b The HRNES (Russell & Starkey, 1993) has the same demographic statistics as the HRNES-R (Russell & Starkey, 2001).

^c The division by sex was restricted to tapping, grip strength and pegboard.

^d USA state or Canada.

^e Only three areas are fully represented (see Russell, 1997).

^f Manual is not clear as to type of subjects.

^g HRB(): number of HRB tests; W: WAIS or WAIS-R subtests; +number: number of non HRB tests added.

psychological or minor physical ailments. None of this group showed evidence of neurological problems at any point during the following year. These data suggest that concerns about undiagnosed pathology among neurologically normal subjects might rarely be correct, but is largely a red herring that draws attention away from the relevance and importance of using such a group of patients for comparison purposes.

As this study indicates, volunteer subjects are not representative of the “normal” person who undergoes clinical evaluation. Norms based on volunteer subjects run the risk of increasing false diagnoses of pathology by raising the bar for what is normal. Furthermore, the scrupulous scrubbing and selecting of volunteer participants apparently leads to a more homogenous group with a restricted range of variation. This range restriction further increases the likelihood of false positive diagnoses whenever cutting scores are derived from the performance of volunteer participants. Reliance on volunteer norms increases the chances that a neurologically normal person of average intelligence will be misdiagnosed as brain damaged. This concern is important in our clinical arena, and even more important in the forensic setting.

References

- American Educational Research Association and the American Psychological Association & the National Council on Measurement in Education (AERA). American Educational Research Association and the American Psychological Association & the National Council on Measurement in Education (AERA). (1999). *Standards for educational and psychological testing*. Washington: American Educational Research Association.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Bornstein, R. A. (1985). Normative data on selected neuropsychological measures from a nonclinical sample. *Journal of Clinical Psychology, 41*, 651–659.
- Dodrill, C. B. (1987). What constitutes normal performance in clinical neuropsychology? In *Paper presented at the 97th annual convention of the American Psychological Association*.
- Fastenau, P. S. (1998). Validity of regression-based norms: An empirical test of the comprehensive norms with older adults. *Journal of Clinical and Experimental Neuropsychology, 20*, 906–916.
- Fromm-Auch, D., & Yeudall, L. T. (1983). Normative data for the Halstead–Reitan Neuropsychological tests. *Journal of Clinical Neuropsychology, 5*, 221–238.
- Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15*(12), 1045–1060.
- Gronwall, D. M. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual & Motor Skills, 44*(2), 367–373.
- Haaland, K. Y., Linn, R. T., Hunt, W. C., & Goodwin, J. S. (1983). A normative study of Russell’s variant of the Wechsler memory scale in a healthy elderly population. *Journal of Consulting and Clinical Psychology, 51*, 878–881.
- Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead–Reitan battery (norms, manual and computer program)*. Odessa, FL: Psychological Assessment Resources.
- Heaton, R. K., Miller, S. W., Taylor, M. J., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead–Reitan battery (norms, manual and computer program)*. Odessa, FL: Psychological Assessment Resources.
- Mitrushina, M. N., Boone, K. B., & D’Elia, L. F. (1999). *Handbook of normative data for neuropsychological assessment*. New York: Oxford University Press.
- Pauker, J. D. (1981). *Adult norms or standards for the Halstead–Reitan Neuropsychological test battery*. Clarke Institute of Psychiatry and University of Toronto Canada, Unpublished manuscript.
- Reitan, R. M., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological test battery: Theory and clinical interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

- Reitan, R. M., & Wolfson, D. (1997). Emotional disturbances and their interaction with neuropsychological deficits. *Neuropsychological Review*, 7, 3–19.
- Russell, E. W. (1990). Three validity studies for negative neurological criterion norming. In *Paper presented at the 98th annual convention of the American Psychological Association*, Boston, Unpublished.
- Russell, E. W. (1997). Developments in the psychometric foundations of neuropsychological assessment. In G. Goldstein & T. M. Incagnoli (Eds.), *Contemporary Approaches to Neuropsychological Assessment* (pp. 15–65). New York: Plenum.
- Russell, E. W. (2003). The critique of the HRNES. In *The handbook of normative data for neuropsychological assessment. Archives of Clinical Neuropsychology*, 18 (2), 165–180.
- Russell, E. W., & Starkey, R. I. (1993). *Halstead, Russell neuropsychological evaluation system [manual and computer program]*. Los Angeles: Western Psychological Services.
- Russell, E. W., & Starkey, R. I. (2001). *Halstead, Russell neuropsychological evaluation system—revised [manual and computer program]*. Los Angeles: Western Psychological Services.
- Stanczak, E. M., Stanczak, D. E., & Templer, D. I. (2000). Subject-selection procedures in neuropsychological research: A meta-analysis and prospective study. *Archives of Clinical Neuropsychology*, 15(7), 587–601.
- Steinmeyer, C. H. (1986). A meta-analysis of Halstead–Reitan test performances on non-brain damaged subjects. *Archives of Clinical Neuropsychology*, 1, 301–307.
- Yeudall, L. T., Reddon, J. R., Gill, D. M., & Stefanyk, W. O. (1987). Normative data for the Halstead–Reitan Neuropsychological tests stratified by age and sex. *Journal of Clinical Psychology*, 43(3), 346–367.
- Wechsler, D. (1955). *Wechsler Adult Intelligence Scale manual*. New York: Psychological Corporation.
- Wechsler, D. (1981). *WAIS-R, Wechsler Adult Intelligence Scale—Revised manual*. San Antonio, TX: Psychological Corporation.