

Some Problems Inherent in Neuropsychological Testing¹

by Brett C. Trowbridge, Ph.D., J.D.

and James W. Schutte, Ph.D.

Introduction

Neuropsychological tests are used to determine whether an individual has brain damage (brain dysfunction) and if so, the extent and cause of that brain damage.

Neuropsychological tests use samples of the individual's behavior (test performance) to make inferences about brain functioning. In order for neuropsychological testing to be of use, the practitioner must be qualified, must use standardized tests with official norms, must be able to address basic statistical issues regarding testing, and must be able to address threats to test validity, including inappropriate norms, limited English proficiency on the part of the patient, and possible malingering.

Qualifications for Neuropsychologists

Neuropsychology is a specialty area in psychology, although there is no generally agreed-upon definition of what training and experience a psychologist must have before he can call himself a "neuropsychologist", and thus there is no legal rule restricting any psychologist from advertising himself as a "neuropsychologist" even if he has little or no training or experience in neuropsychology. Professionals involved in legal cases involving neuropsychologists should inquire what specialized training and experience a psychologist has that he feels justifies using the term "neuropsychologist"; those with specialized full-time post-doctoral training supervised by neuropsychologists in settings

¹ This article was based, in part, on a presentation made by the second author at the 19th Annual Symposium in Forensic Psychology of the American College of Forensic Psychology, Rancho Mirage, CA, in 2003.

that have access to a wide variety of different types of brain-damaged individuals (such as university or hospital-based training programs) are obviously more qualified to call themselves “neuropsychologists” than those who have not had such training.

Reliability

Reliability refers to the degree to which test scores are free from errors of measurement, and is often defined as an indication of a test’s consistency between two or more administrations or ratings of that test. To the extent that a test is unreliable it cannot be valid, because a test’s reliability establishes the upper limit for that test’s validity: psychologists who use tests should know to what extent differences between forms or administrations of a test reflect simply errors of measurement as opposed to signifying actual differences in underlying abilities. Psychological tests are generally more prone to errors of measurement than measurements in “hard sciences”—for example, different measurements of the temperature at which water boils will vary by only fractions of a degree, but when measuring an individual’s intelligence the “standard error of measurement” is, on average, plus or minus 5 points, so that a person who receives a full scale score of 105 on an IQ test likely has an actual IQ ranging between 100 and 110. If the same test is administered on two occasions, obtaining scores of 104 and 109, these scores would be considered in practice to be the same. The “standard error of measurement” is published in the manual or in the professional literature for each test. For those tests requiring scoring on subjectively rated criteria, inter-rater reliability is particularly important, i.e., the extent to which different raters rating the same behavior by the same subject come up with the same score.

For example, an IQ of 69 is indicative of mild mental retardation, whereas an IQ of 72 is not. However, if one takes into account the “standard error of measurement”, these scores are not significantly different from each other, yet each will lead to very different diagnostic conclusions (and may even be a case of life or death, as in death penalty cases). Fortunately, a diagnosis of mental retardation is not based solely on IQ, but also on the presence of deficits in different daily living skills. In order to make an accurate diagnosis, neuropsychological test results should also be consistent with a person’s observed behavior in daily life.

Professionals involved in legal cases need to ask neuropsychologists not just what score an individual received on each test, but what the “standard error of measurement” is, as small differences in scores may actually not represent true differences in underlying ability or performance. Some neuropsychological tests are frequently criticized as having such poor reliability that the instrument is really not appropriate for use. For example, the Wisconsin Card Sorting Test has been criticized as having such poor reliability that it is not appropriate for clinical use (as opposed to research use).² Professionals wishing to reference reliability figures for various neuropsychological tests should refer to standard texts such as *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (2nd ed.) and *Reliability and Validity in Neuropsychological Assessment* (2nd ed.).³

² Bowden, S.C., et al. (1998). The reliability and internal validity of the Wisconsin Card Sorting Test. *Neuropsychology Rehabilitation*, 8, 243-254.

³ Spreen, O., and Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York, Oxford University Press, and Franzen, M.D. (2000). *Reliability and validity in neuropsychological assessment* (2nd ed.). New York, Kluwer Academic, Plenum Publishers

Standardization

To be defined as a “test”, an instrument must be “standardized”, which means that a procedure for administering and scoring the test needs to be specified. Users of the test need to follow the rules for administering and scoring the test precisely, because if they do not the reliability and validity of the test will be compromised. In some instances the administration procedures and scoring may be quite complex, and may require substantial training. It is not at all uncommon for neuropsychologists to administer tests incorrectly, or to make errors in scoring them.⁴ Professionals should ask for the “raw data”, which usually consists of the scoring sheets provided by the test publisher on which the neuropsychologist has written the individual’s responses, and then has scored them. They can then provide the raw data to another neuropsychologist to determine whether the test was correctly administered and scored.

Equipment, test instructions, and testing procedures that neuropsychologists use when administering some tests may vary considerably, sometimes due to the fact that there are different versions of tests, or different administrative instructions (manuals) in use for a given tests. Some test-users devise their own shortcuts, or “short-forms” of tests that do not comply with standardized procedures. Some users have developed computerized versions of tests that the individual completes on his own, so that the test does not have to be individually administered. Some users have developed “paper and pencil” forms of some tests to avoid having to carry around bulky equipment. Such variations can alter test scores considerably.

⁴ See Berent, S. and Swartz, C. (1999). Essential psychometrics, pp. 3-26 in Sweet, J. (Ed.), *Forensic Neuropsychology: Fundamentals and practice*, Swets & Zeitlinger, Exton, PA.

Use of Psychometricians to Administer Tests

Because neuropsychologists often use batteries consisting of large numbers of time-consuming tests, most do not administer some or all of the tests themselves, instead using “psychometricians” or technicians to do the administration, and sometimes even to do the scoring. A few neuropsychologists even have other people do the history taking and interviewing, so that they never actually see the client themselves.⁵ The use of technicians is “not without controversy”, because it decreases neuropsychologists’ contact with the examinees.⁶ There is no standard training for psychometricians, so some neuropsychologists may simply give brief training to secretaries or other members of their office staff, calling into question whether such persons have enough training and skill to follow complex rules of test administration and/or scoring. Professionals working on cases involving neuropsychologists should always find out who administered which tests and what the training of non-psychologists consisted of.

Norms

“Norms” are the tables used to compare an individual’s “raw scores” on tests with the performances of “normal” people. Most psychological tests routinely used by psychologists have official, national norms. Developers of these tests have given the tests to large numbers of people from many different geographic areas of the United States and have included representative numbers of persons of all relevant age groups, both genders,

⁵ See Sweet, J.J. and Mobery, P.J. (1990). A survey of practices and beliefs among ABPP and non-ABPP clinical of neuropsychologists. *The Clinical Neuropsychologist*, 4, 101-120, for data showing that about three quarters of neuropsychologists use technicians to assist in testing, and showing that about 5 percent do not even personally interview the clients.

⁶ Putnam, S.H. and DeLuca, J.W. (1990). The TCN salary survey: A survey of neuropsychologists in primary employment and private practice settings. *The Clinical Neuropsychologist*, 4, 199-244.

all minority groups, and various social classes, as well as people of differing educational levels, and sometimes even people with various levels of intellectual functioning. Ideally, when an individual takes these tests, his score(s) can therefore be compared with others of his gender, race, age, social class, and educational level. Norms developed in this way can be considered representative of the population as a whole. The norms will have to be periodically updated, as the population is gradually changing, so that norms from many years ago may not be truly representative of today's population, and the material for the items on the tests also has to be updated in order that the test-items be appropriate for today's population. For those reasons most psychological tests periodically come out in new iterations. For example, the Wechsler Adult Intelligence Scale (WAIS) was revised to the Wechsler Adult Intelligence Scale-Revised (WAIS-R), and the latest edition of the test is the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III).

Indeed, probably because of better schools, better medicine and nutrition, and better dissemination of information, people are gradually becoming “smarter” on average, at least in the sense that they are scoring higher on average on IQ tests. This effect is so pronounced that its size has even been estimated—it is thought to be on average about three-tenths of an IQ point a year.⁷ This means that an IQ score derived from an older IQ test which has not recently been “re-standardized” will probably over-estimate a person's IQ, on average about three tenths of an IQ point times the number of years ago that the test was published. For example, on the WAIS-R (published in 1981) people will score on average about 4.8 points higher than on the WAIS-III (published in 1997), 16 years later ($16 \times .3 = 4.8$). For that reason, it is not really correct to directly compare an

⁷ Psychological Corp. (1997). WAIS-III, WMS-III. Technical Manual. Harcourt, Brace & Co., San Antonio, Texas, Table 41, p. 79.

individual's IQ as measured at some time in the past on an older version of an IQ test with his current scores on the modern version of the test.

Norms for these types of tests can be considered to be authoritative, as it is clear to everybody just which norms should be used in scoring these tests, and there is little confusion or debate as to what the scores mean. Examples of these types of tests are intelligence tests such as the WAIS-III, and the Kaufman Brief Intelligence Test Second Edition (K-BIT2); achievement tests such as the Wide-Range Achievement Test Third Edition (WRAT-3) and the Wechsler Individual Achievement Test Second Edition (WIAT-II); memory tests such as the Wechsler Memory Scale Third Edition (WMS-III); personality tests such as the Minnesota Multiphasic Personality Inventory-2 (MMPI-2), the Million Clinical Multiaxial Inventory Third Edition (MCMI-III), and the Personality Assessment Inventory (PAI); and tests to determine if the individual is malingering, such as the Test of Memory Malingering (TOMM), the Validity Indicator Profile (VIP), and the Structured Interview of Reported Symptoms (SIRS). All of these tests have authoritative official national norms representative of the population as a whole, so that psychologists will always be in agreement as to how an individual's score or set of scores compares to national averages.

Neuropsychologists evaluate people to try to assess their "neuropsychological" functioning. In doing so they almost always use some of the tests with authoritative and updated national norms representative of the population as a whole mentioned above, but they often also employ special "neuropsychological" tests which were originally developed for research purposes, for which the norms were developed, sometimes many years ago, from small samples that even then were unrepresentative of the population as a

whole. Oftentimes there will be competing sets of norms, such that an individual's score as measured by one set of norms could be above average but could be below average according to another set of norms. These competing norms, and norms derived from small unrepresentative samples, create confusion and debate about what different scores actually mean. Oftentimes these norms are based exclusively on very narrow demographic groups, such as college students, or other highly educated, highly intelligent persons.

Mitrushina and her colleagues report, for example, that there are at least 24 different sets of norms for the Trail-Making Test, 13 different sets of norms for the Controlled Oral Word Association Test (COWA), 37 different sets of norms for the Finger-Tapping Test, and 19 different sets of norms for the Category Test.⁸

In a research report written by Kalechstin and his colleagues, results showed that the same scores on five neuropsychological tests ranged from the impaired range to the average range, depending upon the norms that were used.⁹ The fact that competing norms exist allows neuropsychologists and those who hire them to “norm-shop” depending on whether they want the individual's scores to appear to be relatively lower or higher.

Probably the most commonly used set/fixed battery of neuropsychological tests is the Halstead-Reitan Battery; neuropsychologists often use some of these tests, and sometimes all of them, in order to determine whether an individual is “brain-damaged”, and if so, what his “impairment index” is. The original norms (and associated cut-off scores to use in determining whether a score was in the “impaired” range) for the

⁸ Mitrushina, M.N., Boone, K.B. & D'Elia, L.F. (1999). Handbook of normative data for neuropsychological assessment. New York, Oxford University Press.

⁹ Kalechstine, A.D., VanGorp, W.G. & Rapport, L.J. (1998). Variability in clinical classification of raw test scores across normative data sets. *The Clinical Neuropsychologist*, 12, 339-347.

Halstead-Reitan battery were based on just 29 people. There were no corrections for age or education level, and various norms for the different tests have been developed since then. Recognizing those problems inherent with multiple sets of norms, in 1991 an effort was made to come up with authoritative or “Comprehensive” norms for the battery.¹⁰

The average age for the individuals used for these Comprehensive norms was 42, the average IQ was 113.8 (“normal” IQ is 100) and the average education was 13.6 years, so the sample used for these norms was relatively older, more intelligent, and more educated than the population at large; no data were provided on ethnicity. There has been considerable controversy about whether the norms are truly representative of the population as a whole.¹¹ The developers of the comprehensive norms used both genders as well as 10 age levels and six educational levels, to come up with 120 “cells” (2 times 10 times 6), but to date have declined to state exactly how many participants there were in each cell; they have responded, however, to some of the criticism.¹² In any case, the data from which these “comprehensive” norms were derived is now about 15 years old. Last year an updated version of the “comprehensive norms” was released. While the sample sizes used in these norms have increased, so has the number of cells, as data are now available by ethnicity, as well as gender, age, and education level.¹³

Professionals who are involved in lawsuits involving expert witnesses who have performed neuropsychological assessment might be well-advised to inquire as to which

¹⁰ Heaton, R.K., Grant, I., & Matthews, C.G. (1991). Comprehensive norms for an expanded Halstead-Reitan Battery. Odessa, Florida, Psychological Assessment Resources, Inc.

¹¹ See, for example, Fastenau, P.S. & Adams, K.M. (1996). Heaton, Grant and Mathews’ comprehensive norms: An overzealous attempt? 18 *Journal of Experimental and Clinical Neuropsychology* 444-448.

¹² See Heaton, R.K., Matthews, C.G., Grant, I. & Avitable, N. (1996). Demographic corrections with comprehensive norms; an overzealous attempt or a good start? 18 *Journal of Experimental and Clinical Neuropsychology* 449-458.

¹³ Heaton, R.K.; Miller, S.W.; Taylor, M.J. & Grant, I. (2004). Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults, Odessa, Florida, Psychological Assessment Resources, Inc.

norms were used in scoring each of the neuropsychological tests involved, and as to how the individual's score would rank if other competing published norms for each test had been used instead.¹⁴

Validation of Neuropsychological Tests

The usual experimental method for validating neuropsychological tests is to compare the performance on the test of a group of known brain-damaged individuals with the performance on that test of a group of “normals”. The test is considered to be “valid” in discriminating brain-damaged individuals from normals if a “cut-off” score can be developed such that most of the individuals whose score falls on one side of that score (either above or below the cut-off) turn out to be from the brain-damaged group, and such that most of the individuals whose scores fall on the other side of the “cut-off” score turn out to be from the normal group. For example, suppose that the validation study was composed of 50 brain-damaged individuals and 50 normals, and suppose further that on a hypothetical test, 90 percent (or 45) of the brain-damaged individuals got more than 12 of the 20 items on the test wrong, whereas 90 percent (or 45) of the normals got 12 or less of the 20 items wrong. Thus, using the “cut-off” score of 12, the test's overall “accuracy rate” would be 90 percent, as 90 percent (or 90 out of 100) of the individuals would be correctly classified as either being brain-damaged or normal using the test. The test would incorrectly classify five brain-damaged individuals as being normal, and would incorrectly classify five normal individuals as being brain-damaged. These results could be tabulated in what psychologists call a two-by-two contingency table.

¹⁴ Various norms for most neuropsychological tests are listed in the *Handbook of Normative Data for Neuropsychological Assessment* by Mitrushina, et al., published in 1999 by Oxford University Press.

		Indication of the Test + (Brain-damaged) – (normals)	
True Condition (actual brain damaged)	“Hit”, or true positives “Sensitivity” 45	“Miss”, or false negatives “Type II Error” 5	
No Condition (actual normals)	“False Alarm”, or false positives; “Type I Error” 5	“Correct Reject”, or true negatives; “Specificity” 45	

Thus in this example, there would be 45 “true positives”, or people who were actually brain-damaged who the test identified as brain-damaged, and what psychologists call the “sensitivity”, or the percentage of time the test identifies brain damage, (sensitivity = true positives divided by true positives + false negatives) of the test would be .90 (45 divided by 45 + 5), or 90 percent. There would be 45 “true negatives”, or people who were actually normal and the test identified them as normal, and what psychologists call the “specificity”, or the percentage of time the test identifies the absence of brain damage, (specificity = true negatives divided by false positives and true negatives) of the tests would also be .90 (45 divided by 45 + 5), or 90 percent. There would be five “false negatives”, or people who were actually brain damaged but the test said they were normals; thus the rate of what psychologists call “Type II errors” would be 10 percent (Type II error rate = false negatives divided by true negatives). Finally, there would be five “false positives”, or people who were actually normal, but the test said they were brain damaged; thus the rate of what psychologists call “Type I errors” would also be 10 percent (Type I error rate = false positives divided by true positives).

Positive Predictive Power and Negative Predictive Power

Notice that these formulas can only be applied if the frequency count for each of the four cells is available, as is the case after doing a validation experiment. However, a forensic psychologist attempting to assess someone in a forensic psychology context to determine if that individual is brain-damaged does **not** have all of the information available to him that he needs to fill in the four cells. What he does not know is the base-rate of brain damage among the people he sees in his practice. If he has validation research on the test available to him (as he normally would) he knows how accurately the test identified brain damage under the experimental situation's base rate, which is usually 50 percent (or close to it) as in our example, as 50 percent of the people used in the validation study were brain-damaged. From the validation research he knows the "sensitivity" of the test (knowing that the person has brain damage, the likelihood the test will identify it) and the specificity of the test (knowing that the person is not brain damaged, the likelihood the test will show no brain damage). However, in the forensic context the psychologist needs to answer two other questions. The first question is: knowing that the test shows brain damage, what is the likelihood that the person actually is brain-damaged, or what psychologists call positive predictive power (PPP); and the second question is: knowing that the test does not show brain damage, what is the likelihood that the person is actually not brain-damaged, or what psychologists call negative predictive power (NPP).

In order to calculate PPP and NPP the psychologist must know the base-rate of brain damage among people that he evaluates, and that base-rate may not be close to the 50-50 situation typically used for validation research. If, as in our example, the sensitivity

of the test is .90 and the specificity of the test is .90, with a 50-50 base rate of brain damage among these the psychologist evaluates, PPP will be .90, and NPP will also be .90. However, if in his practice the psychologist sees only 10 brain-damaged people out of every 100 he evaluates (low base-rate condition), PPP will only be .50, so if the psychologist says the test shows the individual is brain damaged he will only be right half the time. In this low base rate condition NPP will be .988, or very good, so if the psychologist says the test shows the individual is not brain damaged, he will be right over 98 percent of the time. Conversely, if in his practice the psychologist sees 90 truly brain-damaged people out of every 100 he evaluates (high base-rate condition), PPP will be .988, or very good, so if the psychologist says the test shows the individual is brain-damaged he will be right over 98 percent of the time, but NPP will be only .50, so if the psychologist says the individual is not brain damaged, he will be right only 50 percent of the time.

When testifying in court psychologists often seek to avoid discussing these issues. Since most validation research is based on differences between the averages of the scores on the test between the brain-damaged and normal groups (“means-testing”), the psychologist will usually refer to the validation studies by saying something like, “The mean score for the brain-damaged sample was significantly lower than the mean score for the normal sample”. If the validation research was done in terms of correlations, the psychologist will usually say something like, “A low score on this test was significantly correlated with brain damage”. However, means testing only tells you whether one group scores higher or lower than the other, and correlations can only tell you whether there is

an association, but neither of these techniques tells you how likely you are to be right if you use the test.

As we have seen, in order for the psychologist to know how likely he is to be right if he uses the test, he needs to know what percentage of the people he sees are actually brain-damaged. If he works in a rehabilitation hospital for brain-damaged individuals he will be in a high base rate situation, so as we have seen, his positive predictive power (knowing the test shows brain damage, the chances the person is actually brain damaged) will be very high, but his negative predictive power (knowing that the test does not show brain damage, the chances the person is actually not brain-damaged) will be poor. If the psychologist is working in a private practice clinic with relatively few brain-damaged individuals, he is in a low base-rate situation, so as we have discussed, his positive predictive power will be poor, but his negative predictive power will be very high. We can see intuitively how this works—in a rehabilitation hospital setting the psychologist will have the highest overall accuracy rate if he says everyone he sees is brain-damaged (regardless of their test scores), and in a private practice setting the psychologist will have the highest overall accuracy rate if he says everyone he sees is not brain-damaged, regardless of their test scores.

What all of this means for a professional who is involved in a legal case with a neuropsychologist expert witness, is that the attorney should ask the neuropsychologist to calculate PPP and NPP. In order to do so the neuropsychologist will have to come up with an estimate of the base-rate of actual brain damage among those he evaluates. Many neuropsychologists will not be able to produce any data on which to justify such an

estimate, and will simply have to “make-up” an estimate of the base-rate in their practice without much real justification.

Detecting Obvious Cases of Brain Damage vs. Detecting Subtle Cases

As mentioned above, neuropsychological tests are “validated” through research in which the test scores of people who are known to be brain damaged are compared to the scores of people who are thought to be normal. Tests that can do a fairly accurate job of discriminating between the normals and the brain-damaged groups are considered to be valid. The brain-damaged individuals used for these validation studies almost always have obvious brain damage, which is evident on brain imaging techniques such as CT-scans, and which is obvious upon even a cursory conversation with them, as they show obvious symptoms such as aphasia (serious speech problems), dementia (gross memory problems), and/or disabling motor problems. Usually these brain-damaged individuals have already been diagnosed as having a serious brain damage causing illness such as a stroke or Alzheimer’s, or a serious head injury from an accident. Indeed, typically these brain-damaged individuals are so severely disabled that there is no question about their brain dysfunction and because of the obvious nature of their symptoms they have no clinical need to be seen by neuropsychologists; they take the neuropsychological tests only for research purposes.

It can be argued that to a considerable extent this type of research has little applicability to legal contexts. When brain damage can be identified definitively through other means, it is not necessary to enlist the services of a neuropsychologist to indicate what is already known. Most cases in which forensic neuropsychologists become

involved are subtle cases, in which it can be reasonably disputed whether brain damage exists. Various studies show lower or poor accuracy when neuropsychological testing is applied to less obvious or gross cases.^{15 16} Researchers can obtain artificially inflated “hit rates” through the selection of only very grossly brain-damaged individuals for use in validation research, but these rates may not be representative of those achieved in typical clinical practice.¹⁷

In such cases the “cut-off” scores for brain damage may not be appropriate.¹⁸ Furthermore, the “normals” used in the validation studies are often high-functioning persons, such as college students, but in a typical litigation case the neuropsychologist is attempting to distinguish whether the individual has subtle brain damage, or whether he has psychiatric problems, because to be referred, someone has suspected a problem exists based on some reported symptoms. Thus, a control group made up of referred subjects would appear to be a more appropriate control group.¹⁹ The point is that in practice, neuropsychologists are not asked to discriminate between those with obvious brain damage and those that nobody thinks has any problem. Instead, they are usually asked to differentiate between those with subtle brain damage and those who are suspected of having brain damage but in fact do not. Neuropsychologists are often involved in cases with no hard evidence of brain injury, such as mild head injury cases or many cases involving exposure to toxins, cases in which the validation research for the tests

¹⁵ Sweet, J.J., et al. (1985). Comparisons of the decision-tree and standard methods of the Luria-Nebraska Neuropsychological Battery. *Journal of Consulting and Clinical Psychology*, 53, 185-188.

¹⁶ Klesges, R.C., et al. (1984). A major validation study of the Halstead-Reitan in the prediction of CAT-scan assessed brain damage in adults. *The International Journal of Clinical Neuropsychology*, 6, 19-34.

¹⁷ Lezak, M.D. (1983). *Neuropsychological assessment* (2nd ed.). Oxford University Press, New York, p. 153.

¹⁸ Golden, C.J. (1986). Forensic neuropsychology: Introduction and overview in C.J. Golden and M.A. Snider (eds.). *Forensic Neuropsychology* (1-47), Plenum Press, New York.

¹⁹ See Klesges, et al. (1984) mentioned above.

employed may not be applicable. All of this points to the fact that the term “brain damage” is such a broad term that it has little specific significance, as there are many types and degrees of brain damage.

Flexible vs. Inflexible Test Batteries

Many neuropsychologists use the same battery of tests each time they do a brain-damage assessment, and that was the original concept behind the commonly used Halstead-Reitan Neuropsychological Battery for adults, which consists of a core of five tests (Category Test, Tactual Performance Test, Seashore Rhythm Test, Speech-Sounds Perception Test, Finger-Tapping Test) from which seven scores are derived that are used in determining the Impairment Index. To compute the Impairment Index one divides the number of scores that fall outside of a cut-off point by seven (the total number of scores), which results in a score ranging from 0 to 1.0, with 0 signifying that no scores were in the impaired range, and 1.0 signifying that all the scores were in the impaired range. Reitan indicates that scores of .5 or above (half or more of the scores in the impaired range) should be classified as indicative of brain damage.²⁰

Thus it is readily apparent that individuals can perform in the impaired range on up to three of the seven tests, and still not be classified as brain-damaged. However, many experts will point to one or two deviant scores, and claim that they establish brain damage. Indeed, it is not uncommon for an expert to give up to 20 tests, and then point to a low score on one test as indicative of brain damage. It should be obvious that even among normal individuals who are not brain damaged, if enough tests are administered,

²⁰ Reitan, R.M. and Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Battery: Theory and clinical interpretation* (2nd ed.). Neuropsychology Press, Tucson, Arizona.

eventually a low score will probably be obtained. Indeed, it is the unusual result that individuals complete a neuropsychological battery without receiving some low or “abnormal” scores.²¹

In addition to the five core tests, additional tests are administered as well for a complete Halstead-Reitan Neuropsychological Test Battery, including the Trail Making Test, Strength of Grip, and the Aphasia Screening Tests. An MMPI-2 and a full-scale intelligence test are usually also administered. However, some neuropsychologists may not administer the entire battery, using what is referred to as a “partial” battery. In these cases research based on the complete battery must be of questionable validity, and it may not be possible to compute the Impairment Index.

The Halstead-Reitan Neuropsychological Test Battery is a “fixed battery” approach, since most users always administer all of the tests in every brain-damage assessment. The other commonly used fixed battery approach is the Luria-Nebraska Neuropsychological Battery, a less lengthy procedure than the Halstead-Reitan Neuropsychological Battery, which seems to be falling into disfavor. In contrast to the fixed battery approach, probably a majority of neuropsychologists use a “flexible” approach in which they vary the combination of tests they use depending upon the referral issues.^{22 23}

The flexible approach presents some problems in the forensic context. Practitioners may invent their own idiosyncratic batteries nearly each time a new person

²¹ See Heaton, R.K.; Grant, I.; and Mathews, C.G. (1991). Comprehensive norms for an expanded Halstead-Reitan Battery. Psychological Assessment Resources, Odessa, Florida.

²² See Lezak, M.D. (1995). Neuropsychological Assessment (3rd ed.). Oxford University Press, New York.

²³ Lezak, M.D. (1984). An individualized approach to neuropsychological assessment. In P.E. Logue and J.M. Schear (Eds.). Clinical neuropsychology: A multidisciplinary approach (pp. 29-49). Charles C. Thomas, Springfield, Illinois.

is evaluated, meaning that those exact procedures have usually never been used by that examiner before. Research consistently shows that standardized or pre-specified procedures lead to higher overall rates of accuracy than more variable or non-specific procedures.²⁴

Some courts will prefer inflexible, or “fixed” batteries, over flexible batteries. This may be especially true in Federal courts and in other jurisdictions that use the *Daubert* standard,²⁵ which requires that the methodology used be testable and that there must be a measurable error rate. For example, in *Chapple v. Ganger* a federal case in the State of Washington,²⁶ two evaluators employed a flexible battery approach, and one employed a fixed battery, and the Court preferred the fixed battery approach.²⁷

Neuropsychologists who rely upon a flexible battery approach must select tests at least partly based on the examinee’s stated complaints or symptoms, so the value of the entire neuropsychological evaluation may depend on the validity of those complaints or symptoms, and may miss the mark in many cases.²⁸ The flexible battery approach necessarily requires the examiner to focus on the reported symptoms of the person being evaluated as a place to start in deciding whether a neuropsychological evaluation should be undertaken, and if so, what tests or procedures to employ. The assumption often seems to be that if an individual reports common symptoms of brain dysfunction, such as memory and attention problems, headaches, word-finding difficulties, and fatigue, then

²⁴ See Faust, D. (1995). Neuropsychological (brain damage) assessment, in Jay Ziskin (Ed.). *Coping with psychiatric and psychological testimony*, (5th Ed.). Law and Psychology Press, Los Angeles, pp. 999-1004.

²⁵ *Daubert v. Merrell Dow Pharmaceuticals*, 113 S. Ct. 2786 (1993).

²⁶ 851 F. Supp. 1481 (E.D. Wa 1994).

²⁷ See Reed, J.E. (1996). Fixed vs. flexible neuropsychological test batteries under the *Daubert* standard for the admissibility of scientific evidence. *Behavioral Sciences and the Law*, 14, 313-322.

²⁸ Chelune, G.J.; Heaton, R.K.; and Lehman, R.A.W. (1986). Neuropsychological and personality correlates of patients’ complaints of disability. In G. Goldstein and R.E. Tauter (Eds.). *Advances in clinical neuropsychology* (Vol 3, pp 95-126). Plenum Press, New York.

he must be suffering from brain damage or brain dysfunction. However, research consistently shows that such symptoms are very common among “normal” persons who have had no head injury at all.²⁹ Such symptoms are also frequently present among patients receiving psychotherapy, so the presence of these symptoms does not necessarily point to brain dysfunction.³⁰

Furthermore, persons who have experienced a head injury which possibly caused brain dysfunction tend to underestimate pre-injury cognitive difficulties, making it difficult to determine whether the reported problems existed before the head injury occurred.³¹ Research shows that self-reports of symptoms are often inaccurate, at least when compared with performances on neuropsychological tests. On one hand, brain-damaged individuals sometimes fail to recognize the presence of deficits.³² On the other hand, brain-damaged individuals can also under-report their symptomatology.³³ All of this research suggests that the type of reported symptoms may not be a useful starting place in investigation of brain dysfunction.

²⁹ Newman, B. (1992). Research in mild traumatic brain injury. New York University Medical Center Research Training Center News, 4, 3; Gouvier, W.D., et al. (1988). Base rates for post-concussional symptoms. *Archives of Clinical Neuropsychology*, 3, 273-278; Gouvier, W.D., et al. (1992). Post-concussion symptoms and daily stress in normal and head-injured college students. *Archives of Clinical Neuropsychology*, 7, 193-211; Wong, J.L., et al. (1994). Base rate and simulated symptoms of mild head injury among normals. *Archives of Clinical Neuropsychology*, 9, 411-425.

³⁰ Fox, D.D., et al. (1995). Base rates of post-concussive symptoms in health maintenance organization patients and controls. *Neuropsychology*, 9, 606-611.

³¹ Mittenberg, W., et al. (1992). Symptoms following mild head injury: Expectation as a etiology. *Journal of Neurology, Neurosurgery and Neuropsychiatry*, 35, 200-204.

³² McGlynn, S.M. and Schacter, D.L. (1989). Unawareness of deficits in neuropsychological syndromes. *Journal of Clinical and Experimental Nueropsychology*, 11, 143-205.

³³ Gass, G.S., et al. (1990). Accuracy of MMPI-based inferences regarding memory and concentration in closed-head trauma patients. *Psychological Assessment: A Journal of consulting and Clinical Psychology*, 12, 175-178.

Extraneous Factors that can Decrease Scores on Neuropsychological Tests

Neuropsychologists often assume that an individual's poor performance on neuropsychological tests automatically means that he is suffering from brain damage. However, there are a number of other factors known to decrease test performance. A thorough neuropsychologist will assess to what extent these factors reduced test performance, but some experts make little attempt to determine what other factors might have lowered test results.

Considerable research shows that emotional problems can reduce performance on neuropsychological tests.³⁴ Depression has been shown to significantly reduce scores on IQ tests.³⁵ ³⁶ Anxiety also interferes with performance on neuropsychological tests.³⁷ During testing many people become anxious when they believe they are performing poorly, or when they expect they will do poorly, and that anxiety can seriously impair test performance. Personality disorders can also impair test performance.³⁸ Anxiety due to post-traumatic stress disorder (PTSD) will also often cause cognitive deficits.³⁹

Past and current alcohol and/or drug abuse frequently depress scores on neuropsychological tests. Once an individual quits using alcohol or drugs, performance

³⁴ See e.g. King, D.A., Caine, E.D., and Cot, C. (1993). Influence of depression and age on selected cognitive functions. *The Clinical Neuropsychologist*, 7, 443-453.

³⁵ Dean, R.S., Gray, J.W., and Seretny, M.L. (1987). Cognitive aspects of schizophrenia and primary affective depression. *The International Journal of Clinical Neuropsychology*, 9, 33-36.

³⁶ Sweet, J.J., Newman, P., Bell, B. (1992). Significance of depression in clinical neuropsychological assessment. *Clinical Psychology Review*, 12, 21-45.

³⁷ Buchelew, S.P. and Hannay, H.J. (1986). Relationships among anxiety, defensiveness, sex, task difficulty, and performance on various neuropsychological tasks. *Perceptual and Motor Skills*, 63, 711-718.

³⁸ Cox, C.D. (1997). Neuropsychological abnormalities in obsessive compulsive disorder and their assessments. *International Review of Psychiatry*, 9-45-60; Swirsky-Sachetti, T., et al. (1993). Neuropsychological function in borderline personality disorder. *Journal of Clinical Psychology*, 36, 3-20.

³⁹ Bremner, J.D., et al. (1995). MRI-based measurement of hippocampal volume in patients with combat-related post-traumatic stress disorder. *American Journal of Psychiatry*, 150, 1015-1019.

may gradually improve, but it may take a considerable time for performance to return to normal.⁴⁰ Various medications can also compromise cognitive functioning, causing incoherent speech, memory loss, and confusion.⁴¹

If the person being tested has slept poorly or is fatigued, performance on neuropsychological tests will also be decreased. Many neuropsychologists schedule testing sessions lasting eight hours or more, leading to fatigue and reduced motivation.⁴²

Performance on neuropsychological testing can be significantly impaired by poor vision and/or poor hearing.^{43 44} Chronic and severe pain can also interfere with an individual's performance, causing distraction and inability to concentrate.⁴⁵

All of these factors that potentially decrease performance point to the necessity of assessing for psychological problems, medications, drug/alcohol abuse/dependence, vision, hearing, pain, and other factors that might contribute to impaired functioning.

Post Traumatic Stress Disorder may Look Like Brain Damage

Persons who suffer from possible brain damage from trauma may show symptoms of post-traumatic stress, and persons with PTSD show cognitive deficits, with memory and attention problems, and disorganization. Significant depression and somatization are also common symptoms. These symptoms may masquerade as brain damage.

⁴⁰ Grant, I. (1987). Alcohol and the brain: Neuropsychological correlates. *Journal of Consulting and Clinical Psychology*, 35, 310-324.

⁴¹ Meyer, J.D. and Fink, C.M. (1989). Psychiatric symptoms from prescription medications. *Professional Psychology, Research and Practice*, 20, 90-96.

⁴² Benton, A. (1992). Clinical Neuropsychology: 1960-1990. *Journal of Clinical and Experimental Neuropsychology*, 14, 407-417.

⁴³ Kempen, J.H., et al. (1994). Effect of visual impairment on neuropsychological test performance. *Journal of Clinical and Experimental Neuropsychology*, 16, 223-231.

⁴⁴ Schear, J.M., et al. (1988). Effect of simulated hearing loss on speech sounds perception. *Journal of Clinical and Experimental Neuropsychology*, 10, 597-602.

⁴⁵ Taylor, A.E., et al. (1996). Persistent neuropsychological deficits following whiplash: Evidence for chronic mild traumatic brain injury? *Archives of Physical Medicine and Rehabilitation*, 77, 529-535.

Furthermore, even when brain damage has occurred, PTSD may be superimposed, resulting in additional disability. Trauma may also trigger the re-emergence of old emotions if the individual suffered other traumatic events in his past.⁴⁶

Testing of Older Individuals

As might be suspected, age is related to performance on many neuropsychological tests, with older persons (even “normal” ones) usually performing more poorly than younger individuals. Thus, failure to use age-corrected norms for evaluating scores can easily mean the difference between a correct and incorrect determination of brain damage. However, much of the research in neuropsychology has involved middle-aged or younger populations, and less is known about the test-scores of older people. For many neuropsychological tests there may be little or no normative data for older populations. Unless age-corrected norms are used, older persons will frequently be misclassified as brain-damaged.^{47 48} Indeed, there are very few neuropsychological tests that have established validity for older populations.⁴⁹ Furthermore, if age-corrected norms for older individuals are available for a given test, neuropsychologists should not use “standard” rather than age-corrected norms for those tests, and if they use standard norms their conclusions should be called into question.

⁴⁶ See Kay, T. (1999). Interpreting apparent neuropsychological deficits, in Sweet, J.J. (Ed.). *Forensic neuropsychology: Fundamentals and practice*. Swets & Zeitlinger, Etton, PA, pp 170-173.

⁴⁷ Long, C.J. and Klein, K. (1990). Decision strategies in neuropsychology II: Determination of age effects on neuropsychological performance. *Archives of Clinical Neuropsychology*, 5, 335-345.

⁴⁸ Bornstein, R.A.; Paniak, C.; and O'Brien, W.O. (1987). Preliminary data on classification of normal and brain-damaged elderly subjects. *The Clinical Neuropsychologist*, 1, 315-323.

⁴⁹ Loewenstein, D.A., et al. (1994). Potential cultural bias in the neuropsychological assessment of the older adult. *Journal of Clinical and Experimental Neuropsychology*, 16, 623-629.

Testing of those with Limited Education and/or Low IQ

Research also shows that less-educated and less-intelligent examinees receive substantially lower scores on many if not most neuropsychological tests. For example, using those tests comprising the Halstead-Reitan Battery, the Halstead-Reitan norms will often result in falsely classifying individuals (especially older individuals) simply because of their level of education and their IQ.⁵⁰

Testing of Minority Group Members

There is considerable literature pertaining to ethnic differences on various tests of cognitive abilities, with various authors pointing out problems in relying upon standard norms for neuropsychological tests when testing minority group members.⁵¹ The majority of the Halstead-Reitan tests show differences between various ethnic group members.⁵² Commentators urge neuropsychologists to carefully evaluate their tests for cultural fairness and the adequacy of their normative base when evaluating Hispanics, African-Americans, Native Americans and Asians, and suggest the development of more adequate normative data.⁵³

Illustrative of these problems, data from the standardization sample for a standard intelligence test, the WAIS-R, shows that whites outperformed blacks by about one to

⁵⁰ Leckliter, I.N. and Matarazzo, J.D. (1989). The influence of age, education, IQ, gender and alcohol abuse on the Halstead-Reitan Neuropsychological Test Battery performance. *Journal of Clinical Psychology*, 45, 485-512.

⁵¹ Adams, K.M. and Heaton, R.K. (1990). Statement concerning the MMH Neuropsychological Battery. *Journal of Clinical and Experimental Neuropsychology*, 12, 960-962.

⁵² Fletcher, J.M. (1989). Child neuropsychology. *The Clinical Neuropsychologist*, 3, 85-92.

⁵³ See, for example, Loewenstein, D.A. et al. (1994). Potential cultural bias in the neuropsychological assessment of the older adult. *Journal of Clinical and Experimental Neuropsychology*, 16, 623-629, and Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, 17, 143-150.

three points on all of the test's sub-tests, and these differences add up, with the cumulative effect being that on average blacks obtain substantially lower overall IQ scores on the test than do whites; on average scoring about 14 to 15 IQ points below the average score for whites.⁵⁴

Evaluating those from other Cultures Whose Native Language is not English

Most psychological tests have not been developed for use with non-English speaking populations or for people who did not grow up in the United States, but many of the people neuropsychologists are asked to evaluate are immigrants who speak limited English or speak English as a second language. Neuropsychologists asked to evaluate such persons, as for example, Spanish-speaking immigrants with limited education in their native country, can simply decline to do so, stating that their tests are inappropriate for such persons. Another approach, assuming that the psychologist or his psychometrician speaks Spanish, would be to administer the usual English-language neuropsychological tests, translating them into Spanish, and to score them using the available U.S. norms. However, it should be obvious that such persons will likely receive fairly low scores when compared to people who grew up in the United States and had more education, even if they are not actually brain-damaged; furthermore, it will not be clear whether the item difficulty level will be the same in the translation. A third approach would be to try to find neuropsychological tests developed and scored in Spanish, but there may still be problems with trying to equate the level of difficulty

⁵⁴ Kaufman, A.S.; McBean, J.E.; and Reynolds, C.R. (1988). Sex, race, residence, region, and educational differences on the II WAIS-R subtests. *Journal of Clinical Psychology*, 44, 231-248.

between the items.^{55 56 57} A fourth approach would be to use non-verbal tests, such as the Test of Non-Verbal Intelligence Third Edition (TONI-3), which are supposedly more “culture free”, but there are very few such tests available.

All of this suggests that in any case in which a neuropsychologist has evaluated an individual whose native language is not English, attorneys involved should carefully scrutinize any testing utilized, and any conclusions drawn from that testing.

Assessment of Pre-morbid Functioning

In purely clinical cases, neuropsychologists may simply focus on what impairments the individual has without concerning themselves with the reasons for the reduced performance. However, in most cases involving litigation, the neuropsychologists must attempt to identify the causes of the impaired functioning. These types of cases include personal injury cases, product liability cases, medical malpractice suits, and workers compensation cases. In such cases it is not sufficient to identify what the subject’s present impairments are, because they could have been present before the alleged injury; the subject could even have been born with congenital brain damage. Thus, in most forensic cases the neuropsychologist must attempt to determine what the individual’s functioning was before the alleged accident or injury, which is referred to as his “pre-morbid functioning”.

⁵⁵ See, for example, Allegri, R.F., Mangone, C.A. Villavicencio, A.F., Rymberg, S., Taragana, F.E., & Baumann, D. (1997). Spanish Boston Naming Test Norms, 11 *The Clinical Neuropsychologist*, 416-420.

⁵⁶ See also Demsky, Y., Gass, C., Edwards, W.T., & Golden, C.J. (1998). Optimal short-forms of the Spanish WAIS (ELWA), 4 *Assessment* 361-364.

⁵⁷ Scott, R.L. & Pampa, W.M. (2000). The MMPI-2 in Peru: A normative study. 74 *Journal of Personality Assessment* 95-105. The Spanish translation of the MMPI-2 has been found to be useful outside the U.S.

This usually involves extensive gathering of historical information about the person's previous functioning, such as school records, military records, hospital and doctors' reports, employment records, and any records of prior psychological testing, which are used to attempt to determine if the subject's functioning before the injury or accident was better than it is after the injury or accident. The neuropsychologist also needs to determine whether there were any other accidents or injuries which could have caused brain damage.

This investigation of other causes of the brain damage, and of the person's pre-morbid functioning, is one of the most difficult issues in forensic neuropsychology. Frequently there is no psychological testing available to show how the subject functioned before the injury or accident, so "pre-morbid" functioning has to be inferred from school performance and employment performance. Even those sources of historical functioning may not be complete or readily available, and neuropsychologists have to rely on what the person tells them about their past functioning, and/or on what significant others, such as their parents, their spouses, their children, etc. tell them about their prior functioning. Clinical interviewing is the most routinely used method of establishing pre-morbid status.⁵⁸

A neuropsychologist who does not obtain a careful history and who does not obtain and review available school records, employment records and medical reports should be vigorously criticized. Some neuropsychologists simply look at the results of testing and claim to be able to determine that the person is brain-damaged due to a specific accident or injury, but such experts have no basis for that opinion, because the

⁵⁸ Smith-Seemiller, L., et al. (1997) Neuropsychologists' practice patterns in assessing pre-morbid intelligence. *Archives of Clinical Neuropsychology*, 12, 739-744.

brain-damage could have been present throughout the person's life, and also could have been caused by other injuries, accidents, or disease processes.

It has been established that performance on some types of tests does not deteriorate much with brain damage, or is "resistant" to brain damage, whereas performance on other types of tests is more "sensitive" to brain damage. For example, a brain-damaged person will still be able to pronounce irregularly spelled words about as well as he could before his brain-injury, but will do much more poorly on memory tasks than he could before his brain injury. Based on the possibly faulty assumption that all of the individual's pre-morbid abilities were originally about at the same level, some neuropsychologists use actuarial formulas to mathematically estimate the level of pre-morbid functioning using scores on the more "resistant" tests. For example, scores on the Wechsler Test of Adult Reading (WTAR),⁵⁹ or the North American Adult Reading Test (NAART) are sometimes used to predict pre-morbid functioning.⁶⁰ Another actuarial approach to estimating pre-morbid functioning uses socio-demographic variables, such as age, education, gender, race, geographic region, and occupation, to estimate what the level of pre-morbid functioning would have been. One such method is the Barona Regression Equation (BRE).⁶¹ Some neuropsychologists use a combination of these two types of approaches, such as the Oklahoma Pre-morbid Intelligence Estimate (OPIE).⁶²

⁵⁹ The Psychological Corporation (2001). Manual for the Wechsler Test of Adult Reading. San Antonio, Texas.

⁶⁰ Blair, J.R. and Spreen, O. (1989). Predicting pre-morbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist*, 3, 129-136.

⁶¹ Barona, J.P., et al. (1984). A demographically-based index of pre-morbid intelligence for the WAIS-R. *Journal of Clinical Psychology*, 52, 885-887.

⁶² Scott, J.G., et al. (1997). Oklahoma Pre-morbid Intelligence Estimate (OPIE): Utilization of a clinical sample. *The Clinical Neuropsychologist*, 11, 146-154.

All of these actuarial approaches have been criticized, but they may still be superior to subjective clinical estimates of pre-morbid functioning.^{63 64}

Ecological Validity

Even if the neuropsychologist is able to establish that the individual is brain damaged, and is able to determine that the brain damage is due to a specific injury or accident, it will still be necessary to determine how much this brain damage actually interferes with the person's everyday functioning, especially in forensic cases in which the amount of damages is an issue. Some individuals who truly are brain damaged may show little or no functional impairment. Even if neuropsychological testing can validly measure memory functions, for example, how well does that testing relate to everyday memory issues such as remembering a work task or remembering to take medications? The issue of the relevance of test scores to real-world behaviors and performances has been dubbed "ecological validity".⁶⁵ Different people have different types of challenges in their everyday lives, so the neuropsychologist may have to find out what the demand characteristics are of the person's home and work environment, such as what they are required to do at work, whether they must drive, and how much they have to remember. Most commentators believe there is little meaningful research addressing the relationship between neuropsychological test and everyday functioning.^{66 67 68} However, there is some

⁶³ Wiens, A.N., Bryan, J.E., and Crossen, J.R. (1993). Estimating WAIS-R FSIQ from the National Adult Reading Test-Revised in normal subjects. *The Clinical Neuropsychologist*, 7, 70-84.

⁶⁴ Silverstein, A.B. (1987). Accuracy of estimates of pre-morbid intelligence based on demographic variables. *Journal of Clinical Psychology*, 43, 409-414.

⁶⁵ Wilson, B.A. (1993). Ecological validity of neuropsychological assessment: Do neuropsychological indexes predict performance in everyday activities? *Applied and Preventative Psychology*, 2, 209-215.

⁶⁶ Dodrill, C.B., and Clemmons, D. (1984). Use of neuropsychological tests to identify high school students with epilepsy who later demonstrate inadequate performances in life. *Journal of Consulting and Clinical Psychology*, 52, 520-527.

research relating neuropsychological test performance to the likelihood of returning to work after an injury.⁶⁹

At the very least, conclusions based on neuropsychological testing should be consistent with observations of the patient's daily behavior. For example, it is highly unlikely that a person who scores in the severely deficient range on tests of memory and executive functioning would be able to perform daily living tasks such as driving (particularly to previously-unfamiliar locations, such as the neuropsychologist's office or the deposition site) or money handling.

Malingering

Malingering is not an inconsequential problem in neuropsychological assessment, especially in forensic cases. In an article reviewing the available literature as to feigned neuropsychological impairment, practitioners subjectively estimated that as many as half of those being evaluated may be faking all or part of their cognitive disabilities.⁷⁰

Various methods have been devised to attempt to detect malingering. Malingerers often score even worse than truly disabled people on many instruments, since malingerers do not know what items even very disturbed or disabled persons can complete correctly, and thus score below the "floor" scores of disturbed or disabled persons. This strategy for detecting malingering is called the "Floor Effect", and it is the principle behind the Rey

⁶⁷ Dunn, E.J. et al. (1990). The relation of the Halstead-Reitan Neuropsychological Battery to functional daily living skills in geriatric patients. *Archives of Clinical Neuropsychology*, 5, 103-117.

⁶⁸ Puente, A.E. (1992). The status of clinical neuropsychology. *Archives of Clinical Neuropsychology*, 7, 297-312.

⁶⁹ Kalechstein, A.D.; Newton, T.F.; vanGrop, W.G. (2003). Neurocognitive functioning is associated with employment status. *Journal of Clinical and Experimental Neuropsychology*, 8, 1186-1191.

⁷⁰ Rogers, R.; Harell, E.H.; and Liff, C.D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, 13, 255-274.

Visual Memory Test (a.k.a. Rey Fifteen Item Test), the Rey Word Recognition Test, and the Rey Dot Counting Test, popularized by Lezak.⁷¹ Many neuropsychological tests can be used to detect malingering in this fashion, including tests in the Halstead-Reitan Neuropsychological Battery.⁷²

Another related strategy, called Symptom Validity Testing, relies on the principles of probability to determine which scores are so poor that they are below what would be expected by chance alone. A forced choice procedure is used, usually with only two choices, in which the subject is told to answer all questions, guessing on those on which he is not sure of the answers. For example, on a 100-item, two-choice, forced-choice test, an average person who was guessing without even reading the questions would get about half of the items right. Those scoring substantially lower are showing that on at least some of the items they know the right answer but deliberately picked the wrong one. Examples of such tests are the Test of Memory Malingering (TOMM) and the Victoria Symptom Validity Test (VSVT).

Another approach involves ordering the items on a test from the very easy to the very difficult, but then mixing those items up when administering them on a two-choice test. A person who is not malingering should get most or all of the easier items correct, but as the items become harder his percentage correct will gradually decrease to approximately 50 percent correct, or a chance level of responding. The Validity Indicator Profile (VIP) uses this strategy, as well as floor effect and symptom validity testing.⁷³

⁷¹ Lezak, M. (1983). *Neuropsychological Assessment*. Oxford University Press, New York.

⁷² See Sweet, J.J. (1999). *Malingering: Differential Diagnosis*, in Sweet, J.J., Ed. (1999). *Forensic Neuropsychology: Fundamentals and practice*. Eyton, PA, Swets & Zeitlinger.

⁷³ Frederick, R. and Crosby, R. (2000). Development and validation of the validity indicator profile. *Law and Human Behavior*, 24: 1.

Overlooking the possibility of malingering is irresponsible.⁷⁴ Neuropsychologists who do not employ specialized tests to determine if malingering is occurring are leaving themselves open to serious criticism, since without the use of such specialized tests neuropsychologists have not been shown to be able to be particularly successful at detecting malingering.⁷⁵ A valuable guide to assessing malingering in brain damage cases is Reynolds, C.R. (1998). *Detection of malingering during head injury litigation*. New York, Plenum.

Conclusions

Professionals involved in cases in which neuropsychologists present evidence should critically evaluate whether the tests used are reliable and valid. They should determine whether the tests were administered and scored correctly by qualified individuals, and whether the norms that were used were appropriate given the subject's age, education, ethnic affiliation, and IQ. It is important to consider whether a standardized battery of tests was employed. An attempt should be made to determine if other factors could have affected test scores, such as fatigue, medications, depression, post-traumatic stress, poor English skills, etc. Professionals need to assess whether the assessment of pre-morbid functioning was done appropriately, and whether the possibility of malingering was adequately considered. Professionals should also consider whether the test scores are consistent with the subject's actual everyday life functioning.

⁷⁴ Binder, L.M. (1990). Malingering following minor head trauma. *The Clinical Neuropsychologist*, 4, 25-36.

⁷⁵ Faust, D.; Hark, K.; Guilmetter, T.J.; and Arkes, H.R. (1988). Neuropsychologists' capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice*, 19, 508-515.