8

Actuarial Assessment of Malingering: Rohling's Interpretive Method

Martin L. Rohling[1]
University of South Alabama

Jennifer Langhinrichsen-Rohling
University of South Alabama

L. Stephen Miller
University of Georgia - Athens

CONSIDERATIONS IN THE DIAGNOSIS OF MALINGERING

In recent years, attorneys, physicians, and psychologists have had to become more familiar with the diagnosis of malingering. Patients who malinger are consciously choosing to feign or exaggerate symptoms, often in order to obtain compensation. These same professionals are often given

---

[1]Please address correspondence to Martin Rohling, Ph.D. Assistant Professor Department of Psychology University of South Alabama 381 Life Science Building Mobile, AL 36688-0002. mrohling@usouthal.edu

the responsibility of determining the degree to which the expressed symptoms represent valid, exaggerated, or feigned deficits. For example, a substantial subgroup of individuals seeking compensation (20% to 25%) may warrant the diagnosis of malingering (Binder, Villanueva, Howieson, & Moore, 1993; Guilmette, Sparadeo, Whelihan, & Buongiorno, 1994). Recently, Binder and Rohling (1996) completed a comprehensive meta-analytic review of the effect of financial compensation on the experience and treatment of head injury. They found that patients who received economic incentives were significantly more likely to persist in their complaints of sequelae, regardless of the severity of their injury. An earlier meta analysis by Rohling, Binder, and Langhinrichsen-Rohling (1995) found comparable results for chronic pain patients. Specifically, compensated chronic pain patients reported more pain than did noncompensated chronic pain patients even when the two groups' severity of injuries were comparible prior to receiving compensation. In a third meta analytic study, Binder, Rohling, and Larrabee (1997) found the effect size for residual cognitive deficits from mild head injury was nearly zero (e.g., few, if any deficits can be expected from this type of trauma). One implication of these authors' body of work might be that, in the absence of

financial incentives, few patients would experience sequella due to mild head injuries.

There are a number of reasons why it is important for professionals to accurately differentiate between malingers and individuals with detectable neurologically based impairments. First, accurate diagnosis is critical because it appears that awarding unnecessary financial compensation can make patients' symptoms worse (e.g., Binder and Rohling, 1996; Rohling et al., 1995). Second, compensating patients who are reporting undetectable impairments likely inflates insurance costs and inequitable distribute of health care dollars. Third, other iatrogenic disorders, such as depression and somatoform disorders may develop as a result of inappropriate distribution of health resources. Fourth, attorneys and neuropsychologists likely lose credibility when they pursue unfounded mild injury lawsuits (e.g., see Faust & Ziskin, 1988).

The focus of this chapter is on the difficulties inherent in diagnosing malingering in the population of individuals who have experienced a mild neurological event and who are seeking compensation for their impairments. We summarize human judgment research as it applies to this complex differential diagnosis and present an actuarial

strategy that will facilitate neuropsychological detection of valid and feigned neurocognitive deficits.

Diagnostic Problems Faced by Neuropsychologists When Feigning Occurs

The assessment of malingering has increasingly become an expected component of neuropsychological assessment (Williams, 1998). As a result, neuropsychologists should be aware of the diagnostic criteria of malingering. *The Diagnostic and Statistical Manual of Mental Disorders* -(4th ed.) [*DSM-IV*]; American Psychiatric Association, 1994) defines three criteria that must be met before a diagnosis of malingering should be applied. Briefly, these criteria require the determination that a patient has feigned or exaggerated symptoms. A patient must have intentionally produced the symptoms. Finally, the patient's motivation for reporting symptoms has to be the acquisition of external incentives. Therefore, when neuropsychologists try to determine if these criteria have been met, several complex decisions must be made. These decisions can be conceptualized as following a two (symptoms are valid vs. feigned) by two (unintentional vs. intentional) by two (no incentives vs. incentives) by two (internal incentives vs. external incentives) matrix that results in sixteen possible outcomes. We offer suggestions that contrast with more traditional

resolutions to each decision (e.g., Lezak, 1995; for a critic of traditional methods see Wedding, 1983). Particular attention is paid throughout this chapter to issues raised by Wedding and Faust (1989) in their research review of neuropsychologists' accuracy of assessment. We show how many of the standard assessment problems that they described are magnified when the differential diagnosis of malingering is involved.

Criterion 1: Determining If Symptoms Are False or Grossly Exaggerated

First, the neuropsychologist attempts to determine whether a symptom is valid, grossly exaggerated, or feigned. Traditionally, this decision has been left to the clinical judgment of the evaluator. The judgment of symptom validity is most complex when the patient's complaint is a subjective experience (e.g., "My personality has changed."). When the neuropsychologist obtains abnormally low test scores, she or he must also judge whether the patient's objective performance is reasonable or exaggerated, given the severity of a lesion. Traditionally, standards of "reasonableness" are determined by the personal or professional judgment of the examiner, rather than by using an actuarial strategy (Wedding, 1983). Actuarial methods are appropriate for a number of acute measures such as a patient's post injury time to follow commands (Dikmen, Machamer, Winn, & Temkin, 1995). Research on human judgment has

shown that trained professionals' clinical judgments are likely to be unreliable (Garb & Schramke, 1996; Oskamp, 1965; for an alternative view see McCaffrey & Lynch, 1992; Trueblood & Binder, 1997). In forensic cases, competent neuropsychologists often disagree about the "reasonableness" of a particular symptom, when the determination is based on their own experience and beliefs.

MEDICAL-LEGAL HINT: Determining whether a complaint is false or exaggerated is best done by using an actuarial strategy. This strategy should be based on published data that have examined patients with similar acute circumstances. Norms gathered from these patients can then be used to predict expected residual deficits. Estimates of residual effects can then be compared to current deficits to determine the degree of incongruity. High levels of incongruity should lead one to consider a diagnosis of malingering.

Normative data have been psychology's "gold standard" as a way of minimizing subjective bias. Epidemiological research data are available to determine the probability of a certain level of deficit being presented post trauma. From normative data, the likelihood or reasonableness of developing any particular symptom post injury can be calculated. Neuropsychologists can then use a standard procedure to determine the

probability that a patient's complaint is valid and the likelihood that any particular symptom resulted from the alleged injury. Norm-based predictions can then be compared to the assessed deficits. Probability can then be assigned to the likelihood that the assessed complaint resulted from the alleged injury. This procedure would result in a uniform standard for determining the likelihood that a patient is feigning or grossly exaggerating a complaint.

The Importance of Neuropsychological Signs as Opposed to Patient-Reported Symptoms. The definition of signs versus symptoms is relevant to the discussion of how neuropsychologists might objectively assess a patient's complaint. A symptom is a subjectively experienced problem that is believed to be causally related to a disorder. For example, a headache is a symptom because it cannot be objectively measured by a physician. We know that a person has a headache because they tell us so. Fever, on the other hand, can be verified by a physician by using an objective method of assessment (e.g., a mercury thermometer). Therefore, fever is considered a sign. Both symptoms and signs can be indications of a specific illness, as a headache and a fever can be a direct result of a cold or the flu. Although physicians may ask patients if they feel warm (i.e., Does the patient exhibit

a symptom?), patients' responses to these questions can be considered independent of whether or not they have a fever. Traditionally, evaluation of feigning has been based on symptom complaints gathered during clinical interview, which are not easily verified.

Although it is more difficult to feign a sign than a symptom, it is easier to feign a neuropsychological sign than a physical sign. For example, a patient may intentionally withhold known answers to a psychometrically sound assessment of memory. When the chance of obtaining exceedingly low scores or an unusual pattern of errors is low, it is likely that factors other than the suspected dysfunction caused the patient's poor performance. Furthermore, a direct examination of test scores and item responses may give the neuropsychologist clues as to whether the patient consciously chose to do poorly on a task or not. It is on this basis that we believe that neuropsychologists have tools available to more objectively conclude whether the patient's test scores and symptomatic complaints are feigned or grossly exaggerated.

Criterion 2: Determining Whether Intention or Awareness Existed

DSM-IV requires that the patient intentionally (i.e., consciously) feign or exaggerate a symptom. How one defines the word "intentional"

ultimately determines if the diagnosis of malingering is applied. A patient's intentions to feign and manipulate are typically not disclosed to the professional, particularly if the patient is attempting to manipulate the contingencies via the evaluation. Therefore, it is common to infer a patient's intentions from his or her behavior. This is the second complex judgment required for the diagnosis.

Neuropsychologists often infer patients' awareness on the basis of their "pattern of responding" to interview questions and test items. Rogers, Harrell, and Liff (1993) suggested methods to assist neuropsychologists with this complex inference process. They proposed that intention to deceive may be assumed if a patient (a) has symptoms with a late onset (b) is resistant to treatment or evaluation (c) has no obvious neurological findings or inconsistent findings on neurological exams (d) presents with bizarre signs or symptoms that are inconsistent with current models of cognitive functioning and/or (e) exhibits discrepancies between what is expected and what is observed. Unfortunately, Rogers et al. did not specify how neuropsychologists should determine when these behaviors have been exhibited by a patient. For example, no procedures were proposed for determining if a patient's results are significantly different from expectation. Furthermore, no

recommendations were provided as to how neuropsychologists are to integrate inconsistent positive and negative findings from these methods. As a result, these excellent suggestions are not often utilized in a reliable manner.

Neuropsychologists' assessment of the intentionality of sign or symptom production has been further complicated by the changing nature of the attorney-client relationship. For example, recent evidence suggests that some attorneys believe it is their professional obligation to educate their clients about the assessment process prior to their being subjected to it (Youngjohn, 1995). Although little is known about the degree to which this educational process alters the validity of the obtained test results some have argued that this process may actually function as a method of "coaching" the client so as to avoid detection of malingering (Youngjohn, 1995).

To resolve the neuropsychologist's dilemma in assessing intention, we make several recommendations. First, neuropsychologists should view a client's intention to deceive on a continuum rather than as a totally present versus totally absent dichotomy. Using this continuum, malingering can be considered as a reasonable diagnosis even if the neuropsychologist has only enough evidence to show that a patient is

beyond the midpoint of this continuum (e.g., it can be inferred that the patient has some level of intention to deceive in order to respond in the manner evident in the test results). Thus, the *DSM-IV* criterion of intentionality could be judged present when only two of Rogers et al.'s (1993) criteria appear to be present. Likewise, intentionality would still be judged as present in situations in which some of the patient's signs and symptoms appear to be valid, if it can be proven that the patient has intentionally feigned or exaggerated other signs and symptoms in order to increase their chances of receiving compensation.

MEDICAL-LEGAL HINT: Neuropsychologists frequently disagree about whether a subjectively assessed sign of feigning exists or not. Their skills at detecting feigning with traditional methods are reported as poor (Faust, Hart, & Guilmette, 1988; Faust, Hart, Guilmette, & Arkes, 1988; Wedding, 1983). Relying on neuropsychological signs rather then symptoms increases the accuracy of diagnosis (Trueblood & Binder, 1997). Furthermore, viewing patient's intention to malinger along a continuum of awareness, rather than as a dichotomy, may help to reduce neuropsychologists' rate of misdiagnosis.

Criterion 3: Determining If Incentives Exist

Because malingering requires that incentives exist, it is not uncommon for patients to avoid disclosing that they are trying to obtain compensation for their alleged injury. If this deception is successful, the

neuropsychologist may be less likely to infer malingering. For example, attorneys may inadvertently facilitate their clients' attempts at deception by having another professional preliminarily evaluate their clients (e.g., neurologist, chiropractor, family practitioner, physical therapist, etc.). This "middle person" then refers the client to a neuropsychologist for an assessment. The neuropsychologist has no direct contact with or knowledge of the attorney. Complex circumstances, including the use of professionals from various disciplines, pose several challenges to accurate diagnosis. First, if the neuropsychologist remains unaware of existing external incentives, their likelihood of accurately detecting malingering may be diminished. Second, when these circumstances occur, neuropsychologists may feel significant pressure to provide diagnoses other than malingering in borderline cases in order to ensure that they are reimbursed for their work. These real-world influences highlight the possibility that judgments made by clinicians can be biased and objective decision strategies are needed.

Consequently, neuropsychologists should directly question their patients about the nature of the referral and if they have retained an attorney. Answers to these questions should be well documented. Documenting patients' responses can minimize later problems if the

patient has not been sincere in her or his answers. Neuropsychologists should make explicit agreements with clients regarding the need to access all pertinent medical records, high school records, and employment information. They should get consent for collateral interviewing. Refusal to agree to these stipulations should results in a refusal to evaluate a patient. These conditions are communicated to all referring professionals prior to evaluation.

MEDICAL-LEGAL HINT: Neuropsychologists are becoming increasingly savvy about the need to consider whether a client has retained an attorney for compensation-seeking litigation. Attorneys will have stronger cases to litigate if they encourage their clients to give accurate information to all professionals involved in a case. Neuropsychologists are better expert witnesses when all of the background information has been provided before they begin an evaluation.

Criterion 4: Determining If Incentives Are Internal or External

The *DSM-IV* requires that the clinician substantiate that any existing incentives to exaggerate symptoms or malinger be external rather than internal. Unfortunately, this concept, often referred to as locus of control, has long been debated by psychologists. At one extreme, behaviorists have argued that all incentives can be considered reinforcers

and that all reinforcement is external. These psychologists would then view all incentives as external. In contrast, psychoanalytic theorists have argued that there is a real difference between external and internal incentives, with internal incentives being driven by unmet emotional needs that were frustrated during childhood. The acquisition of an external incentive would not satisfy these more primitive needs.

As a result of the ongoing professional debate, we recommend that neuropsychologists narrowly define external incentives as economic incentives (e.g., disability payments, health care insurance coverage, civil litigation settlements), making them more concrete and quantifiable. Despite our redefinition, we believe that the neuropsychologist should not have to prove that external incentives exist in order to diagnose malingering. Conversely, we also believe that the presence of these incentives should not be insufficient to establish the existence of malingering. Many seriously injured patients are justifably seeking compensation for obvious and real impairments. Instead, the known existence of external incentives should be used to aid in establishing of a malingering diagnosis. Conversely, the lack of external incentives mitigate the diagnosis of malingering. Furthermore, the probability of diagnosing malingering should be directly related to the quantity, saliency, and

economic value of potential incentives.

MEDICAL-LEGAL HINT: External incentives are best thought of as economic incentives. When economic incentives are known to exist, the likelihood of a patient malingering is increased.

The Problem of Mixed Results in the Examination of Multiple Symptoms. Once primary decisions are established, their interaction must be considered. What is the diagnostic outcome when some signs or symptoms seem to have been feigned, whereas others seem legitimate? No techniques for determining the contribution of conflicting signs or symptoms to the final diagnosis have been specified in the literature. Consequently, this integration has also been left to clinical judgment. As a result, competent neuropsychologists with the same assessment results often come to opposite conclusions if they weigh the influence of these results differently.

To resolve this problem, we recommend that a neuropsychologist attempt to calculate the odds that a particular pattern of signs or symptoms would result from a particular lesion. If signs or symptoms

have unusually low odds of presenting together (i.e., less than 5%), it should be assumed that conscious intention was required for such a pattern to be exhibited. These calculations can be used to give an objective measure of the patient's level of intention. Another objective strategy is to compare a patient's scores to the normative scores obtained from patients who truly suffer from the alternative disorders. This second strategy also gives the neuropsychologist a method of calculating the probability that conscious feigning of signs or symptoms occurred. Finally, comorbidity of disorders is relatively common in this population. Multiple signs and symptoms may not be generated from a single diagnosis. Instead, they may represent multiple disorders. Consequently, for most patients the integration process may best be resolved by diagnosing multiple disorders. The disorder that appears to account for the most variance within a pattern of signs or symptoms should also be identified as primary.

ESTABLISHING CURRENT DIAGNOSIS

When neuropsychologists are asked to evaluate a patient, what if two disorders are suspected? Several categories of psychiatric disorders are often considered during the diagnostic process. These disorders typically include: (a) factitious disorders, (b) somatoform disorders (e.g., pain

disorder), (c) disorders of affect (e.g., major depressive episode), anxiety (e.g., panic disorder), and thought (e.g., paranoid schizophrenia), (d) other physical disorders (e.g., endocrine problems or other metabolic disorders); and (e) neurological dysfunction (e.g., closed head injury, stroke, Alzheimer's disease, learning disabilities). We specifically presented these alternative diagnoses in this order to reflect the underlying assumption about the neurological bases of each (i.e., from least neurologically based to most). Differential diagnosis then involves both determining the validity of patients' signs and symptoms, and basis for the expression of the symptoms. Neuropsychologists who diagnose malingering must also determine if a patient suffers concurrently from other disorders.

Many patients who malinger have valid psychiatric symptoms and may have had valid neurological symptoms at some point. Their psychiatric problems may lead them to misinterpret physical signs and poorly judge environmental contingencies. Consequently, neuropsychologists suspecting malingering should be acutely sensitive to other concurrent disorders, as they are likely to coexist.

MEDICAL-LEGAL HINT: A diagnosis of malingering is most likely to occur in a patient with psychiatric problems who has experienced valid physical signs and symptoms. Dual diagnosis in this population should be expected.

Traditional Methods of Detecting Signs of Feigning During Neuropsychological Assessment

Two reviews of malingering (Rogers et al., 1993; & Williams, 1998) appear in recent scientific literature, with a combination of overlapping recommendations revealing nine methods of detecting feigning (i.e., floor effect; performance curve; magnitude of error; atypical presentation; psychological sequalae; inattention; slow responding; haphazard, systematic, random, or sequential responding; and symptom validity tests). The floor effect was defined by Rogers et al., who suggested that a patient who is attempting to feign dysfunction will even fail tasks that severely impaired patients get right. A second method advocated by Rogers et al. is based on the hypothesis that patients who are attempting to feign often fail easier items whereas they pass more difficult ones. This can be seen if a clinician were to examine a patient's performance curve across test items and compare that to the items' degree of difficulty. Rogers et al. also hypothesized that patients who are feigning respond

with a magnitude of error that is uncommon for brain injured patients. For example, when asked "Who is the president of the United States?" a feigning patient might respond "George Washington," rather than with the more common error of "Ronald Reagan". Another method suggests that patients who feign are inconsistent or atypical in their presentation of signs and symptoms (Rogers et al., 1993). For example, a patient who shows up late for his or her appointment may claim to have forgot the appointment time. However, during the clinical interview, the same patient may give explicit details of the results of prior evaluations as well as the specific dates and times of when these evaluations were conducted. Rogers et al. also hypothesize that patients who feign neuropsychological deficits can be detected by examining their responses on noncognitive or personality inventories, as they give invalid responses to these personality inventories (e.g., Minnesota Multiphasic Personality Inventory, 2nd ed. [MMPI-2]).

Williams (1998) hypothesized that patients who are feigning attempt to do so by intentionally not paying attention. For example, when patients are told to listen to a series of digits that will be read to a patient, they may purposefully distract themselves so as not to hear the digits. Then, when asked to repeat the digits, they can sincerely report being unable to recall the correct series. Furthermore, Williams suggested that

patients who feign impairment respond more slowly to questions than do truly brain-injured patients. Their longer response time is thought to result from the additional time it takes to generate an incorrect response while inhibiting a correct one. Williams also hypothesized that patients who are attempting to feign cognitive impairment often develop a strategy for responding that does not involve giving the correct answer (e.g.,

JUDGMENT PROBLEM: Each of these recommendations is based on item response theory. The proposed methods of detecting malingering require statistics about items and/or tests that are rarely available (e.g., percentage of persons who answer correctly; items' and tests' ability to discriminate between normals, malingerers, and truly brain-injured patients; the correlation between patients' overall performance and proportion of those who scores correctly on an item or test; the frequency of incorrect responses given by different samples). Without these data, clinicians are left to make these judgments subjectively. Research has not supported clincians' ability to determine a patient's pattern of responding (Dawes, Meehl, & Faust, 1989; Goldberg, 1968; Wiggins & Hoffman, 1968).

answering every third item as true). Therefore, if a neuropsychologist examines a patient's response pattern carefully and notices a haphazard, systematic, or random pattern, this should be considered a sign of malingering.

Symptom Validity Tests. The last method proposed by both Rogers et al. (1993) and Williams (1998) recommended the use of symptom validity tests. These are typically forced-choice tests in which 50% of the right answers should be achieved by chance alone (e.g., predicting heads every time a coin is flipped will typically result in half of the predictions being correct). Patients who perform significantly below chance on forced-choice tests (e.g., 20% correct with a probability of less than .05 from the binomial distribution) are assumed to have achieved such poor performance by intentionally giving incorrect responses to questions. Of the nine proposed methods of detecting feigning, this method is the least susceptible to problems in human judgment.

JUDGMENT PROBLEM: The most effective method of detecting feigning involves the administration of multiple symptom validity tests. When results from these tests are inconsistent with one another, clinicians must integrate these divergent findings into a dichotomous decision about feigning. The outcome of these subjective decisions is likely to differ between clinicians.

We believe that the methods reviewed by Rogers et al. (1993) and Williams (1998) are reasonable approaches to detecting feigning. If a neuropsychologist notices any of these signs or symptoms during an evaluation of a patient, the level of suspicion that the patient is feigning should be raised and greater scrutiny of the patient's responses should ensue. One challenge that faces the neuropsychologist, however, is that the statistical information needed to make the judgments is often missing. Another challenge is that many neuropsychologists fail to utilize available statistical procedures, even when data are provided, because calculations are time consuming and non reimbursable. The problem is, however, that without these statistical analyses, neuropsychologists must depend on their subjective clinical experience and judgment to utilize the malingering criteria. These judgments often require the neuropsychologist to interpret

patients' data patterns. Research has shown that neuropsychologists make many mistakes when subjectively interpreting test scores and patterns (Arkes & Faust, 1987; Dawes & Corrigan, 1974). In part, this is because there is significant overlap between the patterns found in patients who are feigning and those found in neurologically impaired patients. These errors also occur because humans, in general, use simplified heuristics in their decision making that poorly detect complex patterns that exist in the data.

MEDICAL-LEGAL HINT: Diagnosis of malingering, based on an analysis of a patient's pattern of sign and symptom presentation, is open to legal challenge. Research has not supported psychologists' capacity to conduct pattern analysis (Arkes & Faust, 1987; Dawes & Corrigan, 1974). Accurate interpretatin of pattern analysis requires statistical analyses.

Assessment Problems Faced by Neuropsychologists Who Diagnose Malingering: Common Human Judgment Errors to Which Neuropsychologists Are Susceptible

When a neuropsychologist has diagnosed malingering without having completed any statistical calculations related to the probability that the diagnosis is accurate, the diagnosis is open to legal challenge. Wedding and Faust (1989) showed that neuropsychologists are just as likely as other

humans to commit judgment errors. They also pointed out that these errors could be avoided if neuropsychologists were to take better advantage of their statistical training and if they would fully use the technology that is readily available (e.g., personal computers and statistical software). Failure to do so reduces neuropsychologists' probability of successfully detecting. Wedding and Faust (1989) described five common errors in human judgment that have been shown to apply to neuropsychologists' diagnostic decisions.

Hindsight Bias. The hindsight bias may cause a neuropsychologist who has reviewed a patient's medical record to diagnose only those disorders that have already been noted in the record. For example, if a computer-aided tomography (CT) scan is read by a radiologist as being indicative of cerebral atrophy caused by Alzheimer's disease, the neuropsychologist will make the same diagnosis. Tests will be administered and scores interpreted in such a way as to make the diagnosis of Alzheimer's disease appear to be correct.

Confirmatory Bias. When a neuropsychologist hypothesizes that a memory disorder was caused by a motor vehicle accident, questionnaires may be administered that ask the patient to rate his or her memory.

Complaints of memory problems on these questionnaire are then used to support the original hypothesis. The problem with the strategy is that the neuropsychologist failed to recognize that complaints of memory problems are fairly common and may be due to other disorders that interfere with memory rather than to the traumatic brain injury (e.g., depression). The overlap between the two distributions may be unknown or ignored by the neuropsychologist. The selective gathering of evidence is particularly common if the hypothesis-testing model of assessment is followed (e.g., see Lezak, 1995).

Overreliance on Salient Data. Some neuropsychologists may believe that a certain test score is a pathognomonic sign of a disorder (e.g., Reitan, 1986). Thus, when a patient exhibits this sign, a neuropsychologist concludes that the disorder is present. Additional test scores that are inconsistent with the diagnosis are ignored. For example, if a neuropsychologist concludes that a patient has feigned memory impairment based on the patient's deceptive responses to questions regarding psychiatric history, s/he may ignore the patient's performance on more objective tests that are valid indicators of neurological dysfunction.

Under Utilization of Base Rates. Neuropsychologists may underutilize base- rate information (Arkes, 1981). For example, if a patient's test scores fall into a range that is common for patients who suffer from a rare disorder, neuropsychologists may be inclined to diagnosis the patient with the rare disorder. The error is that, although patients with the rare disorder may always score in this low range, low scores may have also been obtained by patients who suffer from more common disorders. Typically, the most likely cause for the low score is the common disorder and not the rare disorder. For example, personality change and poor judgment in a male who is in his 50s may be due to Pick's disease a relatively rare disorder. However, this same pattern is also seen in patients suffering from Alzheimer's disease. Although Pick's disease tends to strike persons at an earlier age than does Alzheimer's disease, Alzheimer's disease is five times more common in this age range than is Pick's disease. Therefore, the abnormality is more likely caused by Alzheimer's disease than it is by Pick's disease.

Failure to Analyze Co-Variation. A neuropsychologist may have administered one test that results in poor performance by the patient. To validate this finding, the neuropsychologist administers a second test that

is highly correlated with the first. When a similar pattern of poor performance is found on the second test, the neuropsychologist assumes that his or her diagnosis of dysfunction has been confirmed. However, if the first test did not measure the construct of interest accurately, the correlated second test may simply replicate invalid findings rather than substantiate the diagnosis per se (Chapman & Chapman, 1969).

A STATISTICAL AND ACTUARIAL PROCEDURE FOR THE ANALYSIS OF NEUROPSYCHOLOGICAL DATA

Williams (1997) noted that despite the research that shows how these common human biases in decision making also exist for neuropsychologists, awareness of these biases has not prevented them from occurring. Essentially, admonishing neuropsychologists to think better has not been found to alter their capacity to make diagnoses with any greater reliability or validity. Therefore, Williams recommended using technology be utilized that would function as a decision aid. These decision aids would then help neuropsychologists correctly interpret psychometric test data and increase the accuracy of their diagnoses (Sicoly, 1989). Consistent with several researchers' recommendations (e.g., Garb & Schramke, 1996; Sicoly's, 1989; Williams, 1997), we have developed a

process of data analysis, called the Rohling Interpretive Method (RIM), that can be programmed on most personal computers. This RIM will help neuropsychologists overcome common human judgment errors. The logic of the RIM has been presented in greater detail elsewhere (Miller & Rohling, 2001).

In this chapter, we highlight how the Rohling Interpretive Method (RIM) helps a neuropsychologist avoid the biases noted by Wedding and Faust (1989). The methodology used in the RIM is similar to that recommended for meta-analytic reviews of research literature (e.g., see Glass, McGaw, & M. L. Smith, 1981; Rosenthal, 1984). The linear combination of scores we recommend is supported by the research of Dawes (1979), Dawes and Corrigan (1974), and Heaton et al. (2001). It follows a model similar to that presented by Kiernan and Matthews (1976). We believe it to be a more statistically sound method of analysis which improves upon the Impairment Index (II) of Reitan and Wolfson (1985), and the Average Impairment Rating (AIR) of Russell, Neuringer, and Goldstein (1970). A further advantage is that it does not restrict a neuropsychologist to a particular battery of tests.

The steps of the RIM are listed in Table 8.1. When these steps are conducted on an individual case basis with a calculator, they can be time

consuming. However, by programming a personal computer with commonly available statistical software (e.g., Microsoft Excel; SPSS for Windows; Statview; sample program available from the authors), each of these steps is easily automated. Once automated for an initial case, the time it takes to complete future interpretations is reduced. In fact, it takes less time to conduct a RIM interpretation than it takes to complete a more traditional interpretation that does not require that these calculations be completed.

Each step of the RIM process is illustrated with two example cases. The tables and graphs that are presented for these cases are referred to throughout the case description. Steps 1 through 17 of the RIM process generate a table of summary statistics, which is also put in graphic form. Steps 18 through 24 describe the interpretation of the summary statistics.

Both case examples were referred for assessment by an attorney and litigation was expected. Issues of financial compensation existed in both cases and concern over malingering or symptom exaggeration also applied. In the first case, a traumatic-brain-injured patient is presented who had clear evidence of neurological impairment (e.g., LOC [loss of consciousness] of 17 days, positive findings on CT scanning, MRI [magnetic resonance imaging] scanning, EEG [electroencephalogram]

TABlE 8.1
Steps of Rohling's Evaluation -
Miller's Interpretation of Neuropsychological Data (RIM).

Summary Statistics: Steps 1-17.	13. Determine the upper limit necessary for premorbid performance.
1. Design and administer a flexible test battery.	14. Conduct one-sample *t* tests on each of the means generated.
2. Estimate premorbid general ability (EPGA).	15. Conduct a between-subjects ANOVA with domain means.
3. Convert test scores to a common metric (e.g., T-scores with $M = 50$, $sd = 10$).	16. Conduct a power analysis for each of the domains and the TBM scores.
4. Assign each test score to the cognitive domain for which it has the highest factor loadings.	17. Sort test scores in ascending order for qualitative inspection
5. Calculate cognitive domain means, standard deviations, and sample sizes.	***Interpretation: Steps 18-24***
6. Calculate test battery means (TBM), including an Overall, Domain, and Instrument TBM.	18. Determine the test battery's validity by examining the symptom validity, sample size, and heterogeneity statistics.
7. Calculate the heterogeneity probabilities for each domain.	19. Determine if psychopathology influenced test scores.
8. Assign categories of impairment.	20. Use test battery means to determine if impairment exists.
9. Determine the percentage of test scores that fall in the impaired range.	21. Determine patterns of strengths and weaknesses across the domains.
10. Graphically display all of the summary statistics in T-score format.	22. Examine the mean T-scores from the noncognitive domains.
11. Calculate effect sizes for all domains and TBM scores.	23. Explore low power comparisons for Type II errors.
12. Calculate a confidence intervals for all domains and TBM scores.	24. Examine the response operating characteristics of sorted T-scores.

results, and neurological examination). Case 1 is presented to give the reader a clear understanding of how a brain-injured patient's RIM output

is examined. Case 2 is that of a suspected malingering patient. In Case 2, in contrast to Case 1, there were no positive findings upon presentation at the emergency room, nor on CT scanning, MRI scanning, EEG readings, nor during the neurological exam. Case 2 is meant to illustrate the advantages of the RIM process over traditional interpretations when milder injures are involved.

Case 1: Traumatic Brain Injury Caused by a Motor Vehicle Accident

Case 1 is that of Truman Hurt, a 22-year-old white right-handed man who allegedly suffered a head injury in a motor vehicle accident (MVA). Mr. Hurt's attorney referred him for an evaluation. Litigation processes were begun when Mr. Hurt's insurance claims were denied. The assessment issues focused on Mr. Hurt's premorbid use of alcohol and illicit drugs. Mr. Hurt's attorney believed that the defense might argue that Mr. Hurt's cognitive deficits preceded his MVA and were due to alcohol and/or substance abuse. If this explanation were accepted, the insurance company could not be held responsible for Mr. Hurt's neuropsychological deficits; the company could then justify their denial of compensation.

Mr. Hurt was administered a battery of tests to assess the extent of his cognitive deficits, the likelihood that these deficits could be

TALBE 8.2a
RIM scores for Mr. Hurt

	M	*sd*	*n*	*Hetero. p value*	*Classify*	*%TI*
Sx Validity (SV)	60.0	0.0	2	---	Superior	0.00
Emotion/Person (EP)	55.4	12.3	11	---	High Avg.	0.00
EPGA	48.1	4.4	6	---	Average	0.00
Overall TBM (OTBM)	37.2	11.4	63	<.005	Mild	0.46
Domain TBM (DTBM)	35.5	6.8	7	---	Mild	0.86
Instrument TBM (ITBM)	38.0	10.4	14	---	Mild	0.50
Verbal (VC)	36.5	10.3	9	---	Mild	0.56
Performance (PO)	45.0	13.5	8	<.10	Average	0.25
Executive (EF)	38.1	8.1	11	---	Mild	0.54
Memory (ML)	36.8	15.5	9	<.025	Mild	0.33
Attention (AW)	35.8	7.5	14	---	Mild	0.50
Process Speed (PS)	21.5	2.1	2	---	Mod-to-Sev	1.00
Language Aphasia (LA)	36.0	13.3	10	<.10	Mild	0.50

attributed to the MVA, and the truthfulness of his responses. Data generated from this test battery were then subjected to the RIM process. Tables 2a and 2b show the summary statistics generated by the RIM process. Figure 8.1 is a graphical portrayal of these same results. The mean for the symptom validity domain (SV) fell within the above-average range. This indicates that Mr. Hurt was probably responding truthfully to the assessment measures and that the results can be meaningfully interpreted

as representing neurocognitive functioning. Sample sizes (*n*) appeared relatively adequate, particularly for those domains in which there was significant heterogeneity (see column headed Hetero. *p* value). Finally, the heterogeneity of the test results could be explained by Mr. Hurt's lateralized left-hemisphere lesion, focused at the intersection of the frontal, parietal, and temporal lobes. When the variability of this lateralization was considered, the cause of the heterogeneity of his results seemed apparent.

Regarding psychopathology, the mean score for the *emotional-personality (EP)* domain is in the high-average range and not indicative of significant psychopathology that might have influenced Mr. Hurt's test results. Looking at the tests score more closely, all validity indices were within the normal range and his pattern of scores did not suggest any obvious DSM-IV diagnostic category.

Mr. Hurt's *Estimated Premorbid General Ability (EPGA)* indicates that his general ability was within the average range. Furthermore, there is little variability in the *EPGA*. This indicates that it is unlikely that Mr. Hurt had a preexisting ability deficit (e.g., a learning disability). Small variability in the *EPGA* also increases the ease of interpreting the current assessment. Because there is little variability in the *EPGA* and the estimate

TABLE 8.2b
RIM Scores for Mr. Hurt (continued)

	ES	CI	PreM Nec.	1-sample *t* Test	ANOVA S&W	Power (1-b)
Domain						
SV	1.00	0.0	60.0	---	---	---
EP	0.54	6.7	62.1	---	---	---
EPGA	-0.19	3.6	44.5	---	---	---
OTBM	-1.09	2.4	39.6	<.0001	---	---
DTBM	-1.26	5.0	40.5	<.0020	---	---
ITBM	-1.01	4.9	42.9	<.0000	---	---
VCl	-1.16	6.4	42.9	<.0050	---	---
PO	-0.31	9.0	54.0	---	S	0.32
EF	-1.00	4.4	43.5	<.0010	---	---
ML	-1.13	9.6	46.4	<.0300	---	0.97
AW	-1.23	3.6	39.4	<.0030	---	---
PS	-2.66	9.5	31.0	<.0200	W	---
LA	-1.21	7.7	43.7	<.0100	---	---

is in the normal range, it is unlikely that Mr. Hurt was showing significant impairments in his ability prior to the accident that could be attributed to drug and alcohol use.

The next three rows in Tables 8.2a and 8.2b (*Overall Test Battery Mean [OTBM], Domain Test Battery Mean [DTBM],* and *Instrument Test Battery Mean [ITBM]* depict an overall view of Mr. Hurt's performance on

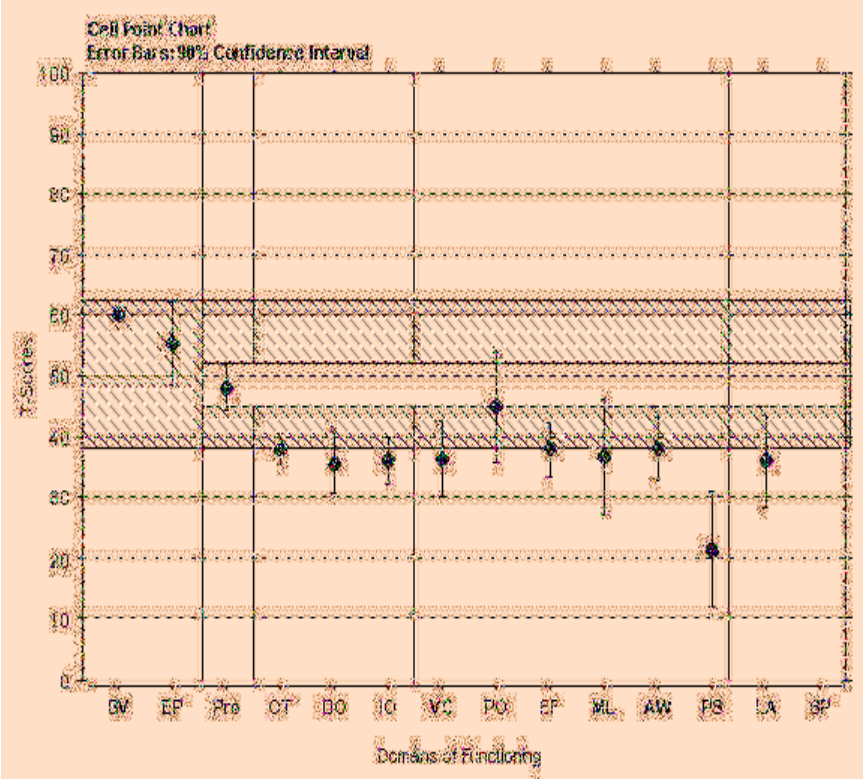

FIG. 8.1. Graphic Profile of Mr. Hurt.

the assessment battery. Ideally, these three numbers should yield highly similar results. This the case for Mr. Hurt. Specifically, when examining the *ES*s for each of these estimates (Table 8.2b), they are all of almost equal magnitude (-1.09, -1.26, and -1.01, respectively). Furthermore, as can be seen in Figure 8.1, all three global measures of performance on the test

battery fall well below the average range. Finally, the percentage of tests impaired also suggests significant impairment across the test battery (46%, 86%, and 50%, respectively). Examining the premorbid level necessary for these results to be significant (PreM-Nec, Table 8.2b), all of the test battery means fell below the required 44.5. Finally, the one-sample *t*-test results (Table 8.2b) show significant impairment that is quite unlikely to have occurred by chance alone. These results strongly support the conclusion that Mr. Hurt has suffered a head injury that has resulted in residual cognitive deficits.

However, as depicted in Table 8.2a, Hetero *p* value, recall that there was significant variability in the *OTBM*. When this is the case, there is a possibility that the individual has significant cognitive strengths and weaknesses. This is the case for Mr. Hurt. Verbal Comprehension - Aphasia language measures on the right side of Fig. 8.1 display. Specifically, Mr. Hurt has a significant deficit in his processing speed. Conversely, his performance domain is a relative strength, showing less impairment than in other areas. Finally, as discussed previously in this chapter, interpretations of neuropsychological test data are enhanced by direct comparisons of the obtained results to the results that could be

predicted from the literature. Using data from the nature of the injury Mr. Hurt sustained (e.g., time to follow commands of 17 days), and interpolating from data published by Rohling, Millis, and Meyers (2000), it was estimated that Mr. Hurt's *T*-score should be 35.8 (ranging from a *T*-score of 27.4 to 42.7 at the 90% confidence interval). Once again, the obtained *T*-scores (Table 8.2b) were exactly as would be expected, with the exceptions noted earlier (i.e., performance skills less impaired and processing speed more impaired).

Looking next at Mr. Hurt's language domain, there was evidence of heterogeneity in his overall score. This was evident in his test scores that were most likely to be lateralized. That is, he demonstrated problems with naming, reading, writing, and fluency whereas there was no evidence of impairment in his gestural and graphic abilities.

No evidence of a learning disability or premorbid strengths or weaknesses was evident in Mr. Hurt's academic record. This came from several standardized test results he had been administered (e.g., Otis-Lennon Test, Stanford Achievement Test, ACT [American College Test], and high school grades). None of his prior records would have predicted a significant strength on the low power domain of performance skills

(Power, Table 822b). Therefore, it is believed that this is a consequence of his accident and argues against impairments being due to alcohol or substance abuse, which tends to have a nonlateralized pattern of impairments.

Finally, in looking at the numerically sorted *T*-scores, there was no evidence that the pattern of scores found in the summary statistics was contradicted by the performance curve found in his numerically sorted *T*-scores. This is the same pattern described previously. Thus, there is no reason to believe that his test results were inadvertently misinterpreted by using inappropriate tests, norms, or caused by impaired motivation.

As a result of this interpretation, the neuropsychologist was confident in asserting that Mr. Hurt was experiencing impairment due to a neurological injury. Clearly, this individual was not malingering. Instead, his pattern of performance was what one would anticipate as a result of the head injury he experienced in the MVA. Moreover, there was no evidence to support the insurance company's assertion that these deficits existed prior to the accident and were a result of Mr. Hurt's drug and alcohol use.

Case 2: Malingering of Memory Impairment

Mea Fake was a 32-year-old white left-handed woman who claimed to have suffered a mild head injury at work. She was referred by a managed care company that was responsible for handling her workers' compensation claim. The primary issue in Ms. Fake's case was the determination of the cause of her persistent symptoms (e.g., headache and memory problems). Were these complaints due to (a) neurological injury, (b) preexisting psychiatric dysfunction (e.g., depression), or (c) malingering? This was the second referral of Ms. Fake's case; the first involved simply a review of records to determine the quality of her prior assessments and/or treatments, and to recommend what, if anything, should be done for her in the future. The only neuropsychological testing following her initial accident was with the Wechsler Adult Intelligence Scale -Revised (WAIS-R), Wechsler Memory Scale-Revised (WMS-R), and the D-scale of the MMPI-2. While reviewing Ms. Fake's records, it was noted that the prior evaluation was considered to have been inadequate for this type of case. First, there was no estimate of her premorbid general ability. Second, no tests were administered to assess the validity of her signs or symptoms. Third, the D-scale of the MMPI-2 was considered to have been an inadequate assessment of her emotional or personality status. Finally, although the psychologist's conclusions may have been correct,

based on the available data, the potential for litigation in this case also caused concerned about the brevity of the evaluation.

As a result, the review of records resulted in a second assessment that was more comprehensive than that administered by the first psychologist. Consistent with the RIM process outlined previously, Ms. Fake was administered a battery of tests to assess symptom validity, personality functioning, and cognitive ability. Appendix A lists each of the test instruments administered and the scores obtained. Data were collected to derive estimates of Ms. Fake's premorbid level of functioning. Information was also collected about the work-related head injury in order to generate estimates of the extent of deficits that could be expected from this type of trauma.

Tables 8.3a, 8.3b, and Fig. 8. 2 show Ms. Fake's results as generated by the RIM process. The first thing to notice is Ms. Fake's symptom validity results. Ms. Fake's SV domain score fell into the impaired range (Table 8.3a). This indicates that there is a strong probability that Ms. Fake was exaggerating or feigning signs and symptoms. The diagnosis of malingering should be considered at this point. In addition, the invalidity of these measures suggests that all other data generated from the assessment should be viewed with caution. According to Step 18 of the

TABLE 8.3a.
RIM scores of Ms. Fake

	M	*sd*	*n*	*Hetero p value*	*Classify*	*%TI*
SV	17.3	20.1	11	<.0001	Severe	0.73
EP	44.4	13.5	10	<.0500	Below Avg.	0.30
EPGA	44.9	1.2	6	---	Below Avg.	0.00
OTBM	39.5	16.3	53	<.0001	Mild	0.41
DTBM	42.8	11.2	6	---	Below Avg.	0.33
ITBM	44.4	13.4	10	---	Below Avg.	0.40
VC	39.4	7.4	6	---	Mild	0.50
PO	43.2	13.2	8	<.1000	Below Avg.	0.38
EF	59.3	16.8	7	<.0500	Superior	0.14
ML	24.8	15.2	16	<.0050	Mod-to-Sev	0.75
AW	43.5	6.4	15	---	Below Avg.	0.27
LA	46.7	---	1	---	Average	0.00

RIM process, it is important that the neuropsychologist next review the heterogeneity of variance (Hetero*p* value). This tells the examiner whether the data from all the tests that were combined told the same story (e.g., did all the measures of personality have the same results?). Considerable variability was obtained on the symptom validity measures, the emotion/personality domain, the OTBM, as well as the performance skills, executive skills, and memory domains. A review of Ms. Fake's

TABLE 8.3b.
RIM scores of Ms. Fake continued

	ES	*CI*	*PreM Nec.*	*1-sample t Test*	*ANOVA S&W*	*Power (1- b)*
SV	-3.27	11.0	28.3	<.0002	W	---
EP	-0.56	7.8	52.2	<.1100	---	0.66
EPGA	-0.51	1.0	44.0	<.0001	---	---
OTBM	-0.54	3.8	43.3	<.0100	---	---
DTBM	-0.21	9.2	52.0	---	---	0.21
ITBM	-0.05	7.8	52.2	---	---	0.13
VC	-0.55	6.1	45.5	<.0700	---	0.48
PO	-0.17	8.8	52.1	---	---	0.20
EF	1.44	12.4	71.7	<.0400	S	---
ML	-2.01	6.7	31.5	<.0001	W	---
AW	-0.14	2.9	46.5	---	---	0.22
LA	0.18	---	---	---	---	---

performance on the symptom validity tests indicates that although most of the symptom validity tests identified Ms. Fake as an invalid responder (i.e., 73% in %TI), she did less exaggerating on the more obvious tests of malingering. The variability in Ms. Fake's emotional/personality data appeared to be due to Ms. Fake's higher scores on measures of hypochondriasis, depression, and hysteria on the MMPI-2. This led the

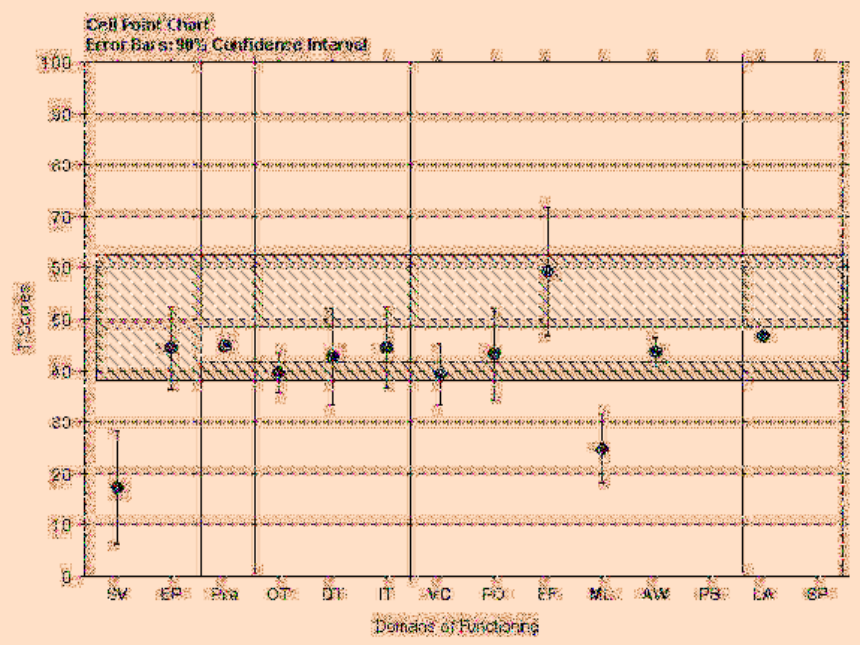

FIG. 8.2. Graphic Profile of Ms. Fake

neuropsychologist to be concerned that test scores may also have been influenced by the non-neurological factors of depression and somatosization. The variability in Ms. Fake's OTBM indicates that Ms. Fake is presenting a pattern of relative strengths and weaknesses on the cognitive data. As is evident from Fig. 8.2, Ms. Fake is performing significantly poorly on the memory and learning measures, but did particularly well on the measures of executive skills. All attempts to eliminate this heterogeneity were unsuccessful. That is, there were no signs

of lateralization in her test scores; nor was there evidence of a preexisting learning disability. Furthermore, when the effect of a disproportionate number of tests that were administered in the domains of memory and attention were reduced by examining the ITBM (i.e., one mean score per instrument administered), the *ES* shrank from -.54 to -.05. This smaller effect size was not significantly different from her EPGA and raises doubt about the existence of residual cognitive deficits caused by a neurological injury.

Furthermore, Ms. Fake's EPGA clearly and consistently placed her in the low-average range of ability. Little variability was obtained in Ms. Fake's EPGA, estimate indicating that she probability had few noticeable strengths or weaknesses in her cognitive profile prior to the alleged head injury. In contrast, her current assessment results indicate a great deal of variability, both across the various domains and within domains, as there was also excess variability on Ms. Fake's performance skills, executive skills, and memory domains. A review of the specific test results suggests the following: Ms. Fake was an unsophisticated test taker. Consistent with her presenting complaint, she performed poorly on all obvious measures of memory. However, she did not perform nearly as poorly when memory

functioning was assessed more subtly.

Using the OTBM results for Ms. Fake, and interpolating from data published by Rohling, Mills, and Myers (2000), it was expected that Ms. Fake should have been unconscious for four days, with a 90% confidence interval ranging from 16 hours to 9 days. Ms. Fake reported that she was unconscious for 10 minutes and recalls traveling in the ambulance to the emergency room. The amount of time she claimed to be unconscious is 500 times less than the mean that was predicted from her OTBM. This strongly suggests that Ms. Fake was feigning or exaggerating symptoms.

Several of the domains may have had insufficient statistical power to detect small residual deficits. This should raise the neuropsychologist's concern about making a Type II error (i.e., concluding no cognitive deficits exist when they actually do). The chances of detecting *ES*s in the range seen in Ms. Fake's data are quite low (i.e., see Power, Table 8.3b), ranging from .13 to .21 for her DTBM and ITBM, respectively. If the neuropsychologist were to try and administer enough tests for these ESs to be reliably detected, the number of tests administered would range between 165 and 2,650. Considering that the current test battery, with a sample size of 53, took approximately 8 hours to complete, this would

mean that Ms. Fake would have to be assessed from between 3 days and 10 weeks. Clearly, this amount of time is an unreasonable solution to the problem of low power, particularly because the validity indices of the test battery raise serious doubt as to the motivation of Ms. Fake to perform to the best of her ability.

Finally, examining the numerically sorted *T*-scores, Ms. Fake did particularly poorly on some of the more simple tests (e.g., PDRT, [Portland Digit Recognition Test]) and quite well on the more difficult tests (e.g., WCST, [Wisconsin Card Sorting Test]). This floor effect is further evidence of the probability that she was attempting to feign cognitive impairment.

As a result of this interpretation, the neuropsychologist should strongly consider a diagnose of malingering. Her obtained test scores were inconsistent with the alleged injury. Furthermore, given past research on the influence of compensation on mild head injury, it seems likely that awarding compensation to this patient would only enhance her experience of symptoms.

Case Examples Summary

Summarizing the findings of the two cases, it seems clear that Mr. Hurt and Ms. Fake each came to the evaluation sessions via similar referral

mechanisms. Both had external incentives, however the resulting diagnoses were clearly different. Mr. Hurt did not receive a psychiatric diagnosis and all of his residual deficits were attributed to the injuries he suffered in his MVA. On the other hand, Ms. Fake was given three psychiatric diagnoses, which included malingering, depression, and dependent personality disorder. None of her cognitive complaints were believed to be due to a neurological injury.

Competent neuropsychologists who use their statistical training, well-standardized test instruments, up-to-date knowledge of brain-behavior relationships, and technological aids such as personal computers and statistical software, can complete a state-of-the-art assessment that is both reliably and valid. The avoids most of the problems highlighted by Wedding and Faust (1989), is objective, can be replicated, and is easy to present in court. It operationalizes many of the suggests made by other researchers (e.g., Rogers et al., 1993; Williams, 1998) in the assessment of malingering. The RIM process does not depend on the personal insight or experience to complete. Clear decision rules can be applied. It helps a neuropsychologist avoid overconfidence in his or her diagnoses. It facilitates the separation of statistical fact from artifact, without allowing a neuropsychologist to become overreliant on highly salient data. It

statistically considers covariation in test measures and facilitates the neuropsychologist's use of only the most valid of data. Finally, it assists the neuropsychologist in assessing alternative or concurrent diagnoses and the use of base rates of these diagnoses.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*, (4th ed.). Washington, DC: Author.

Arkes, H. R. (1981). Impediments to actual clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology, 49*, 323-330.

Arkes, H. R., & Faust, D. (1987, August). Cue utilization in neuropsychology. In D. Faust (Chair), *Clinical judgment in neuropsychology. How good are we?* Symposium conducted at the annual convention of the american psychological association, New York.

Binder, L. M. (1993). *Portland Digit Recognition Test* (2nd ed.). Portland, OR: Author.

Binder, L. M., & Rohling, M. L. (1996). Money matters: A meta-analysis of the effect of financial incentives on recovery from closed-head injury. *American Journal of Psychiatry, 153*, 7-10.

Binder, L. M., Rohling, M. L., & Larrabee, G. (1997). A review of mild head trauma. Part I: A meta-analytic review of neuropsychological studies. *Journal of Clinical and Experimental Neuropsychology, 19*, 421-431.

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

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*, 217-280.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*, 571-582.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95-106.

Dawes, R. M., Meehl, P. E., & Faust, D. (1989). Clinical versus statistical prediction of human outcomes. *Science, 243*, 1668-1674.

Dikmen, S. S., Machamer, J. E., Winn, H. R., & Temkin, N. R. (1995). Neuropsychological outcome 1-year post head injury. *Neuropsychology, 9*, 80-90.

Dunn, L. M., & Dunn, E. S. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MN: American Guidance Service.

Faust, D., Hart, K., & 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., Guilmette, T. J., & Arkes, H. R. (1988). Neuropsychologists' capacity to detect adolescent malingering. *Professional Psychology: Research and Practice*, 19, 508-515.

Faust, D., & Ziskin, J. (1988). The expert witness in psychology and psychiatry. *Science, 241*, 31-35.

Garb, H. N., & Schramke, C. J. (1996). Judgment research and neuropsychological assessment: A narrative review and meta-analysis. *Psychological Bulletin, 120*, 140-153.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Goldberg, L. R. (1968). Simple models or simple processes; Some research on clinical judgments. *American Psychologist, 23*, 483-496.

Greene, R. L. (1979). Response consistency on the MMPI: The TR index. *Journal of Personality Assessment, 43*, 69-71.

Guilmette, T. J., Sparadeo, F. R., Whelihan, W., & Buongiorno, G. (1994). Validity of neuropsychological test results in disability evaluations.

Perceptual and Motor Skills, 78, 1179-1186.

Heaton, R. K., Gladsjo, J. A., Palmer, B. W., Kuck, J., Maracotte, T. D., & Jeste, D. V. (2001). Stability and course of neuropsychological deficits in schizophrenia. *Archives of General Psychiatry, 58*, 24-32.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological tests norms above age 55: COWAT, BNT, Token, WRAT-R Reading, AMNART, Stroop, TMT, JLO. *The Clinical Neuropsychologist*, 10, 262-278.

Kiernan, R. J., & Matthews, C. G. (1976). Impairment index versus T-score averaging in neuropsychological assessment. *Journal of Consulting and Clinical Psychology, 44*, 951-957.

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

McCaffrey, R. J., & Lynch, J. K. (1992). A methodological review of method skeptic reports. *Neuropsychology Review, 3*, 235-248.

Miller, L. S., & Rohling, M. L. (2000). A statistical interpretive method for neuropsychological test data. *Neuropsychological Review, 11*, 141-167.

Millis, S. R. (1994). Assessment of motivation and memory with the Recognition Memory Test after financially compensable mild head injury. *Journal of Clinical Psychology, 50*, 601-605.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting and Clinical Psychology, 29*, 261-265.

Reitan, R. M. (1986). Theoretical and methodological bases of the Halstead-Reitan Neuropsychological Battery. In G. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 3-29). New York: Oxford University Press.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery. Theory and clinical interpretation.* Tucson, AZ: Neuropsychological 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.

Rohling, M. L., Binder, L., & Langhinrichsen-Rohling, J. (1995). Money matters: A meta-analysis of the association between of financial compensation and the experience and treatment of chronic pain. *Health Psychology, 14*, 537-547.

Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Russell, E. W., Neuringer, C., & Goldstein, G. (1970). *Assessment of brain damage: A neuropsychological key approach*. New York: Wiley-Interscience.

Sicoly, F. (1989). Computer-aided decisions in human services: Expert systems and multivariate models. *Computers in Human Behavior, 5*, 47-60.

Trueblood, W., & Binder, L. M. (1997). Psychologists' accuracy in identifying neuropsychological test protocols of clinical malingerers. *Archives of Clinical Neuropsychology, 12*, 13-27.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. San Antonio TX: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised*. San Antonio, TX: The Psychological Corporation.

Wedding, D. (1983). Clinical and statistical prediction in neuropsychology. *Clinical Neuropsychology, 5*, 49-54.

Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology, 4*, 233-265.

Wiggins, N., & Hoffman, P. J. (1968). Three models of clinical judgment. *Journal of Abnormal Psychology, 73*, 70-77.

Williams, J. M. (1997). The prediction of premorbid memory ability. *Archives of Clinical Neuropsychology, 12*, 745-756.

Williams, J. M. (1998). The malingering of memory disorder. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation. Critical issues in neuropsychology*. (pp 105-132). New York: Plenum.

Youngjohn, J. R. (1995). Confirmed attorney coaching prior to neuropsychological evaluation. *Assessment, 2*, 279-283.