INDUSTRIAL REHABILITATION QUARTERLY
Volume 1, Number 1
M & A Publication. Spring, 1988

"How do you know that he tried his best?"

The Reliability Crisis in Industrial Rehabilitation

Leonard N. Matheson, Ph.D.

The issue of reliability is pertinent to every evaluation procedure. Reliability can be defined in terms of the stability or dependability of a measure; that the measure will be stable over time and without regard to the evaluee, to the evaluator, or to the circumstances of testing.

The developers of most of the performance evaluation techniques and devices that are commercially available have demonstrated the reliability of their procedures in studies which assume that the evaluee is providing maximum voluntary effort (MVE) in the evaluation task. This is not an acceptable assumption in industrial rehabilitation. Most experienced industrial rehabilitation professionals recognize that the medico-legal context of practice makes it more reasonable to doubt the evaluee's motivation to fully participate than to accept it as a declared fact. This sceptical position is nurtured by the adversarial nature of most industrial rehabilitation practiced in the United States and Canada. Although we do not take sides; although we strive to provide information that is independent of the circumstances of the case, those who use our information are on one side of the issue or on the other. As a consequence, our information is used (or suppressed) by one side against the other.

Given the importance of the information that we provide, the side for whom it is potentially damaging will attempt to discredit the work on which the information is based. One of the easiest and most certain methods to cast doubt on the professional opinion of an industrial rehabilitation professional is to attack the reliability of the findings that the professional's evaluation procedures have generated. In fact, I must admit that, when hired as a consultant to review a medico-legal case, I often recommend an attack on the reliability of the test results. If the industrial rehabilitation professional is one of that currently rare breed who has taken steps to assure reliability, I am left to attack the validity of the findings, a difficult and often unsuccessful gambit If we (me and the attorney who has hired me as a Benedict Arnold to my profession) are unable to assail the reliability of the other side's findings, we are reduced to what attorneys call a "swearing contest" in which the professional opinions of one side's experts are pitted against the other side's experts. You may have thought that this is what the tort process was in the first place but you were wrong. Attorneys are not gamblers, nor are the insurance companies who hire them. Their attack on your reliability is their last technical bastion, the last place they can "knock you out of the box" and avoid a swearing contest that will be decided by a completely independent jury ... who may not find in their favour. And so you must recognize an attack on the reliability of your results for what it is: The last opportunity for the opposing attorney to control your testimony. Once you clear the reliability hurdle, your opinion

counts. Until then, you are not truly an expert

Performance measures are highly dependent on the effort that is expended by the evaluee. The question, "How do we know that he tried his best?" is a challenge to the reliability of the performance that is being measured and, ultimately, to the professional acumen of the industrial rehabilitation practitioner.

Threats to the reliability of performance measures that are obtained in an industrial rehabilitation setting include several that have been identified as evaluee-resident threats to reliability (Matheson, 1986):

1. Unidentified impairment;
2. Easy fatigue-ability;
3. Fear of re-injury or pain;
4. Test anxiety
5. Symptom magnification syndrome

The first two threats to reliability are straightforward and relatively easy to identify. The third and fourth, rooted in the evaluee's response to the testing situation, have been demonstrated to be amenable to controls, such as the preparatory procedures now used before any well-designed cardiac stress test or psychological evaluation. The last factor, symptom magnification syndrome (SMS), defined as "a conscious or unconscious self-destructive socially reinforced behavioural response pattern consisting of reports or displays of symptoms which function to control the life circumstances of the sufferer" (Matheson, 1987), is difficult to evaluate.

However, according to research conducted at the Employment and Rehabilitation Institute of California in 1987, SMS is a pervasive problem. This survey was based on case reviews of more than 800 evaluees and initially identified 377 cases with data sets which were sufficiently complete to make a determination. This subset was composed of industrial rehabilitation evaluees who had been off of work for an average of two years and one month. Of these cases, 91 evaluees (24%) were classified as suffering from the symptom magnification syndrome. Further research will focus on gender differences and the effect of time since injury and age on incidence.

Coefficient of Variation

The direct measurement of maximum voluntary effort should be an integral part of every industrial rehabilitation evaluation. One of the most straightforward methods to measure maximum voluntary effort is based upon the assumption that repetitive trials within a brief span of time will be stable. A statistical measurement of stability that is used frequently in performance testing is the coefficient of variation (CV). This statistic is the quotient of the standard deviation of a set of scores divided by its mean. It is expressed as a percentage and is a convenient notation that can be used to compare the stability of measures on a ratio scale (with a true zero). The correct formula is:

$$CV = \frac{\text{Standard Deviation}}{\text{Mean}}$$

Although most people are using the correct formula to calculate standard deviation, I have found several people who are using the formula that includes a correction factor, so that the standard deviation becomes an unbiased estimate of a population parameter. While this is appropriate if you have selected only certain scores from a larger population of scores and want to be confident that the standard deviation that you calculate of the sample of scores you have collected is a good estimate of the population's standard deviation, it is not the proper statistic to use in the maximum voluntary effort process.

The standard deviation that is used in the maximum voluntary effort process reflects the variability of the single evaluee's scores and is not based on a sample of his scores but, rather, on all of his scores for that particular trial. Those who are using a scientific calculator with built-in standard deviation function should perform a calculation by hand for comparison. Many of the scientific calculators use the formula that provides a correction factor which, as I have noted above, is inappropriate The correct formula is:

The formula below is for the standard deviation that is an unbiased estimate of a population parameter and is incorrect for use in calculation of the coefficient of variation of a set of scores. While use of this formula is a simple error. it results in scores that are incorrect, producing a coefficient of variation that is inflated.

General Guidelines for Test Selection

While many performance evaluation instruments are available, only a few are appropriate for the measurement of maximum voluntary effort in an Industrial Rehabilitation setting. The “General Guidelines for the Selection of Measures of Maximum Voluntary Effort” (Matheson, I986) indicate that such instruments must:

1. Not require cardiovascular effort that exceeds 65% of predicted maximum heart rate.
2. Not directly involve an impaired component of the biomechanical system..
3. Be controlled by the evaluee.
4. Have low error variance.
5. Have high inherent stability relative to the range of recorded values.
6. Allow short-term (brief rest) replication.
7. Give the evaluee minimal visual or proprioceptive feedback regarding the results of his or her effort.

The E.R.I.C. Maximum Voluntary Effort Test Battery cm-ploys a series of instruments that are

currently available and that meet the above criteria. These instruments present tasks to the evaluee that require maximum strength efforts. Three serial measurements of each task are taken. Body position during the evaluation is kept consistent to control for muscle groups and joint angles used. Information regarding the measurement of effort made for each instrument is not revealed to the evaluee during testing. Instruments used in this battery include:

Jamar Hand Dynamometer
WEST4 UpperExtremity Strength Testing Device
BTE Work Simulator

Evaluation Procedures

All maximum voluntary effort test procedures, especially those which are isometric, have the potential to cause physical injury to the evaluee. While each procedure in the E.R.I.C, battery has been used for more than two years in not be used with evaluees who:

a. Have impaired cardiovascular or metabolic systems, even if such impairment is due only to extreme deconditioning and not to a disease process.
b. Are suspected of having cardiovascular or cerebrovascular disease or any cardi-orespiratory impairment who have not previously been cleared by a physician for participation tn maximum strength isometric tasks. Even with this clearance, these procedures should not be used without real-time monitoring and immediately available and competent emergency assistance.
c. Display the valsalva maneuver, in which the evaluee will close off the glottus and exert pressure on the diaphragm in order to increase intra-thorax pressure. This has tremendous and potentially dangerous cardiovascular repercussions.

General Guideline #2 requires that testing "must not directly involve an impaired component of the biomechanical system." This sort of testing is not only potentially injurious but it also is unreliable and will produce an un-interpretable result. Testing of an individual with, for example, carpal tunnel syndrome in the right upper extremity, therefore, would be limited to testing of the unimpaired left upper extremity. Similarly, testing of an individual with spinal disc impairment would avoid placing maximum loads on the spine and would concentrate on upper extremity function.

General Guideline #3 requires that these procedures "be under the control of the evaluee." Therefore, each evaluee must participate of his or her own free-will and not under coercion. Each evaluee must also be encouraged to limit his or her performance to a level that is acceptable to the evaluee

Reliability is at the core of the utility of an industrial rehabilitation evaluation. Without reliability, we can offer little of value. It is time that we approach the problem of the reliability of the evaluee with the same effort and success that has been mustered to solve the problem of the reliability of the test instruments.

Perhaps one day in the not ton distant future we will be able routinely to answer the query: "How do you know that he tried his best?"with: “Because I tested his effort!”