Elemental Considerations in Validity

Validity is a term that is often invoked in decisions to use neuropsychological tests. Unfortunately, the context of this use is usually negative, as when a test is cited as invalid. The use of the term implies that a test can be determined to be either valid or invalid. Of course, most clinical neuropsychologists agree that a test that is “valid” for one population may be “invalid” for another. If this is true, can a test ever be evaluated as universally valid or invalid? A second question relates to how a test is evaluated as valid or invalid. This is a question of both method (How do we evaluate a test?) and of epistemology (How do we know what we know?). Although method may be discussed separately from epistemology, the obverse is not necessarily true. That is, how we know something is highly related to how we investigate that something. This chapter discusses general issues in the relationship between epistemology and method, and Chapter 5 discusses the methodological issues more directly.

Historically, validity in clinical neuropsychological research has involved either the demonstration that scores derived from a test can accurately separate neurologically impaired individuals from unimpaired individuals or the demonstration of a statistical relationship between scores on a neuropsychological test and the results of a medical neuro-diagnostic procedure such as postmortem surgical investigations or CT scans. We say that we know the validity of a test by systematic, empirical investigation. Limiting neuropsychological validity studies to these variables was the result of the questions posed to the neuropsychologist in the clinical setting. Clinical neuropsychological assessment did not have its own canon of methods or its own set of mature scientific principles. To a large extent, clinical neuropsychology still does not have these. However, along with the development of clinical neuropsychology as a form of behavioral science with unique training requirements and professional identity has come a growth in methods that, although not unique in principle, are unique in application.

These developments have made necessary the examination of the concepts of both reliability and validity as applied to clinical neuropsychological assessment. The methods for investigating these concepts are formed partly by the nascent body of neuropsychological assumptions and principles and partly by the changing questions that are posed to the neuropsychologist in the clinical setting. Instead of being asked to localize the site of a lesion, clinical neuropsychologists are being asked to predict the limits of the behavior of a patient in the open environment or to determine whether a substantial change in skill level has occurred as the result of applications of a rehabilitation strategy.

Earlier, we stated that we know the validity of a test by empirical observations. How-
ever, there is a leap that needs to be made before statements can be made regarding the validity of a test. That leap is from the specific results of a particular procedure to statements regarding the test used as part of the procedure. The investigation is actually an evaluation of the conclusions drawn from the use of the test. These conclusions may relate to localization or to prediction issues, but they always depend on the procedure used and the context in which the procedures are used. These issues are usually discussed in terms of internal and external validity, and they are applied to the interpretations of the results of empirical investigations. Threats to internal validity are presented by those events or processes that cast doubt on the reasonableness of the conclusions drawn. Threats to external validity are presented by those events or processes that cast doubt on the generalizability of the results to other populations. These terms may be easily applied to the investigations of neuropsychological tests and may also be appropriate to discussions of the conclusions drawn from the use of these tests in a clinical situation. It may be misleading to speak of a test as valid or invalid when our research actually investigates specific hypotheses.

There is yet another consideration linking validity to method. In discussing personality constructs, Fiske (1971) argued that there was too much variance in the results when constructs are measured by different methods. Instead, Fiske proposed that the unit of analysis be the construct-operation unit. Huba and Hamilton (1976) replied that there was too much convergence among the data to support such a notion. Even though different instruments give slightly different results, they seem to share a central construct, as demonstrated by covariation among the instruments. Huba and Hamilton implicitly suggested that the best way to measure a construct is through multioperationalization; however, they did not suggest a way to concatenate the data into a single index. Fiske (1976) replied that the presence of even small variations in relationships among different methods of measuring constructs indicates the need to include the method as an integral part of the measuring unit.

Not all of these arguments may be applied to clinical neuropsychological assessment; however, parts of the arguments are very pertinent to the present discussion. Clinicians are familiar with the pattern of results when a patient performs well on a test of verbal memory but not on a test of visual memory. Alternately, the patient may perform well on a test of recognition memory, but not on a test of free recall. When these results occur, clinical neuropsychologists do not generally throw up their hands and conclude that the results are due to method variance but that the construct is singular. Instead, more than one construct is used to explain the pattern of results.

Clinical neuropsychologists often attempt to delineate the actual skill or ability that is deficient by presenting a task to a patient under different conditions. A useful method for conceptualizing this set of relationships is to consider aspects of the stimulus (e.g., the sensory modality used and the potency of the stimulus compared to other stimuli in the environment), aspects of the processing required to perform the task (mental arithmetic vs. the use of paper-and-pencil or verbal encoding vs. abstract visual encoding), and aspects of the response (motoric, verbal, and recognition). In this way, we arrive at an assessment of the ability of the individual to copy abstract line drawings and not an assessment of the construct of construction apraxia. The construct-operation unit may be specified by means of the three aspects of the behavior requested: stimulus, processing, and response. Validity investigations can then be aimed at the evaluation of the conclusions drawn from the results of a specific test procedure.