Challenging expert testimony on intelligence and mental retardation

BY I. BRUCE FRUMKIN, PH.D., ABFP

An accurate assessment of a defendant's intellectual functioning is oftentimes needed to help in the assessment of a variety of psycholegal issues. Psychologists' evaluations and testimony are sometimes based upon misinterpretation of data. A host of factors may not be considered that could influence how the psychologist, and ultimately the court, view a defendant's intellectual capabilities. This article will discuss definitions of intelligence and mental retardation, tests of adaptive functioning, the Flynn and practice effects, abbreviated scales of intelligence, crosscultural issues in intellectual assessment, and malingering of cognitive abilities.

An accurate assessment of a defendant's intellectual functioning is oftentimes needed to help in the assessment of competency to stand trial, criminal responsibility, Miranda comprehension and appreciation, mitigating issues in sentencing, and now, based upon *Atkins v. Virginia* (2002) decision, mental retardation. In *Atkins*, the United States Supreme Court reversed its decision in *Penry v. Lynaugh*

(1989) and held that it is unconstitutional to execute the mentally retarded.

The formal assessment of an individual's intelligence is an important skill psychologists master. Nonetheless, these skills are often taught while in graduate school. Many psychologists' evaluations are not based upon the most current research in the area, or their assessments, while appropriate for clinical interventions, leave them with a vulnerability in sufficiently backing up opinions in a courtroom.

This article will provide attorneys with useful information regarding the concept of mental retardation, IQ and adaptive behavior testing, cultural and ethnic considerations, as well as common mistakes clinicians make when evaluating a defendant's intellectual functioning.

Definitions of intelligence and IQ tests

Before defining mental retardation, one should explore the concept of *intelligence*. What attributes would a person need to have high intelligence? Often people believe that good vocabulary levels, reasoning skills, judgement, memory, etc. are characteristics of smart people. These all fall into the category of verbal-related skills. There are also non-verbal skills, such as perceptual organization, visual-motor coordination, and abstract visual reasoning. Good IQ tests measure both verbal and non-verbal domains across a variety of dimensions. The tests are individually administered, standardized (administered, score, and interpreted the same way), and scores are compared to normative group (usually a cross-section of the United States population). Wechsler (1944) defined "intelligence" as "the capacity to act purposely, to think rationally, and to deal effectively with his environment."

The most commonly used IQ tests are the Wechsler intelligence scales. There are three of these tests. The Wechsler Adult Intelligence Scale-III (WAIS-III) is for use with individuals ages 16 through 89. The Wechsler Intelligence Scale for Children-IV (WISC-IV) is for use with individuals ages 6 through 16. The Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III) is for use with individuals ages 2 through 7.

On the WAIS-III, there are 13 subtests that measure attributes of intelligence across many dimensions. Seven of these subtests fall into the category of Verbal Subtests and six of these subtests fall into the category of Performance or nonverbal subtests. From these subtests, one obtains a Verbal IQ, a Performance IQ, and a Full Scale IQ score. Intelligence is measured across a variety of skills, aptitudes, and abilities. Nevertheless, IQ tests do not measure everything. IQ tests do not measure creativity, leadership, "street smarts," or other very important areas. To attempt to define "intelligence" is to embark on a fruitless philosophical endeavor full of tautological reasoning.

Definitions of mental retardation

Mental retardation is defined in part by one's IQ scores. If a clinician, when asked why a defendant has an extremely low IQ score, responds by saying it is because of mental retardation, that clinician has engaged in circular reasoning. The causative factor of very low IQ is never mental retardation. The diagnosis of mental retardation is based upon specific criteria, not etiological considerations.

If this gets confusing, that is the point. Clinicians oftentimes confuse diagnoses with what causes the mental disorder or defect. Mental retardation, as defined by criteria espoused by the American Association on Mental Retardation (AAMR) and used in the *Atkins* decision, or the Diagnostic and

Statistical Manual of Mental Disorders-4th Edition, Text Revision (DSM-IV-TR) (American Psychiatric Association, 2000) as used by psychologists and psychiatrists, make the concept of mental retardation a statistical one. Arbitrarily, mental retardation is defined in part as "IQ of approximately 70 or below on an individually administered IQ test (DSM-IV-TR)." An average or mean IQ is 100 with a standard deviation of 15. An IQ score of 70 corresponds to the lower 2% range of intellectual functioning for that person's age group, or two standard deviations below the mean. If one takes into consideration measurement error (with which every test has), then IQ is never an absolute score. Rather, IQ should be interpreted within a range of scores. There is built in error in any type of measurement. For example, there can be slight differences in scoring between clinicians, varying levels of motivation and/or fatigue within subjects, or there can be differences in the quality of the testing environment. Thus, there is a 95 percent chance that any particular "true" Full Scale IQ score falls approximately within five points in either direction of the obtained score.

As pointed out, a decision was made to classify IQs in the mentally retarded range as being from those in the lower two to three percentile range of functioning. Using the DSM-IV-TR criteria, three thresholds must be met. First, there must be an IQ of approximately 70 or below. Second, concurrent deficits in present adaptive functioning in two areas (communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health, and safety) must be present. Third, the onset of the deficits in intelligence and adaptive functioning must have occurred before age 18.

In addition, the DSM-IV criteria specify types of mental retardation. Mild Mental Retardation consists of an IQ level of 50-55 to approximately 70, Moderate Mental Retardation consists of an IQ level of 35-40 to 50-55, Severe Mental Retardation consists of an IQ level of 20-25 to 35-40, and

Profound Mental Retardation consists of an IQ level below 20-25. Someone with Mild Mental Retardation is not functioning at a mild range of impairment. That individual is functioning at the lowest 2 to 3 percent of the population. An individual with Moderate Mental Retardation is functioning at the lowest one-tenth of one percentile range, 999 out of a 1000 people will be brighter than that person. I doubt that there will ever be a prosecution of an individual functioning whose valid functioning is lower than Moderate Mental Retardation. Almost certainly that person would have been institutionalized and the level of functioning below that of a five-year old.

In *Atkins*, the AAMR's 1992 definition of mental retardation was used, which consisted of ". . . significantly subaverage intellectual functioning, existing concurrently with related limitations in two or more . . . adaptive skill areas . . . manifests before age 18." Note that the AAMR definition does not give approximate IQ scores. In 2002, AAMR revised the 1992 definition to incorporate social and cultural influences in the individual's functioning, a ". . . multidimensional and ecological approach that reflects the interaction of the individual with the environment, and the outcomes of that interaction with regards to independence, relationships, societal contributions, participation in school and community, and personal well-being."

What are the legal implications of a defendant obtaining a Verbal IQ score in the Low Average range (such as 80), a Performance IQ score in the Extremely Low range (such as 60), and a Full Scale IQ score which would fall into the range of mental retardation (such as 70)? Why did *Atkins* create an automatic prohibition from executing individuals with mental retardation? Well, the United States Supreme Court said that those individuals were more likely to not understand the legal process, more likely not to understand *Miranda* warnings, thereby invalidly waiving their rights, more likely to falsely confess and be susceptible to coercion, and more likely to

have issues related to criminal responsibility at the time of the commission of the offense. Therefore, they should not be held as responsible as higher functioning defendants and should not be given the ultimate penalty, that of death! All of the Court's concerns regarding the functioning of those with mental retardation have to do with "Verbal" skills. Performance IQ has little relevancy to these perceived limitations. One can make an argument that if a defendant's Verbal IQ does not fall in the mentally retarded range, then he should not require the special protection enunciated in *Atkins*. Clinically, it is Verbal IQ or Verbal Comprehension that is relevant to the *Atkins* decision, not the Full Scale IQ score.

Tests of adaptive functioning

Apart from IQ determination, concurrent deficits in adaptive functioning must be objectively assessed. One of the major pitfalls in the evaluation of mental retardation is the inappropriate assessment of adaptive functioning. Clinicians must be aware of the limitations of these assessments. There exist standardized tests developed to measure deficits in adaptive functioning. The most commonly used test is the Vineland Adaptive Behavior Scales-II (Vineland-II). It was revised in early 2005. A second test is the Adaptive Behavior Scale-Residential and Community (ABS-SC:2). Both tests consist of a semi-structured interview of a family member or other party who knows the subject well. Questions are asked about adaptive functioning in a variety of areas and responses are scored. Scores on the various scales can then be compared to a normative sample of other people the same age, education, and circumstance.

The Adaptive Behavior Assessment System-2nd Edition (ABAS-II) is touted by its publisher as the only instrument that incorporates the language of AAMR's 2002 definition of mental retardation. Unlike the Vineland and the ABS, the test

consists of checklists which can be given to the client, parent, teacher, or caregiver.

Since one cannot make a diagnosis of mental retardation relying on IQ scores alone, an assessment of present adaptive functioning must be made. The use of psychological tests is the most objective method of evaluating such concurrent deficits in adaptive functioning.

The most common mistake forensic clinicians make in these assessments is to evaluate adaptive functioning by giving the Vineland or ABS-SC:2 tests to the defendant rather than to a family member or individual who knows the defendant well. Such administration procedures invalidate the testing for a variety to reasons in addition to violating the instructions for test administration contained in the tests' manuals.

Nevertheless, there are serious unfortunate complications with administering the test to family members or others as well. First, mental retardation is diagnosed based upon *present* deficits in adaptive functioning. If a defendant has been incarcerated for a number of years and that person's functioning has improved, the informant is providing responses based upon knowledge of the subject years earlier.

Second, many of the behaviors questioned about in the tests may be impossible for a defendant to have an opportunity to engage in while incarcerated, such as being able to ride safely and independently using public transportation, preparing one's own meals, handling money, etc. These are not relevant behaviors to be measured with someone incarcerated although they constitute a major portion of the test.

Third, who are the informants and how reliable are they in describing the defendant? There has been no research on how accurate family members or others are in describing deficits. Are certain individuals more accurate in relating history than others? Also, how biased are these family members in

providing accurate answers regarding deficits? Might the family member exaggerate the deficits so the loved will be found mentally retarded. Some clinicians may interview correctional officers. They may also be biased in a different direction or may not really have had the opportunity to observe closely the defendant to accurately answer many of the questions. It is unlikely they are willing to spend 30 minutes to an hour being interviewed about a detainee.

A fourth problem is what is called in the psychology profession, a *sampling* problem. There have been no norms developed to compare a subject's scores with a correctional population. While the Vineland and ABS-RC:2 appear to meet the *Frye* (1923) standard, they are commonly used and are really the only tests available to objectively measure adaptive functioning, it is doubtful these tests would meet *Daubert* (1993) criteria.

The ABAS-II suffers from the same deficiencies as the Vineland and ABS-RC:2 with several additional problems. The checklist format is even more susceptible to response distortion or malingering, particularly when given to the defendant himself. Finally, since the test consists of a checklist, it is more appropriate to look at the results as more of a screening assessment, rather than a comprehensive assessment of adaptive functioning.

Practice effects and the Flynn effect

Scores on the WAIS-III should remain fairly stable over a period of time. Nevertheless, *practice effects* (significant gains in IQ between testings) have been well documented, particularly for Performance IQ (e.g., Kaufman & Lichtenberger, 1999). Verbal IQ scores increase two to three points while Performance IQ scores increase nine to ten points. Full Scale IQ scores increase six to seven points. The reason for the larger increases in Performance IQ is that the

tasks are no longer novel and subjects remember the strategy they used to solve the problems involving blocks, story cards, abstract symbols, etc. If one retests a subject, particularly within six months of the first administration, one should anticipate practice effects. Clinicians may erroneously over-interpret large differences between Performance and Verbal IQ without taking these effects into consideration. There is an age-related pattern in practice effects. The size of the practice effect decreases with age, particularly with those individuals over 75 years old.

Oftentimes attorneys, in wanting a client to be found mentally retarded, rely on IQ testing done long ago when the defendant was in school. The recently hired psychologist is told not to repeat IQ tests because the first test showed IQ levels in the mental retardation range and the defense attorney wants to rely on that score. This is problematic for a number of reasons. First, oftentimes the training and expertise of the school psychologist is not known. Many school systems rely on master's level clinicians to do their psychological evaluations. The clinicians may or may not have the skills of the doctoral level psychologist. Second, many times the raw psychological test data is unavailable from the school. It is difficult for the psychologist to determine why a subject's IQ score was low. A low score can be due to poor effort, improper test administration, or errors in scoring. It is also unlikely the school psychologist would ever administer tests to assess malingering or lack of effort. Also, poorly validated tests may be used in the schools in an attempt to better measure intelligence. There may be more language and cultural issues for a student, recently having immigrated to the United States, than that same individual who is now being retested as an adult.

Psychologists also need to be aware of the *Flynn effect* (Flynn, 1987). As a group, individuals' IQs increase approximately 1/3 point each year. Thus it is important for a psychologist to use the most updated version of an IQ test. A

15-year-old student tested in 2002 would have been administered the WISC-III, which was published in 1991. Thus the IQ score on the testing done in 2002 would have overestimated IQ approximately 4 points. The IQ score reflects that one is comparing a 15-year-old subject to peers some 12 years earlier. The population becomes more intelligent each year. This is because of better nutrition, better exposure to information, better schooling, etc. Consequently, in IQ assessment, it is important to compare a person's scores with peers today. Each successive edition of an IQ test has updated normative data. Since tests are not revised yearly, the *Flynn* effect comes into play except the first year or two a revised version of an IQ test comes out in distribution. The WAIS-IV, which is the most updated version of the adult intelligence test, was published in 1997 (and the norms on which the test were based developed several years earlier). Therefore an IQ score obtained today will overestimate IQ by some three points. This is not to say that a psychologist should "adjust" the IQ score to take into consideration the Flynn effect. But it is something a psychologist should know about, particularly when making fine-line cutoffs for what scores should indicate mental retardation or particular levels of intellectual functioning.

Assessment of malingering and deception

Frumkin (2003) in an earlier publication, emphasized that psychologists who assess for mental retardation (or any cognitive deficits for that matter), must rely on third party data to help assess whether a defendant may be attempting to exaggerate or feign intellectual problems. If at all possible, interviews must be conducted with those who know the defendant well. School records and prior testing results need to be reviewed if available.

It is incumbent that the psychologist formally assess for malingering using standardized tests and procedures. Most

attorneys are familiar with the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) as a psychological test which includes special scales to assess exaggeration and minimization of psychological problems. This test, which was designed to assess for personality characteristics and psychopathology, is not the appropriate test to rule out or assess for malingering of intellectual deficits. While it is beyond the scope of this article to detail all the available tests to evaluate malingering of intellectual problems, attorneys need to be aware that the MMPI-2 is not one of them.

A number of tests exist to help assess for malingering of intellectual and memory deficits. Although tests such as the Rey 15-item Test and the Dot Counting Test are screening tests that may provide useful hypotheses regarding malingering (they both have high rates of misclassifying individuals as malingerers when they are not, or failing to classify true malingerers), clinicians should also use tests which have better validation.

Tests referred to as *forced-choice* or *symptom-validity* are ones with potentially the most promise in helping to assess malingering of cognitive deficits. These are tests in which a subject must choose between one of (usually) two choices. Typical of these tests are the Portland Digit Recognition Test (PDRT), the Test of Malingered Memory (TOMM) and the Validity Indicator Profile (VIP).

Forced-choice or symptom validity testing involves measuring a subject's performance and assessing whether the score is lower-than expected based on normative data (comparing that person's score to other similar individuals). The same test can be also be used to examine a statistically improbable failure rate. Regardless of the difficulty of the test items, even a very impaired individual should not produce scores significantly below chance levels of responding.

Although Rogers (1997) has summarized accepted definitions of malingering as "the deliberate fabrication or gross exaggeration of psychological or physical symptoms for the fulfillment of an external goal," the clinician must distinguish between malingering on these tests and suboptimal effort or poor motivation. On the TOMM, the subject is shown 50 pictures of common objects. Then the subject has 50 trials of picking one of two visual stimuli, one of which had just been presented. The individual can get a score from 0 to 50. The mean (average) score for individuals with no brain damage is 50.0 for the second administration of this test. This is an extremely easy task. Those with cognitive impairment obtain a mean score of 48.6. Only 9.6% of cognitively impaired adults get a score lower than 45. This same test can be used to look at a statistically improbable failure rate. A score of less than 17 in an individual responding randomly, will occur less than 2% of the time. Thus, it is statistically probable that a person obtaining a score less than 17 is malingering.

If a test contains rather difficult items, an impaired individual may not try as hard on the difficult items. It is not because he or she is trying to fake, but that the person gives up easily on items which require a lot of thought. Authors of "malingering tests" such as the VIP, which contain difficult items, as well as easy ones, caution use with mentally retarded individuals. Whereas most mentally retarded individuals will have "careless" or "invalid" profiles, a "malingered" profile in which easy items are missed and difficult items are responded at chance levels provides data that the individual is most likely faking or exaggerating intellectual problems.

Rogers and Bender (2003) provide the definitive review of malingering and deception although it is even more technical for non-psychologists than this article.

Abbreviated scales of intelligence

Many psychologists elect to administer an abbreviated or shortened version of an intelligence test. The WAIS-III may take up to several hours to administer to an individual. Some clinicians do not want to spend that amount of time with IQ testing for forensic purposes. Interestingly, in clinical and school settings, these same psychologists do not hesitate to give the full intelligence testing regimen. The most commonly used abbreviated scale of intelligence is the Wechsler Abbreviated Scale of Intelligence-III (WASI-III). This test is considered the best of the group of short-form intelligence tests. Nevertheless, there are serious problems with using the WASI-III and tests like it, particularly in forensic settings.

The WAIS-III contains 11 subtests used to obtain a Verbal IQ, Performance IQ, and Full Scale IQ score. There are also three supplementary scales. The WASI-III, in contrast, contains only four subtests. As noted in the WASI-III manual (Psychological Corporation, 1999), the test is to be used as a screening instrument, an estimation of general intellectual functioning for research purposes, or as a reassessment for someone who had previously been given a more comprehensive evaluation. The test is not meant to replace more comprehensive measures of intelligence. When someone's life or liberty is at stake, why would a psychologist elect to administer a screening test of intelligence?

Axelrod (2002) compared performance on the WAIS-III with performance on the WASI-III and other shortened measures of intelligence. The results of the research "suggest that clinicians should use the WASI cautiously, if at all, especially when accurate estimates of individuals' WAIS-III results are needed." One problem is that when the four subtests are used, the WASI-III scores overestimate Performance IQ by eight points and Full Scale IQ score is overestimated by three points. The Verbal IQ score is underestimated by two points.

What is worse, the WASI-III demonstrated poor accuracy in estimating true IQ scores. Whereas 68% of Full Scale IQ scores on the WAIS-III vary three points or less in either direction (the standard error of measurement), only 30% of WASI-III scores vary three points or less in either direction. That means there is a 70% chance that a true Full Scale IQ score on the WASI is more than three points in either direction of the obtained score. Moreover, on the WASI-III, there is a 34% chance that the Full Scale IQ vary more than six points in either direction and a 16% chance that the obtained Full Scale IQ score vary by ten points in either direction. Many clinicians opt to use only two subtests when administering the WASI-III, one each for Verbal and Performance IQ, thereby shortening the administration time from approximately 30 to 15 minutes. When this transpires, there is a 27% chance that the obtained Full Scale IQ score varies by more than ten points in either direction. These error rates are unacceptable.

Cross-cultural factors in intelligence testing

There are many limitations in assessing intellectual functioning in cultural and linguistic minorities. Even assuming the defendant speaks English, language limitations may nonetheless exist. Some words cannot be translated exactly or may have different meanings, values, or frequency in the defendant's cultural milieu.

Individuals vary across cultures in how they perform on intelligence tests. It is not that a particular cultural group is less intelligent than the mainstream culture. Rather, a cultural group may have less experience or practice with paper and pencil tests, computer tasks, even testing in general. One cannot assume what is normal or average in one culture is abnormal or deficient in another culture. Motivational differences also abound. Japanese children outperform Caucasian American children on most intelligence measures.

Few would make the argument that Japanese are genetically superior intellectually than Caucasian American counterparts. The issue is that in the Japanese culture, academics and test taking are so much more instilled in children at a young age. Failure is not an option. There is a different motivational set in how one manages test taking. Also, there are different interpersonal expectations in how individuals react to testing. Shyness, lack of assertiveness, or "not guessing" at answers on measures of intelligence may be considered a "value" in cultures different than American mainstream. There are also interactive effects between the ethnicity of the examiner and that of the subject. The meaningfulness of the test stimuli can also be a factor. So what is a psychologist to do?

Some psychologists use the standardized English-version IQ tests and translate that test themselves simultaneously or on-the-spot if they speak the language of the defendant. This procedure is the least desirable of any attempt to measure IQ in linguistically different populations. First the test has not been standardized in this fashion. Second, words cannot be translated with exact equivalency and frequency in the other language. For example, in the Vocabulary subtest of the WAIS-III, subjects are shown and read words in which they must state what the word means. One of the easier words to define in English may turn out to be a more difficult word in that other language. Third, translating the word in Spanish, for example, may provide the definition if no common synonyms exist for that word in Spanish. Fourth, some tasks may prove easier or more difficult when translated. For example, Digit Span, which requires one to repeat digits just read, might be a more difficult task in Spanish than in English. Most digits, one through nine, are multisyllabic in Spanish. In English, the same digits are generally monosyllabic. It may be more difficult to remember digits that take longer to rehearse in memory because they contain more syllables. Fifth, although the WAIS-III has more culturally-appropriate items than its predecessors, some questions may be more difficult or have different meanings to someone from

a different culture. Asking a teenager who had lived in a rural area of Honduras prior to arriving in the United States two months earlier, "What is the thing to do if you find a gift-wrapped package laying on the floor of a grocery store?," the "correct" answer may differ than if you asked that same question to an American youth.

The Escala de Inteligencia Wechsler Para Adultos (EIWA) is the only standardized "Spanish version" of the Wechsler intelligence tests for adults in this country. The lack of appropriate test instruments to assess Spanish-speaking adults' intellectual functioning is quite problematic. Various countries have adapted the WAIS-III for their own individual use. These adaptations generally have country-specific normative data in which to compare scores. Yet there has been little research validating these tests, have generally not been subjected to peer review, and do not pertain to these ethnic minorities residing in the United States.

Finally a true Spanish version of the WISC-IV (for children ages 6 to 16), the Wechsler Intelligence Scale for Children-IV-Spanish (WISC-IV-Spanish) was published in the last year. This test is much needed in that it is more than just a linguistic translation of the English version. Items have been adapted to make them relevant to Spanish-speaking children residing in the United States from various countries of origin, including Mexico, Cuba, Puerto Rico, Dominican Republic, and other countries in Central and South America. One can compare scores with other Spanish-speaking children of like age and United States educational experience and that of their parents, as well as with English-speaking examinees. This type of test is needed for adults.

So what we have left is the EIWA. This test was developed using normative data from a rural Puerto Rican population in the early 1960's. You would expect, based upon the *Flynn* effect alone, that the IQ scores would be greatly overestimated. In fact, this is true. It is well documented (e.g.

Lopez & Romero,1988; Melendez, 1994) that scores on the EIWA overestimate IQ by at least 20 points. Therefore, an individual's Full Scale IQ score of 85 on the EIWA is equivalent to a score of 65 or so. Although the clinician should not adjust or change the IQ score on the EIWA by subtracting 20 or more points, this is something that must be taken into consideration when interpreting IQ scores on the EIWA. The EIWA is useful for several different reasons although cautions are suggested in interpretation. First, one can compare strengths and weaknesses in various cognitive areas. One can see whether the defendant has better verbal versus nonverbal skills. One also has behavioral data regarding the actual responses or behavior given during the various tasks. Thus, a response to a test item such as "a pencil and pen are both alike in that they are both brown in color" provides data regarding this individual's concrete level of verbal abstract reasoning. Using the EIWA, a clinician obtains scores and behaviors in a standardized, structured fashion.

A good practice is for the clinician who administers the EIWA to also administer one or more nonverbal measures of intelligence. Although tests such as the Test of Nonverbal Intelligence-III (TONI-III) and Raven's Progressive Matrices provide data on nonverbal intelligence or abilities, these tests are also not totally devoid of the influence of culture. Also, the nonverbal tests are not assessing for the verbal cognitive skills needed to be competent to stand trial, waive Miranda rights, or any of the other myriad legal capacities usually required.

For English-speaking adults from various other countries who score low on IQ tests, the clinician should not be allowed to minimize a low score's importance because of cultural factors. This is the defendant's functioning when compared to people in the United States. Although he may not be impaired when compared to people of his country of origin, most forensic questions pertain to capacities or abilities compared

to those in the United States. Thus, a defendant whose trial competency is at issue, must be compared to the minimal level of competency required of American defendants, not to those from another country.

LaCalle (1987) and to a lesser extent, Frumkin & Friedland (1995), enumerate problems associated with the use of interpreters. Once a decision is made to use an interpreter, several factors need to be considered. An interpreter must be professionally-trained and certified or registered. Except in the rarest of situations, interpreters should not be friends, employers, or family members of the defendant. The interpreter should also not be one who is an employee of the public defender or state attorney's office. The former group may not have the training to adequately provide translation services and may consciously or unconsciously translate to the benefit of the defendant, whereas the latter group may give the appearance of bias when translating. Interpreters should be told to translate every utterance, even if does not make much sense. A clinician oftentimes evaluates by not only analyzing the content of the communication, but how the content was communicated, assessing subtleties and peculiarities in speech. This can be lost if the interpreter paraphrases what the defendant says to make it more understandable. The interpreter should also be asked to provide simultaneous translation, not consecutive translation. This means that the translation should take place as the party is speaking, not after the party is finished speaking.

Often a dilemma occurs whether a referral is to be made to a psychologist who is culturally and linguistically proficient in that culture and language of the defendant. Depending upon the country of origin, there may be few if any psychologists available. Even if one is located, is that clinician a forensic specialist who has the knowledge and experience to address the specific psycholegal issue. Sometimes it is preferable to hire a psychologist who is a specialist in the particular psycholegal issue to be addressed and for that clinician to use

an interpreter. It is generally recommended that the psychologist hire or work in conjunction with a psychologist who speaks the language of the defendant. Therefore certain psychological tests can be administered by the linguistically-proficient psychologist and an interview is conducted by that person to insure the English-speaking psychologist has not missed out on any subtleties of behavior that was not apparent when an interpreter was used. Moreover, the psychologist should consult or hire a cultural consultant versed in the particular culture of the defendant. It is not enough to consult with someone from the country of origin. That person should have knowledge of the specific cultural group within that country. An individual who is from a rural mountain village in the Peruvian Andes is quite different from a Peruvian banker from Lima.

Concluding remarks

The above summarizes the information a testifying psychologist should possess regarding "intelligence" and "intelligence testing." This article is meant as a guide to help attorneys better understand expert testimony on intellectual functioning. Nevertheless, it is essential a lawyer hire a consulting psychologist to review the expert's work product. Attorneys are not trained to interpret raw psychological test data and in some jurisdictions, are not allowed access to them. A qualified psychologist is capable of addressing errors in scoring and administration of the psychological tests and should be retained.

Over 80 years ago, Boring (1923) attempted to crystallize an issue hotly debated among psychologists of the time, "What is intelligence?" His somewhat sarcastic tautology to a lay audience was ". . . intelligence as a measurable capacity must at the start be defined as the capacity to do well in an intelligence test. Intelligence is what the tests test." Yet even today, experts frequently use a tautology when defining

concepts such as mental retardation. Intelligence is not a single construct and an IQ score has many meanings. Oftentimes the lawyer forces the clinician into stating an opinion in dichotomous terms and tries to finagle a yes or no response from what may have no simple answer. Nevertheless, testimony on psychological issues, when not presented in a misleading fashion, educates the trier of fact, and is a valuable resource.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders-Text Revision*, Washington, DC: American Psychiatric Association.

Atkins v. Virginia, 260 Va. 375, 534 S.E. 2d 312, reversed and remanded (U.S. Supreme Court, No. 00-8452, decided July 21, 2002).

Axelrod, B. (2002). Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms estimating intellectual functioning, *Assessment, 9*, 17-23.

Boring, E. (1923). Intelligence as the tests test it, *The New Republic, June 6, 1923*, 35-37.

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

Flynn, J. (1987). Massive IQ gains in 14 nations: What IQ tests really measure, *Psychological Bulletin, 101*, 171-191.

Frumkin, I.B. (2003). Mental retardation and the law: Common misconceptions, *Cornerstone, 25, #2*, Washington, DC: National Legal Aid and Defender Association.

Frumkin, I.B. & Friedland, J. (1995). Forensic evaluations in immigration cases: Evolving issues, *Behavioral Sciences and the Law, 13*, 477-489.

Frye v. United States, 54 App.D.C., 46, 47, 293 F. 1013, 1014 (1923).

Kaufman, A. & Lichtenberger, E. (1999). *Essentials of WAIS-III Assessment*. New York, NY: John Wiley & Sons.

LaCalle, J. (1987). Forensic psychological evaluations through an interpreter: Legal and ethical issues, *American Journal of Forensic Psychology, 5*, 29-43.

Lopez, S. & Romero, A. (1988). Assessing the intellectual functioning of Spanish-speaking adults: Comparison of the EIWA and the WAIS, *Professional Psychology: Research and Practice, 19*, 263-270.

Melendez, F. (1994). The Spanish version of the WAIS: Some ethical considerations. *The Clinical Neuropsychologist, 8,* 388-393.

Penry v. Lynaugh, 492 U.S. 302 (1989).

Psychological Corporation, (1999). *Wechsler Abbreviated Scale of Intelligence (WASI) Manual*, San Antonio, TX.

Rogers, R. (Ed.), (1997). *Clinical Assessment of Malingering and Deception, 2d ed.).* New York, NY: Guilford Press.

Rogers, R. & Bender, S. (2003). Evaluation of malingering and deception, *Handbook of Psychology, Vol. 11-Forensic Psychology)* A. Goldstein (ed.), New York, NY: John Wiley & Sons.

Wechsler, D. (1944). *The Measurement of Adult Intelligence, 3d ed.*, Baltimore, MD: Williams & Wilkens.