CHALLENGES IN IMPLEMENTING THE *ATKINS* DECISION

J. Gregory Olley, Ph.D.

The states' interpretation of the Atkins v. Virginia *(1) decision has led to problems of implementation as noted by Widaman and Siperstein (2). In addition to the problems of subjective modification of adaptive behavior scores to reflect cultural differences, which these authors noted, expert witnesses have encountered many other challenges. Although definitions and clinical diagnostic criteria are nearly uniform, these criteria can be difficult to apply in* Atkins *cases. The problem of determining a diagnosis in childhood and at the time of the crime when many years have passed challenges the usefulness of some common clinical methods. The interpretation of intelligence test scores typically takes into account the standard error of measurement, and this source of error is greater in people with low IQ. Score interpretation is further complicated by the Flynn effect. Nevertheless, courts may interpret IQ scores rigidly. Adaptive behavior must be assessed based on knowledge of community functioning, which can be difficult to obtain for incarcerated individuals. Definitions of mental retardation overlap with definitions of other disorders, and dual diagnoses are common in people with mental retardation. Available methods of assessing malingering mental retardation lack adequate validity. A satisfactory process for the diagnosis of mental retardation cannot be assured until these problems are resolved.*

Two recent articles in this Journal by Denkowski and Denkowski (3, 4) have provided good illustrations of the challenges and disagreements in implementing the Supreme Court's 2002 *Atkins v. Virginia* decision. Disability advocates universally praised the decision that prohibited the death penalty for individuals with mental retardation (often more preferably called intellectual disability), but they did not anticipate the difficulties in translating customary clinical standards of diagnosis to the courtroom. The response to Denkowski and Denkowski's (4) recommendations for assessing adaptive behavior in *Atkins* cases (2) highlights only a few of several serious challenges that have resulted from the Supreme Court's decision to leave the implementation of the *Atkins* decision to the states. This article provides an

overview of several other problems in the role of psychologists and other experts in *Atkins* cases.

Whether it takes place pretrial or as part of the appeal of a death sentence, the determination of mental retardation under *Atkins* typically is done in a hearing before a judge, although in some circumstances a jury may make the decision. The testimony of expert witnesses in *Atkins* cases typically begins with a description of the contemporary definitions of mental retardation (5-9) and their similarities among each other and to the applicable state definition.

Most states use definitions very similar to that of the American Association on Intellectual and Developmental Disabilities (AAIDD) (formerly American Association on Mental Retardation); however, a few states have developed idiosyncratic definitions. Recently the AAIDD and the Arc of the United States wrote an *amicus* brief to the United States Supreme Court in opposition to the Texas definition, which they argued attempted to minimize the protections of the *Atkins* decision by redefining mental retardation (10). Definitions by professional organizations were developed for diagnosis in clinical settings, so their application in *Atkins* cases requires some adaptation, and that is where the problems begin. These problems fall into several categories.

RETROSPECTIVE EVALUATION

The clinical diagnosis of mental retardation typically takes place for the first time in childhood, because by definition, the disorder begins during the developmental period, and that is the time that developmental problems are usually first noticed. However, in capital cases, the defendant is an adult. The definition requires significantly impaired functioning in childhood, and the *Atkins* decision referred to functioning at the time of the crime. The psychologist must determine whether functioning was significantly impaired in childhood and at the time of the crime, which may be as little as a few months or as much as 20 or more years earlier. The diagnosis of mental retardation in *Atkins* hearings is, thus, always retrospective.

Childhood Origin

Although all contemporary definitions of mental retardation require that the disorder originate in childhood, the retrospective nature of *Atkins* pro-

ceedings often makes the documentation of that requirement difficult. Widaman and Siperstein (2) discussed the importance of school records. Patton and Keyes (11) and Olley and Cox (12) also reviewed the value of school records and some of the difficulties in interpreting them. Reschly (13) provided a good overview of the challenges in obtaining retrospective evidence for childhood onset.

Information from school records should be supplemented with interviews with educators who knew the defendant as a student. If these individuals are available, they may provide information on school policies and procedures at the time that the defendant was in their school and interpret school records that may be unclear.

Other records may also be valuable. Medical, mental health, social services, or other records should be sought. Defendants who received Social Security benefits because of a disability should have records that document the disability (14). Although such records are helpful in demonstrating early onset, they are often missing or misleading; therefore, the absence of records or test scores cannot be used to disprove the diagnosis of mental retardation.

Without such records the psychologist may rely on the memories of individuals who knew the defendant at the relevant times, or he/she may test the defendant in prison and interpret the findings to apply to earlier years. Both approaches have their limitations. Olley and Cox (12) described methods for gathering adaptive behavior information from varied sources, including the retrospective administration of adaptive behavior rating scales, such as the Adaptive Behavior Assessment System (second edition) (15).

Adaptive Behavior

The defendants in *Atkins* proceedings are the diagnostic "close calls." They are likely to have mixed evidence for a diagnosis, which makes the assessment of adaptive behavior especially important. Such individuals with mild mental retardation or borderline intellectual functioning do not have the obvious stigmata associated with more severe mental retardation and do not fit the public's stereotype of mental retardation (16). Thus, information on their everyday functioning may be the critical diagnostic information. Widaman and Siperstein (2) have addressed the challenges in the valid assessment of adaptive behavior. Their critique of Denkowski and Denkowski (4)

addressed problems of subjectivity in assessment and the comparison of the defendant's functioning to a restricted population, rather than to the general population of the United States.

The assessment of adaptive functioning also means that many psychological or neuropsychological tests that correlate poorly with everyday functioning are not valid for this purpose. Further, knowledge is not the same as performance. Tests of knowledge that involve answering questions or pointing to pictures are poor indicators of adaptive functioning. Denkowski and Denkowski (17) reviewed the problems in the use of one such test, the Street Survival Skills Questionnaire (18).

DISAGREEMENTS REGARDING THE NATURE OF MENTAL RETARDATION

Differences in the testimony of experts and the arguments of opposing attorneys in *Atkins* hearings often can be reduced to two issues regarding the nature of mental retardation. The first issue is whether the defendant's deficits are significant enough to merit this diagnosis, and the second is whether the observed deficits indicate mental retardation or some other condition. These differences are usually argued based on available information about adaptive behavior and include the problems mentioned in the previous section. Some of the differences focus on the validity of the tests or other assessment methods used, but the fundamental difference is in the way that intellectual disabilities are defined and conceptualized. Again, the chief problems lie in the translation of the clinical definition of mental retardation to the courtroom.

Is the Limitation Severe Enough?

Ellis (10) described these problems very well in the *amicus* brief to the U.S. Supreme Court in the case of *Briseno v. Quarterman* (10). With regard to the first issue, the severity of the impairment is measured by IQ tests and by adaptive functioning. The measurement of IQ has generated its own disagreements to be addressed later in this article. Most of the conflicting testimony on this subject concerns the role of adaptive behavior. As noted earlier, the AAIDD definition (5, 6) clearly stated that the diagnosis of mental retardation requires significant impairment in the skills of everyday living in one's community, and every individual has a different pattern of abilities and limitations. In other words, the contemporary definitions emphasize what the

individual *does not or cannot do*. Arguments for the prosecution often put the emphasis on what the defendant *can* do, rather than on deficits. Thus, the prosecution commonly argues that because the defendant was known to do certain things (e.g., have a driver's license, get married, or be employed), he cannot have mental retardation.

This approach highlights specific instances of competent performance and argues that the selected examples rule out mental retardation. The most common version of this approach is to argue that the defendant's role in the crime involved such high cognitive functioning that it is incompatible with mental retardation. Ellis (10) addressed this issue in detail, noted the varied skills of people with mental retardation, and described the use of isolated examples as "unsupported by, and, totally at odds with, the well accepted clinical understanding of mental retardation" (p. 12).

A close variant of this issue is the difference between what an individual does and his or her potential. Definitions of adaptive behavior emphasize performance; however, it is often argued by the prosecution that the individual could have done more or had the potential to do more. In other words, he simply chose not to pay his bills, maintain a bank account, or participate in community activities. To make the contrasting point, experts arguing for the diagnosis of mental retardation note that adaptive behavior is defined as the individual's *typical* performance without reference to potential (5, 6). Of course, it is true that some of the behaviors that evidence poor adaptive functioning are carried out by choice. Making poor choices is quite characteristic of mental retardation (19).

This argument is often extended to raise the issues of cultural relativity that the Denkowskis argued (4). Mental retardation or intellectual disability is defined as a significant impairment relative to the general population. However, Denkowski and Denkowski (4) and others have argued that if certain adaptive deficits are common in the defendant's cultural group, they are not evidence for mental retardation. This argument contrasts the definition of the disorder as impairment relative to the general population or relative to a smaller group.

Is It Mental Retardation or Something Else?

Mental retardation can coexist with other disabilities. In fact, the prevalence of mental illness in people with intellectual disabilities is several times

higher than in the general population (20). Despite the well-known prevalence of dual diagnoses, the argument against the mental retardation diagnosis commonly takes the position that the documented limitations in adaptive behavior are indications of a mental illness (e.g., antisocial personality disorder) and not mental retardation. Antisocial personality disorder and mental retardation can, of course, coexist, and their diagnostic criteria overlap, for instance, "impulsivity or failure to plan ahead," "consistent irresponsibility, as indicated by failure to sustain consistent work behavior or honor financial obligations" (7, p. 706).

Another variation of this argument against the diagnosis of mental retardation is, "There is impairment, but it isn't caused by mental retardation." This position fails in at least two ways. First, the impairment is not caused by mental retardation. Mental retardation is simply the name of the impairment. Second, the argument fails to recognize that definitions of mental retardation do not require that a cause be identified. Mental retardation is diagnosed by low intelligence associated with low adaptive behavior, regardless of cause. Although hundreds of factors are known to cause this disorder, people with mild mental retardation (with IQs between roughly 55 and 70) and the "close call" cases that come to *Atkins* hearings seldom have an identifiable cause for their limitations.

DISAGREEMENTS REGARDING THE MEASUREMENT OF INTELLIGENCE

Although the study of modern tests of intelligence dates back to Alfred Binet over 100 years ago, and the measurement of intelligence is generally regarded as more precise than the measurement of adaptive behavior, *Atkins* hearings are often punctuated by disagreements over IQ scores. There are arguments over the qualifications of the examiner, the choice of tests, and even the scoring of particular test items, but the central theme of these disagreements has to do with the precision of any given score.

Application of the Standard Error of Measurement

Scores on intelligence tests are customarily interpreted with reference to their standard error of measurement to acknowledge that there is error variance in the scores. On this basis, the American Association on Intellectual and Developmental Disabilities (5, 6) has recommended that a diagnosis of mental retardation can be made with an IQ as high as 75 if there are accom-

panying deficits in adaptive behavior. The court with jurisdiction over the case typically decides this issue. Some jurisdictions have accepted the IQ 75 cutoff, while others have recognized no flexibility in interpretation of scores.

Application of the Flynn Effect

James Flynn has documented an international finding in which the scores on IQ tests increase over time (21, 22). This increase is a major reason why IQ tests must be re-normed every 10 years or so. In *Atkins* hearings, the difference of even one point can be a matter of life or death. As a result, experts for the defense often argue that obtained IQ scores should be adjusted to take into effect the inflation of scores referred to as the Flynn effect. Denkowski and Denkowski (3) and others have argued against the application of the Flynn effect to individual scores or argued that the inflation of scores has stopped in recent years. In reply, Flynn (23, 24) has argued that although an adjustment for the Flynn effect is an approximation when applied to individual scores, an approximate adjustment is better than using a score known to be invalid. Further, Flynn (23) recently presented data to indicate that the Flynn effect is still evident in the difference between scores on the Wechsler Adult Intelligence Scale versions III and IV.

The AAIDD has supported the application of the Flynn effect (25), and some courts have accepted this position; however, there appears to be no consistency in such rulings.

Interpreting Multiple Scores

One of the reasons that *Atkins* cases are "close calls" is that there is conflicting evidence regarding a diagnosis. The 2008 publication of the WAIS-IV (26) and the 2003 publication of the Stanford-Binet Intelligence Scales (27) assures that there are two recent and well standardized tests available for adults, although one can never be confident that two or more tests scores from different tests or times are comparable.

A further source of conflicting information is the existence of multiple IQ scores with some below the cutoff for mental retardation and some above. How should these multiple scores be interpreted? One argument is that the highest score is the valid one, because "You can fake low, but you can't fake high." In other words, the lower scores are interpreted as either malingering or poor effort. A recent study by Whitaker (28) provided useful data on this

question. Whitaker identified several influences that can contribute to variability among test scores. In addition to malingering and such factors as improper test administration, Whitaker presented data demonstrating that the variability in test performance is much greater for those with low IQs, such as the defendants in *Atkins* hearings. It may be impossible to know for sure why an individual's test scores vary, but the variability is greater in people with low IQ, and one cannot assume that malingering or lack of effort are the only possible causes.

MALINGERING

Testing for an *Atkins* claim raises the possibility that the defendant is producing purposefully poor performance in order to fake mental retardation and avoid the death penalty. In light of this concern, it is important to address the possibility of malingering. Unfortunately, the existing tests that have been used to identify malingering of mental retardation have shown poor validity (29-32). This finding leaves the expert few choices. Expert opinion of whether the defendant was malingering may be based on little more than the examiner's observation of test performance. Research to develop and validate tests of malingering mental retardation is greatly needed.

In the absence of valid tests, the best evidence for or against malingering of mental retardation is a well-documented developmental history. As noted earlier, such documentation is usually very difficult to find.

CONCLUSIONS

The issues raised by Denkowski and Denkowski's (4) proposed method for adjusting adaptive behavior scores have been addressed by Widaman and Siperstein (2). However, many more challenges exist in the valid diagnosis of mental retardation in *Atkins* cases. Some of these problems will be resolved by consensus in the field; some will be resolved by research. However, as Ellis (10) so clearly pointed out, courts are making decisions throughout the country, and they often make decisions based on outdated stereotypes and information that is not consistent with the clinical or scientific knowledge of the nature of mental retardation. Psychologists who provide information to the courts in *Atkins* cases have an enormous responsibility to use the available clinical and scientific knowledge to resolve these problems. The lives of many people are at stake.

REFERENCES

1. Atkins v. Virginia, 536 U. S. 304 (2002)
2. Widaman KF, Siperstein GN: Assessing adaptive behavior of criminal defendants in capital cases: a reconsideration. American Journal of Forensic Psychology 2009; 27:2:5-32
3. Denkowski GC, Denkowski KM: WAIS-III IQs of criminal defendants with a mental retardation claim should not be reduced for the "Flynn effect." American Journal of Forensic Psychology 2007; 25:4:41-63
4. Denkowski GC, Denkowski KM: Adaptive behavior assessment of criminal defendants with a mental retardation claim. American Journal of Forensic Psychology 2008; 26:3:43-61
5. American Association on Mental Retardation: Mental Retardation: Definition, Classification and Systems of Support (9th ed.). Washington, DC, Author, 1992
6. American Association on Mental Retardation: Mental Retardation: Definition, Classification, and Systems of Support (10th ed.). Washington, DC, Author, 2002
7. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev.). Washington, DC, Author, 2000
8. Jacobson JW, Mulick JA (Eds.): Manual of Diagnosis and Professional Practice in Mental Retardation. Washington, DC, American Psychological Association, 1996
9. Jacobson JW, Mulick JA, Rojahn J (Eds.): Handbook of Intellectual and Developmental Disabilities. New York, Springer, 2007
10. Ellis JW: Amicus brief. Briseno v. Quarterman, No. 08-6631, U.S. 12/8/2008 (2008)
11. Patton JR, Keyes DW: Death penalty issues following Atkins. Exceptionality 2006; 14:237-255
12. Olley JG, Cox AW: Assessment of adaptive behavior in adult forensic cases: the use of The Adaptive Behavior Assessment System-II, in Adaptive Behavior Assessment System-II: Clinical Use and Interpretation. Edited by Oakland T, Harrison PL. San Diego, Elsevier, 2008; 381-398
13. Reschly DJ: Documenting the developmental origins of mild mental retardation in SSI eligibility and death penalty appeals. Applied Neuropsychology 2009; 16:2:124-134
14. National Research Council: Mental Retardation: Determining Eligibility for Social Security Benefits. Committee on Disability Determination for Mental Retardation. Edited by Reschly DJ, Myers TG, Hartel CR. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press. 2002
15. Harrison PL, Oakland T: Adaptive Behavior Assessment System (2nd ed.). San Antonio, Harcourt Assessment, 2003

16. Siperstein GN, Norins J, Corbin S, Shriver T: Multinational Study of Attitudes Toward Individuals with Intellectual Disabilities: General Findings and Call to Action. Washington, DC, Special Olympics, 2003
17. Denkowski GC, Denkowski KM: Misuse of the Street Survival Skills Questionnaire (SSSQ) for evaluating the adult adaptive behavior of criminal defendants with mental retardation claims. Intellectual and Developmental Disabilities 2008; 46:2:144-149
18. Linkenhoker D, McCarron L: Adaptive Behavior: The Street Survival Skills Questionnaire. Dallas, TX, McCarron-Dial Systems, 1993
19. Greenspan S: Foolish action in adults with intellectual disabilities: the forgotten problem of risk-unawareness, in International Review of Research in Mental Retardation (Vol. 36). Edited by Glidden LM. Boston, Elsevier, 2008; 147-194
20. Fletcher R, Loschen E, Stavrakaki C, First M (Eds.): Diagnostic Manual—Mental Retardation (DM-ID): A Textbook of Diagnosis of Mental Disorders in Persons with Mental Retardation. Kingston, NY, NADD Press, 2007
21. Flynn JR: The mean IQ of Americans: massive gains 1932 to 1938. Psychological Bulletin 1984; 95:29-51
22. Flynn JR: What is Intelligence? Beyond the Flynn Effect. Cambridge, Cambridge University Press, 2007
23. Flynn JR: The WAIS-III and WAIS-IV: Daubert motions favor the certainly false over the approximately true. Applied Neuropsychology 2009; 16:2:1-7
24. Flynn JR, Widaman KF: The Flynn effect and the shadow of the past: mental retardation and the indefensible and indispensable role of IQ, in International Review of Research in Mental Retardation (Vol. 35, pp. 121-149). Edited by Glidden LM. Boston, Elsevier, 2008
25. American Association on Intellectual and Developmental Disabilities: User's Guide: Mental Retardation: Definition, Classification and Systems of Support—10th edition. Washington, DC, Author, 2007
26. Wechsler D: Wechsler Adult Intelligence Scale—Fourth Edition. San Antonio, Pearson, 2008
27. Roid GH: Stanford-Binet Intelligence Scales (5th ed.). Itaska, IL, Riverside, 2003
28. Whitaker S: The stability of IQ in people with low intellectual ability: an analysis of the literature. Intellectual and Developmental Disabilities 2008; 46:120-128
29. Dean AC, Victor TL, Boone KB, Arnold G: The relationship of IQ to effort test performance. The Clinical Neuropsychologist 2008; 22:705-722
30. Graue LO, Berry DTR, Clark JA, Sollman MJ, Cardi M, Hopkins J, Werline D: Identification of feigned mental retardation using the new generation of malingering detection instruments: Preliminary findings. The Clinical Neuropsychologist 2007; 21:929-942

31. Hurley KE, Deal WP: Assessment instruments measuring malingering used with individuals who have mental retardation: potential problems and issues. Mental Retardation 2006; 44:112-119
32. Marshall P, Happe M: The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. The Clinical Neuropsychologist 2007; 21:826-840

ABOUT THE AUTHOR

J. Gregory Olley is Clinical Professor in the Department of Allied Health Sciences and Clinical Scientist in the Clinical Center for the Study of Development and Learning, University of North Carolina at Chapel Hill. He is Chair of the Committee on Mental Retardation and the Death Penalty and President-Elect of Division 33 of the American Psychological Association.

Correspondence concerning this article should be addressed to J. Gregory Olley, Center for Development and Learning, CB 7255, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7255. E-mail: greg.olley@cdl.unc.edu