

# INTELLIGENCE TESTING AND *ATKINS*: CONSIDERATIONS FOR APPELLATE COURTS AND APPELLATE LAWYERS

LaJuana Davis\*

## I. INTRODUCTION

In *Atkins v. Virginia*<sup>1</sup> the Supreme Court held that the execution of any individual with mental retardation violates the Eighth Amendment's prohibition on cruel and unusual punishment. The holding of *Atkins* is straightforward: Persons with mental retardation cannot be subject to the death penalty. In keeping with its tradition of allowing the lower courts to determine how new constitutional rules will be implemented, however, the *Atkins* Court intentionally gave them little direction on how to apply that apparently simple, but profoundly important, prohibition. As the *Atkins* ruling passes its one-year anniversary, the decision has in consequence presented a number of challenges for the legal community, especially for the appellate courts that will be charged with establishing the procedures used to give effect to its rule.

For states like Alabama, which currently has no statute addressing capital defendants with mental retardation and a legislature that will not reconvene for its regular session until the spring of 2004, appellate courts are likely to make the initial

---

\* LaJuana Davis is a staff attorney at the Equal Justice Initiative of Alabama in Montgomery, Alabama. EJI is a non-profit organization that provides legal representation to indigent defendants and prisoners, and as part of that work, provides representation to prisoners sentenced to death in Alabama. The author thanks Ruth A. Friedman for her guidance and encouragement, Professor Randy Hertz, Director of Clinical and Advocacy Programs, New York University School of Law, for his tremendous assistance in organizing this topic, and J. Drew Colfax for his comments and generous help with access to research.

1. 536 U.S. 304 (2002).

determinations about *Atkins* and its implications. However, determining mental retardation will require appellate courts to undertake a more complicated analysis than that required by cases like *Thompson v. Oklahoma*<sup>2</sup> and *Coker v. Georgia*,<sup>3</sup> which addressed other Eighth-Amendment concerns. This Article explores some of the considerations and challenges that appellate courts and the lawyers that practice before them will face in undertaking the complicated analysis that *Atkins* requires: What are the constitutional limits on defining mental retardation, how can the legal system avoid the pitfalls inherent in seeking a definitive answer to a fluid concept such as intelligence, and how does the history of mental retardation inform the appellate courts' assessment of mental retardation?

The *Atkins* Court extended the exemption from execution to all capital defendants who "fall within the range of mentally retarded offenders about whom there is a national consensus,"<sup>4</sup> effectively reminding courts that the states cannot create laws to deprive mentally retarded defendants of an exemption guaranteed by the Eighth Amendment; whatever procedures they use to define and prove mental retardation, those procedures must satisfy the Constitution.<sup>5</sup>

However, because the Constitution "'places a substantive restriction on the state's power to take the life' of a mentally retarded offender,"<sup>6</sup> the Eighth Amendment also restricts states' ability to stray from the accepted mental-retardation definitions or to establish procedures that restrict advocates' abilities to present evidence of mental retardation. If states deviate

---

2. 487 U.S. 815 (1988) (holding that the imposition of the death penalty on persons fifteen years of age or younger violates the Eighth Amendment).

3. 433 U.S. 584 (1977) (prohibiting death penalty for rape as excessive punishment under the Eighth Amendment).

4. *Atkins*, 536 U.S. at 317.

5. There will be undoubtedly be decisionmakers unhappy with the decision in *Atkins* who will be tempted to make both definitions of mental retardation and the procedures required to prove an *Atkins* claim more restrictive. In Alabama, for example, a post-*Atkins* bill proposed that the IQ cutoff be established at "below 70," even though the prevalent clinical definitions of mental retardation include an IQ of 70 within the range of mental retardation. Ala. H. 340, 2003 Reg. Sess. at 2 (Mar. 11, 2003) (proposing the establishment of procedures in death penalty cases to determine whether a defendant is mentally retarded); see also Ala. Sen. 323, 2003 Reg. Sess. (Mar. 18, 2003) (proposing to establish an exemption from capital punishment for defendants with mental retardation).

6. *Atkins*, 536 U.S. at 321 (quoting *Ford v. Wainwright*, 477 U.S. 399, 405 (1986)).

significantly from national standards of defining mental retardation, they risk higher courts' concluding that their definitions are outside the national consensus.

## II. IQ SCORES AND MENTAL RETARDATION

In fashioning a legal definition of mental retardation, the *Atkins* Court recognized that state statutory definitions generally conform to the clinical definitions set forth by the American Association on Mental Retardation (AAMR) and the American Psychiatric Association (APA), the two leading organizations that set definitions and standards for determining mental retardation. The three core components of these definitions are (1) significantly subaverage general intellectual functioning (2) that is accompanied by significant limitations in adaptive functioning in at least two of the following skill areas: communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health, and safety, and (3) onset of the condition prior to 18 years of age.<sup>7</sup> Intellectual functioning is weighed equally with adaptive behavior in determining mental retardation.<sup>8</sup>

The APA describes the "significantly subaverage general intellectual functioning" associated with an IQ in the range that includes a score of 70 as "approximately two standard deviations below the mean [score of 100]."<sup>9</sup> In accordance with that definition, the *Atkins* Court identified an IQ between 70 and 75 or lower as "significantly subaverage general intellectual functioning."<sup>10</sup> The Court noted that only an estimated one to three percent of the population would have an IQ score at 70 or below.<sup>11</sup>

---

7. *Atkins*, 536 U.S. at 308 n. 3 (setting out AAMR definition).

8. Am. Assn. on Mental Retardation, *Mental Retardation: Definition, Classification, and Systems of Supports* 80 (10th ed., Am. Assn. on Mental Retardation 2002) [hereinafter *Mental Retardation*].

9. Am. Psych. Assn., *Diagnostic and Statistical Manual of Mental Disorders-Text Revision* (4th ed., Am. Psych. Press 2000) [hereinafter *DSM-IV-TR*].

10. *Atkins*, 536 U.S. at 309 n. 5 (setting out APA definition).

11. *Id.*

In the past, persons within this range would have been categorized as having “mild” mental retardation, but the “mild” label is being reconsidered in the psychological community today. In 1992, the AAMR dropped the terms classifying persons with mental retardation by levels of severity (profound, severe, moderate, and mild) from its definition. The AAMR cited two major reasons for this move: that the “mild” designation was based only on IQ while an assessment of mental retardation must include adaptive functioning, and that the term “mild” was at times a misnomer for “a condition that represents a considerable disadvantage.”<sup>12</sup> However, levels of severity are still used in the current version of the DSM-IV.<sup>13</sup>

In the summer of 2002, after *Atkins* was decided, the AAMR announced a new definition of mental retardation: “Mental retardation is a disability characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills. This disability originates before age 18.”<sup>14</sup> While statutory definitions of mental retardation may continue to reflect that cited in *Atkins*, then, appellate courts and appellate attorneys should be aware that psychologists and other mental health professionals may use the updated definition.

Both of the most accepted clinical definitions of mental retardation cited in *Atkins* place the IQ cutoff for mental retardation between 70 and 75.<sup>15</sup> Although the Court did not mandate that states use a particular definition of mental retardation, the states that have definitions of mental retardation prohibiting imposition of capital punishment generally fall within that outlined in *Atkins*: Arkansas,<sup>16</sup> Arizona,<sup>17</sup> Colorado,<sup>18</sup> Connecticut,<sup>19</sup> Delaware,<sup>20</sup> Florida,<sup>21</sup> Georgia,<sup>22</sup> Idaho,<sup>23</sup>

---

12. *Mental Retardation*, *supra* n. 8, at 26.

13. *DSM-IV-TR*, *supra* n. 9, at 42.

14. *Mental Retardation*, *supra* n. 8, at 1.

15. 536 U.S. at 308 n. 3.

16. Ark. Code Ann. § 5-4-618 (LEXIS 2003).

17. Ariz. Rev. Stat. § 13-703.02 (LEXIS 2003).

18. Colo. Rev. Stat. § 18-1.3-1101 (LEXIS 2003).

19. Conn. Gen. Stat. § 1-1g (LEXIS 2003).

20. Del. Code Ann. tit. 11 § 4209 (as amended by 74 Del. Laws 174) (LEXIS 2003).

21. Fla. Stat. § 921.137 (LEXIS 2003).

22. Ga. Code Ann. § 17-7-131 (LEXIS 2003).

Indiana,<sup>24</sup> Kansas,<sup>25</sup> Kentucky,<sup>26</sup> Louisiana,<sup>27</sup> Maryland,<sup>28</sup> Missouri,<sup>29</sup> Nebraska,<sup>30</sup> Nevada,<sup>31</sup> New Mexico,<sup>32</sup> New York,<sup>33</sup> North Carolina,<sup>34</sup> South Dakota,<sup>35</sup> Tennessee,<sup>36</sup> Utah,<sup>37</sup> Virginia,<sup>38</sup> and Washington<sup>39</sup> are all in this category. Arkansas's statute has a rebuttable presumption of mental retardation when there is an I.Q. score of 65 or below.<sup>40</sup> South Dakota employs a presumption that a capital defendant with an IQ over 70 is presumed not to meet the definition of "significant subaverage general intellectual functioning."<sup>41</sup> The remaining states that have capital punishment statutes but have not yet enacted statutes exempting persons with mental retardation from the death penalty (as of September 2003) are Alabama, California, Illinois,<sup>42</sup> Iowa, Mississippi, New Hampshire, New Jersey, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, Texas, and Wyoming.<sup>43</sup>

Despite some state legislative efforts to nail down a definitive IQ score to define subaverage intellectual functioning, the psychological community recommends a more flexible

---

23. Idaho Code § 19-2515A (LEXIS 2003).

24. Ind. Code § 35-36-9-2 (LEXIS 2003).

25. Kan. Stat. Ann. § 21-4623 (LEXIS 2003).

26. Ky. Rev. Stat. Ann. § 532.130 (LEXIS 2003).

27. 2003 La. Acts 698 (LEXIS 2003).

28. Md. Crim. L. Code Ann. § 2-202 (LEXIS 2003).

29. Mo. Rev. Stat. § 565.030 (LEXIS 2003).

30. Neb. Rev. Stat. § 28-105.01 (LEXIS 2003).

31. Nev. Rev. Stat. § 174.1 (as amended by 2003 Nev. Stat. 137) (LEXIS 2003).

32. N.M. Stat. Ann. § 31-20A-2.1 (LEXIS 2003).

33. N.Y. Crim. P. L. § 400.27(1) (LEXIS 2003).

34. N.C. Gen. Stat. § 15A-2005 (LEXIS 2003).

35. S.D. Codified Laws § 23A-27A-26.3 (LEXIS 2003).

36. Tenn. Code Ann. § 39-13-203 (LEXIS 2003).

37. Utah Code Ann. § 77-15a-101-102 (LEXIS 2003).

38. Va. Code Ann. § 19.2-264.3:1.1 (LEXIS 2003).

39. Wash. Rev. Code § 10.95.030 (LEXIS 2003).

40. Ark. Code Ann. § 5-4-618 (LEXIS 2003).

41. S.D. Codified Laws § 23A-27A-26.3 (LEXIS 2003).

42. Bills addressing *Atkins* were introduced in the Illinois legislature but were not enacted into law. See Ill. H. 104, 93d Gen. Assembly (Jan. 8, 2003); Ill. H. 3218, 93rd Gen. Assembly (Feb. 27, 2003); Ill. H. 3369, 93d Gen. Assembly (Feb. 28, 2003).

43. Legislative developments in these states bear watching. Lawyers and judges who are likely to be involved in capital appeals and post-conviction proceedings should follow the activities of their state legislatures.

approach. Clinical psychology takes into account a standard error of measurement of plus or minus five for an IQ score on a standardized IQ test.<sup>44</sup> Thus clinical measurement of IQ scores is properly viewed as within a range, rather than as a rigid single number.<sup>45</sup> The DSM-IV-TR requires that any intelligence test used to measure intellectual functioning must be “standardized” and “individually administered.”<sup>46</sup> IQ tests that are given in groups, therefore, are considered unreliable to establish a valid IQ score.

The AAMR cautions that determinations of mental retardation cannot be based solely on the results of an IQ test.<sup>47</sup> While a bright-line IQ score cutoff may appear to be a neat and quick solution for appellate courts facing these difficult cases, the Supreme Court has indicated that wholesale rejection of *Atkins* claims is inappropriate when the decisions are based solely on defendants’ IQ scores. In fact, in the wake of *Atkins*, the Court remanded several cases in which the defendants’ reported IQ scores were above 75.<sup>48</sup>

Psychology’s definition of mental retardation has been the foundation of the corresponding legal definitions. However, the definition of mental retardation will continue to change within the psychological community and may leave some states with outdated statutory definitions of mental retardation that may be no longer relevant to advancing psychological measurements. Although states with outdated definitions of mental retardation could alter their statutes to comport with new mental retardation definitions, given the difficulty of passing new bills in the state legislatures, those—including appellate judges—who are crafting new statutes, court rules, and procedures to give effect to *Atkins* should strive to create definitions and to implement processes broad enough to last through changing definitions.

---

44. *DSM-IV-TR*, *supra* n. 9, at 41.

45. James W. Ellis & Ruth A. Luckasson, *Mentally Retarded Criminal Defendants*, in *Symposium on the ABA Criminal Justice Mental Health Standards*, 53 *Geo. Wash. L. Rev.* 414, 475 (1985).

46. *DSM-IV-TR*, *supra* n. 9, at 41.

47. *Mental Retardation*, *supra* n. 8, at 80.

48. See *Moore v. Tex.*, 535 U.S. 1044 (2002) (Scalia, & Thomas, JJ., & Rehnquist, C.J., dissenting from grant of stay of execution) (indicating that petitioner had IQ scores of 68 and 76); *Perkins v. Ala.*, 536 U.S. 953 (2002), on remand, *Ex parte Perkins*, 851 So. 2d 453 (Ala. 2002) (indicating that petitioner had an IQ score of 76).

### III. SPECIAL AREAS OF INTEREST FOR APPELLATE COURTS AND APPELLATE LAWYERS

#### *A. Adaptive Functioning*

A key factor in *Atkins* was the component of adaptive functioning—that while mentally retarded persons “frequently know the difference between right and wrong and are competent to stand trial,” because of their impairments,

by definition they have diminished capacities to understand and process information, to communicate, to abstract from mistakes and learn from experience, to engage in logical reasoning, to control impulses, and to understand the reactions of others.<sup>49</sup>

While not warranting an exemption from criminal sanctions, the Court found that these deficiencies do “diminish [mentally retarded persons’] personal culpability.”<sup>50</sup> As one team of experts observes,

[I]t is people who have mild mental retardation that is, those who are able to function with the least assistance, who present the greatest obstacle to lawyers, the criminal justice system, and even their own defense.<sup>51</sup>

Given these factors, both lawyers and appellate courts must seek a thorough understanding of how well a capital defendant has been functioning in his environment.

Mental retardation’s adaptive functioning component, that is, a person’s skills in the practical aspects of everyday living—communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health, and safety—is one of the tricky areas in the application of *Atkins*. Two frequently used scales for measuring adaptive behavior are the AAMR’s Adaptive Behavior Scale and the Vineland Adaptive Behavior Scales.

---

49. *Atkins*, 536 U.S. at 318.

50. *Id.*

51. Denis W. Keyes, William J. Edwards & Timothy J. Dering, *Mitigating Mental Retardation in Capital Cases: Finding the “Invisible” Defendant*, 22 *Mental & Physical Disability L. Rev.* 529, 530 (1998) [hereinafter *Invisible Defendant*].

The AAMR points out that an assessment of adaptive functioning “must be considered within the context of community environments typical of the individual’s age peers and culture.”<sup>52</sup> As the Supreme Court has noted, people with mental retardation “have a reduced ability to cope with and function in the everyday world.”<sup>53</sup> Capital murder defendants suspected to have mental retardation, like the members of the general population who may have mental retardation, are likely to be within the highest-functioning level of persons with mental retardation, because eighty-five percent of those with mental retardation have IQs in the range of 55 to 75.<sup>54</sup> These persons are the least likely to have any physical attributes or any outward manifestation that would indicate a developmental disability.<sup>55</sup>

Intellectually, most in the legal profession would avow an understanding that persons with mental retardation can be otherwise capable persons who cannot function in many life-skill areas. However, some early *Atkins* decisions indicate a disconnect between our intellectual knowledge and our application of those understandings to cases involving capital defendants. Some state appellate courts have denied findings of mental retardation without further hearings, on the basis of evidence that defendants had obtained GEDs,<sup>56</sup> maintained employment,<sup>57</sup> or were married or had romantic relationships.<sup>58</sup> The psychological community rejects as too simplistic any analysis that would discount mental retardation because a person had obtained some life skills. A leader in the field of intelligence testing noted in 1905 that persons with mental retardation have “certain aptitudes” such as

good auditory or musical memory, and a whole repertoire of songs; others have mechanical ability. If all were

---

52. *Mental Retardation*, *supra* n. 8, at 8.

53. *City of Cleburne v. Cleburne Living Ctr.*, 473 U.S. 432, 442 (1985).

54. *DSM-IV-TR*, *supra* n. 9, at 43.

55. *See Invisible Defendant*, *supra* n. 51, at 530.

56. *Emmett v. Commonwealth*, 569 S.E.2d 39, 47 (Va. 2002); *Morrisette v. Commonwealth*, 569 S.E.2d 47, 56 n. 8 (Va. 2002).

57. *Emmett*, 569 S.E.2d at 47 (Va. 2002); *Ex Parte Perkins*, 851 So. 2d 453, 456 (Ala. 2002).

58. *Perkins*, 851 So. 2d at 456.

carefully examined, many examples of these partial aptitudes would probably be found.<sup>59</sup>

Advocates for persons with mental retardation have been presenting persuasive evidence for decades that, with support, persons with mental retardation can and do function with success in areas of everyday living. In fact, a core assumption of the understanding of mental retardation is that “[w]ithin an individual, limitations often coexist with strengths.”<sup>60</sup> The DSM-IV-TR also notes that person with “mild” mental retardation “may, with appropriate training and opportunities, develop good adaptive skills in other domains.”<sup>61</sup> This is something that fiction writers seem to understand. Literature does not present persons with mental retardation as incapable of action or emotions. Like characters in John Steinbeck’s *Of Mice and Men* and Daniel Keyes’s *Flowers for Algernon*, persons with mental retardation have relationships and jobs that they can perform within their limitations.

Appellate courts and appellate lawyers faced with capital defendants who may have mental retardation must keep in mind, then, that just as with any living skills, the appearance of a particular person’s ability to function in a life-skill area does not necessarily mean that he or she has complete command of that skill or has mastered the necessary related skills. Persons with mental retardation have been observed to attempt to obscure their condition behind a “mask,”<sup>62</sup> to present to others as more capable than they may be. “Masking” can be an understandable response to a sometimes dangerous and exploitative world in environments where there are few supports for persons with mental retardation.<sup>63</sup> One group of experts explains that

[i]t is rarely considered by lawyers and judges that people with mental retardation will attempt to hide their disability. Defendants who have mental retardation often are

---

59. Alfred Binet, *New Methods for the Diagnosis of the Intellectual Level of Subnormals*, in *The Development of Intelligence in Children* (Elizabeth S. Kite trans., Publications of the Training Sch. of Vineland) (reprinting material first published in *L'Année Psychologique* 12, 191-244 (Presses Universitaire de France 1905)).

60. *Mental Retardation*, *supra* n. 8, at 1.

61. *DSM-IV-TR*, *supra* n. 9, at 47.

62. See *Invisible Defendant*, *supra* n. 51, at 530; Ellis & Luckasson, *supra* n. 45, at 430-31.

63. See *Invisible Defendant*, *supra* n. 51, at 530.

characterized as quiet and cooperative. By attempting to “look like” persons of average ability, they are masking their disabilities because of the associated stigma of retardation. This “cloak of competence,” similar to the concept of “passing” is best done by keeping a low profile, not drawing attention to oneself, and agreeing with those asking the questions. As a result, people with mild mental retardation rarely are suspected of, evaluated for, or identified as having severe deficits and mental retardation. This ironic twist is called “cheating to lose” because the defendant with mental retardation deflects attention from his or her disabilities rather than bringing them to the attention of his or her lawyer or the court.”<sup>64</sup>

However, masking creates an added responsibility for judges and attorneys to look beyond first appearances of a defendant’s expressions or conduct if there is an indication that mental retardation may be present.

One common label that appellate judges and appellate lawyers should be wary of is “streetwise” or “street-smart” as used in a clinical evaluation of mental retardation that appears in a trial transcript. “Street-smart” and “streetwise” are used casually in ordinary language, but clinical psychological literature seems to ascribe no meaning to these concepts, probably because they are untestable and ultimately unhelpful to a full understanding of an individual’s intelligence and real-world capabilities. Opinions about a person’s abilities based on a report of about his or her “street-smarts” (such as a witness’s expressed belief that people who sell drugs cannot be retarded),<sup>65</sup> should be tested with an eye toward how the person was actually operating in his or her environment. For example, questions about whether the person followed a routine or whether the activities that the person was involved in actually required complicated analysis and decisionmaking would be appropriate.

---

64. *Id.* (internal citations omitted).

65. See *Ex parte Smith*, 2003 WL 1145475 at \*10 (Ala. Mar. 13, 2003) (denying *Atkins* remand because defendant’s IQ score was 72 and he did not meet the adaptive behavior component for mental retardation: “Smith was involved in an interstate illegal-drug enterprise. Smith testified that at the time of the murders he was under stress because he owed a Jamaican drug supplier in Jacksonville, Florida, \$27,000. Smith admitted that at the time of the murders he was addicted to cocaine and that he was using \$400.00 worth of crack cocaine per day; he said that in order to maintain that habit he ‘distributed’ drugs.”)

Because few judges, court personnel, lawmakers, or lawyers have any background in mental retardation evaluations or testing protocols (and indeed, relatively few psychologists and psychiatrists have extensive training in the detection of mental retardation<sup>66</sup>), appellate courts should make use of experts on developmental disabilities and testing. Common knowledge about mental retardation should be considered no more authoritative than lay understanding of any scientific issue, such as DNA evidence or pathologists' findings. Grave mistakes could be made if appellate courts base determinations about mental retardation on intuitive feelings about mental retardation or the ways in which people with mental retardation should act. Just as with any scientific, medical, or psychological inquiry, horse sense cannot substitute for clinical judgments by appropriately qualified and experienced experts in mental retardation.

### *B. Malingering*

Justice Scalia expressed in *Atkins* a concern that the decision would result in an onslaught of capital defendants faking mental retardation, or as he put it, "malingering," because he believes that the "symptoms of this condition can readily be feigned."<sup>67</sup> At least one state court has reacted to concerns about malingering by ordering capital defendants to take personality tests along with intelligence tests to determine

---

66. The research shows, in fact, that

[b]ecause mental retardation represents approximately 2 percent to 3 percent of the population, it is a small, specific subspecialty for mental health professionals. Typically, these professionals have little experience and training with people who have mental retardation. The assessment, diagnosis and treatment of this population is sometimes very different than the "typical" patient. Training in traditional mental health graduate programs includes little, if any, information about mental retardation.

*Invisible Defendant*, *supra* n. 51, at 535. Indeed,

mental retardation differs sufficiently from other forms of mental disability that training in mental illness cannot, without more, qualify a physician to provide useful information about a mentally retarded person. . . . The *Mental Health Standards* recognize that the field of mental retardation requires particular training and experience and that relatively few professionals have expertise in both mental illness and mental retardation.

Ellis & Luckasson, *supra* n. 45, at 487.

67. *Atkins*, 536 U.S. at 353 (Scalia & Thomas, JJ., & Rehnquist, C.J., dissenting).

mental retardation,<sup>68</sup> despite the fact that personality tests “have warnings that they should not be administered to individuals who are mentally disabled or who have not successfully completed at least six years of schooling.”<sup>69</sup>

Although malingering has been a concern in forensic evaluation of mental illness,<sup>70</sup> it is not likely to be a problem in mental-retardation assessments. It is unlikely that a defendant can successfully fake having mental retardation because the condition’s components provide a check against malingering. Evidence of mental retardation must be present before the onset age, which is 18 in most definitions. In other words, mental retardation must emerge at an age long before even the most cynical observer could say a defendant had reason to pretend to have mental retardation. As Jim Ellis, a leading author in the field of developmental disability and the law has pointed out,

any concerns that an individual could somehow manage to feign cognitive impairment, undetected by clinical evaluators, should be dispelled by the fact that such deception would have had to begin during the individual’s childhood. There are no reports in the clinical literature indicating that this is a practical problem in the assessment of individuals who are thought to have mental retardation.<sup>71</sup>

Additionally, the composition and sequence of questions and tasks given during IQ tests make them resistant to attempts to malingering.<sup>72</sup> The defendant’s intellectual capability and adaptive

---

68. In *Russell v. State*, 849 So. 2d 95 (Miss. 2003), the Mississippi Supreme Court remanded a case for an *Atkins* determination and further ruled that “the Minnesota Multiphasic Personality Inventory-II (MMPI-II) is to be administered since its associated validity scales make the test best suited to detect malingering.” *Id.* at 148.

69. The MMPI, for example, has some 566 items that must be read to or by the defendant, which can become incredibly tedious.

70. See e.g. *Drope v. Mo.*, 420 U.S. 162, 181 n. 16 (1975) (noting in a case involving a mentally ill defendant who apparently attempted suicide that “a self-inflicted wound near vital organs does not suggest malingering”); *Stewart v. U.S.*, 366 U.S. 1, 24 (1961) (Clark & Whittaker, JJ., dissenting) (pointing out that a defendant who sometimes “acted normal” could have been “feigning mental illness” just before his trial).

71. James W. Ellis, *Special Feature: Mental Retardation and the Death Penalty: A Guide to State Legislative Issues*, 27 *Mental & Physical Disability L. Rep.* 11, 14 (Jan./Feb. 2003).

72. See e.g. *U.S. v. Stevenson*, 135 F. Supp. 2d 878, 881 n. 2 (N.D. Ill. 2001) (“It is possible to test for malingering by noticing deviations from expected responses by people who actually have certain mental defects.”); *State v. Raiford*, 846 So. 2d 913, 917-18 (La.

functioning throughout his life must be carefully examined by experts, again, including periods when the defendant had no motivation to pretend to have mental retardation.

*C. Test Accuracy: The Flynn Effect and the Practice Effect*

Appellate courts should be aware that scores on IQ tests have been rising with every passing decade.<sup>73</sup> This phenomenon is called the “Flynn Effect,” after James Flynn, a psychologist who studied IQ testing in twenty (largely Western) countries.<sup>74</sup> Flynn’s research showed IQ test scores rose from nine to twenty points across nations and cultures over a thirty year period; IQ scores for Americans rose over thirteen points in forty-six years.<sup>75</sup> As of yet, no definitive explanation for the Flynn Effect has been discovered, although Flynn continues to study the issue.<sup>76</sup> An appellate court considering an *Atkins* claim might thus need to ask questions about when the relevant intelligence tests were last re-normed when older IQ test scores are compared in a particular case to more recent scores.<sup>77</sup>

Appellate courts should also be aware that when a person is given the same or an equivalent IQ test in a short span of time, higher scores may be achieved on the later tests, but those increased scores do not reflect improved intelligence or even greater effort. This situation is called “practice effect,” which refers to “gains in scores on cognitive tests that occur when a

---

App. 4th Cir. 2003) (documenting different tests that psychologist used during evaluation to ensure that responses were consistent with ability shown).

73. See e.g. Marguerite Holloway, *Flynn’s Effect* (available at <http://www.sciam.com/article.cfm?articleID=00037F65-D9C0-1C6A-84A9809EC588Ef21&catID=2>) (profile of Flynn first published in *Scientific American* magazine of January 1999) (accessed Nov. 25, 2003; copy on file with Journal of Appellate Practice and Process).

74. *Id.*

75. *Id.*

76. See Kristen Leutwyler, *New Model Solves IQ Paradox* (available at <http://www.sciam.com/article.cfm?articleID=0004E175-930C-1C5E-B882809EC588ED9F&catID=1>) (report of Flynn’s recent research first published in *Scientific American* magazine of April 2001) (accessed Nov. 25, 2003; copy on file with Journal of Appellate Practice and Process).

77. Renorming is the process by which test scores obtained from a sample set of people taking the same test are plotted on a bell curve in order to determine average scores.

person is re-tested on the same instrument, or tested more than once on very similar ones.”<sup>78</sup> One observer notes that

[t]hese gains are due to the experience of having taken the test previously; they occur without the examinee being given specific or general feedback on test items, and they do not reflect growth or other improvement on the skills being assessed.”<sup>79</sup>

In a situation in which a defendant is given multiple tests within a short interval and appears to perform better on the later tests, courts and lawyers should consider whether the practice effect may be affecting the later scores.

#### *D. Cultural Bias and Testing Subjectivity*

The scales that are the foundation of contemporary IQ tests did not originally include non-whites as test subjects. African Americans, for example, were not represented in numbers proportionate to their presence in the population until well into the 1970s.<sup>80</sup> And although efforts were made to balance the scoring on the tests when it was discovered that some questions produced different answers based on gender, relatively little effort was made until the 1970s and 1980s to re-norm IQ tests when non-whites routinely gave different answers to specific questions than did whites.<sup>81</sup>

Some argue that the intelligence tests have not fully outgrown their more exclusionary origins<sup>82</sup> and that test

---

78. Alan S. Kaufman, *Practice Effects*, in *Psychological Forum* (available at <http://www.psychologicalforum.com/articles/aug03.asp>) (reprinting text from 2 *Encyclopedia of Human Intelligence* 828-33 (Robert J. Sternberg, ed., Macmillan Publ. Co. 1994)) (accessed Nov. 25, 2003; copy on file with *Journal of Appellate Practice and Process*).

79. *Id.*

80. *Larry P. v. Riles*, 495 F. Supp. 926, 957 (N.D. Calif. 1979), *aff'd in part, rev'd in part*, 793 F.2d 969 (9th Cir. 1984).

81. *Id.*

82. Henry Goddard, Lewis Terman, and David Wechsler, among the pioneers of intelligence testing, all subscribed to then-socially acceptable belief that non-whites, and indeed, all non-Anglo-Saxons, were genetically inferior to Anglo-Saxon whites. See e.g. *Larry P.*, 495 F. Supp. at 936 (quoting Lewis Terman, *The Measurement of Intelligence* 91 (Houghton Mifflin Co. 1916) (declaring that borderline feeble-minded represented “the level of intelligence which is very, very common among Spanish-Indian and Mexican families of the South-west and also among negroes,” and that “[t]heir dullness seems to be racial or at least inherent in the family stocks from which they come”)).

designers have not sufficiently scrutinized IQ tests for aspects of cultural bias.<sup>83</sup> For example, one question on a version of the Stanford-Binet that was being administered in the 1980s assumed that there were native Americans in the United States unfamiliar with bicycles:

An Indian who had come to town for the first time in his life saw a boy riding along the street. As the boy rode by, the Indian said, "The white boy is lazy; he walks sitting down." What was the boy riding on that caused the Indian to say "he walks sitting down?"<sup>84</sup>

One judge observed in a 1979 case that

[c]ultural differences can also be found in specific test items. Some of these items have in fact become rather notorious, such as the "fight item" on WISC tests. This question asked children what they would do if struck by a smaller child of the same sex. The "correct" answer is that it is wrong to strike the child back. . . . Young black children aged six and seven "missed" this item more than twice as often as their white counterparts. . . . The difference can only be attributed to a cultural variation at that age. Similarly, it may be that such questions as who wrote *Romeo and Juliet*, who discovered America, and who invented the light bulb, are culturally biased.<sup>85</sup>

An example of a right answer being in the eye of the beholder is a question on the WAIS-R for Children. The second item given in the vocabulary section is for the child to define "umbrella." In one case, the test scoring system was explained this way:

83. Most IQ tests were not normed with non-white populations. See *Larry P.*, 495 F. Supp. at 957 (quoting David Wechsler, *The Measurement of Intelligence* 107 (Williams & Wilkins Co. 1944) (concluding that the WAIS was not valid for "the colored population of the United States" and acknowledging that "we have eliminated the colored vs. white factor by admitting at the outset that our norms cannot be used for the colored population of the United States"))).

84. *Parents in Action on Spec. Educ. (PASE) v. Hannon*, 506 F. Supp. 831, 869 (N.D. Ill. 1980) (finding standard intelligence tests administered by the Chicago Board of Education not to be, in the main, culturally biased against black children).

85. *Larry P.*, 495 F. Supp. at 958 (internal citations omitted). According to the tester's guide, acceptable answers for the question "Who discovered America?" are "Columbus . . . Leif Erickson, Vikings (Norsemen), Amerigo Vespucci. (If a child says 'Indians,' say 'yes, the Indians were already there, but who sailed across the ocean and discovered America?')." *PASE*, 506 F. Supp. at 838.

Two-point responses are: "Use it to keep the rain off . . . protects you when it rains . . . put it over your head when it rains . . . so you don't get wet when it rains." One-point responses are: "Carry it when it rains . . . big round thing that can fold up . . . put it over your head . . . to keep off the sun . . . you hold it up (gives appropriate demonstration) . . . helps you if it starts raining . . . keeps you dry. All of these one-point responses are marked with "Q" in the manual, indicating that the examiner should follow up the response with another question as to what the child means. If a child says, "Put it over your head," the examiner should ask, "Explain what you mean." If the child says something like, "You know, like when it rains," he is given two points for the response."<sup>86</sup>

The inherent bias in this scoring system is obvious: If a child responds that umbrellas are used to keep off the sun, that child may get only a one-point score for that apparently partial answer in Seattle, while in Alabama, that answer might be considered perfectly acceptable and credited for two points, since some people in Alabama carry umbrellas to deflect the sun's rays. Although test designers assure that some of the more outdated and culturally suspect questions have been eliminated from the latest versions of intelligence tests, it cannot be forgotten that IQ tests remain highly subjective evaluations of achievement.

#### *E. The Checkered Past of IQ Testing in the United States*

Appellate courts and lawyers who rely on test scores as indications of mental retardation should be aware that they may be placing their faith in the results of mental-retardation measurements that were never intended to be used in a legal context. The tests themselves were designed only to be academic screening devices but were later used for what by contemporary standards would be considered pernicious purposes. A very brief look back at the history of IQ testing reveals the complicated ground upon which appellate courts that rely on test results as indicia of mental retardation are basing their life-and-death decisions.

---

86. *PASE*, 506 F. Supp. at 844.

### *1. European Beginnings: The Binet Scale*

While humans have contemplated methods of determining intellect since the beginning of recorded history, French psychologist Alfred Binet is regarded as a pioneer of modern intelligence testing. In 1904, Binet, the director of the psychology laboratory at the Sorbonne who was also trained as a lawyer, was asked by the French minister of education to devise a method to direct special education services to schoolchildren who were performing poorly in the public schools. Binet saw his task as quantifying “basic processes of reasoning” and “everyday tasks” rather than learned intelligence, or, in other words, he sought to measure a child’s potential, rather than how much classroom information the child had retained. Binet developed an intelligence test with a collaborator that became known as the Binet-Simon Scale, and which, in a revised version, is still widely used.

In 1905, Binet began administering the scale to Parisian schoolchildren. The Binet-Simon Scale had thirty tasks, beginning from the most simple, such as having a child follow a lit match with her eyes, to more difficult tasks of copying patterns, naming objects, having the child form sentences from provided words, repeating phrases and figures in sequence and answering questions such as “My neighbor has been receiving strange visitors. He has received in turn a doctor, a lawyer, and then a priest. What is taking place?”<sup>87</sup> Standards were created based on the children’s performance. “For example, if 70 percent of 8-year-olds could pass a particular test, then success on the test represented the 8-year-old level of intelligence.”<sup>88</sup>

Binet was very concerned about the test being misused, writing that he was concerned about the authorities rationalizing their use of the test as an “‘excellent opportunity for getting rid of all the children who trouble us,’ and without the true critical spirit, they designate all who are unruly, or disinterested in the

---

87. J. A. Plucker, *Human intelligence: Historical Influences, Current Controversies, Teaching Resources* (2003) (available at <http://www.indiana.edu/~intell.binet.shtml>) (accessed Nov. 26, 2003; copy on file with Journal of Appellate Practice and Process).

88. PBS Online, *A Science Odyssey: People and Discoveries: Binet Pioneers Intelligence Testing* (WGBH 1998) (available at <http://www.pbs.org/wgbh/aso/databank/entries/dh05te.html>) (accessed Nov. 26, 2003; copy on file with Journal of Appellate Practice and Process).

school.”<sup>89</sup> As his test came into wider use, particularly in the United States, Binet’s fears were to be realized.

## 2. Early IQ Testing in the United States

Discussing the history of mental retardation in the nineteenth century, Justice Marshall once characterized the treatment of persons with mental retardation as “grotesque.”<sup>90</sup> He observed that

a regime of state-mandated segregation and degradation soon emerged that in its virulence and bigotry rivaled, and indeed paralleled, the worst excesses of Jim Crow. Massive custodial institutions were built to warehouse the retarded for life; the aim was to halt reproduction of the retarded and “nearly extinguish their race.” Retarded children were categorically excluded from public schools, based on the false stereotype that all were ineducable and on the purported need to protect nonretarded children from them. State laws deemed the retarded “unfit for citizenship.”<sup>91</sup>

Early IQ testing in America was to be used to support the goals of Social Darwinism—the belief that persons with mental retardation had to be identified and segregated from society.<sup>92</sup>

### a. Goddard and Ellis Island

In 1908, an American psychologist, Henry H. Goddard, then the director of the Vineland Training School in New Jersey, translated the Binet-Simon Scale into English and headed a

---

89. Stephen Jay Gould, *The Mismeasure of Man* 181 (W.W. Norton & Co. 1996).

90. *City of Cleburne*, 473 U.S. at 461 (Marshall, J., concurring in part and dissenting in part).

91. *Id.* at 461-63 (Marshall, J., concurring in part and dissenting in part) (internal citations omitted).

92. *Id.* at 461 (Marshall, J., concurring in part and dissenting in part). One author defined Social Darwinism as

a general term for any evolutionary argument about the biological basis of human differences, but the initial nineteenth-century meaning referred to a specific theory of class stratification within industrial societies, and particularly to the idea that there was a permanently poor underclass consisting of genetically inferior people who had precipitated down into their inevitable fate.

Stephen Jay Gould, *Curveball*, in *New Yorker* 139, 139 (Nov. 28, 1994) (reviewing Richard J. Herrnstein & Charles Murray, *The Bell Curve* (Free Press 1994)).

wave of efforts using intelligence testing for social purposes.<sup>93</sup> A supporter of the theory that “feeble-mindedness” was due to heredity, Goddard proposed isolating persons with mental retardation from society so that they would not have children. He also declared that mentally retarded people constituted a “menace to society and civilization . . . responsible in a large degree for many, if not all, of our social problems.”<sup>94</sup> In support of his theory, Goddard authored a highly influential study about a young woman residing at Vineland and the incidence of mental retardation in her family.<sup>95</sup> The popularity of his studies of persons with mental retardation led to an invitation from the Commission of Immigration at Ellis Island for Goddard to develop testing procedures to identify immigrants who had mental retardation. Before Goddard’s involvement, immigration officers rejected immigrants they suspected of being “mentally defective” on the basis of visual inspections, a notion that Goddard had partially repudiated in *The Kallikak Family*.<sup>96</sup> Goddard began his assignment by carrying out a study of immigrants in public institutions.<sup>97</sup> His first conclusion was that the then-prevailing public opinion that majority of immigrants were mentally retarded was “grossly overestimated.”<sup>98</sup>

Even at this early stage of mental-retardation testing, practitioners were “acutely aware of the political implications of

---

93. John T.E. Richardson, *Howard Andrew Knox and the Origins of Performance Testing on Ellis Island*, in *History of Psychology* vol. 6, 143, 147-48 (Educ. Pblg. Found. 2003).

94. Ellis & Luckasson, *supra* n. 45, at 418.

95. Henry H. Goddard, *The Kallikak Family: A Study in the Heredity of Feeble-Mindedness* (Macmillan Co. 1927). Decades later, researchers alleged that the young woman who was the source of the study was in fact learning-disabled, not retarded, and further alleged that her reported family history was altered to support Goddard’s conclusions. See Kwame Dakwa, Amber Esping & Jonathan Plucker, *The Kallikak Family in Human Intelligence: Historical Influences, Current Controversies, Teaching Resources* (available at <http://www.indiana.edu/~intell/kallikak.shtml>) (accessed Nov. 26, 2003; copy on file with Journal of Appellate Practice and Process).

96. Richardson, *supra* n. 93, at 150, 156. As Goddard wrote,

A large proportion of those who are considered feeble-minded in this study are persons who would not be recognized as such by the untrained observer. They are not the imbeciles nor idiots who plainly show in their countenances the extent of their mental defect.

Goddard, *supra* n. 95, at 104.

97. Richardson, *supra* n. 93, at 148.

98. *Id.*

their work,”<sup>99</sup> concerns that echo in current applications of *Atkins*. One of Goddard’s contemporaries worried that the IQ tests administered to the immigrants, in addition to being limited to those who spoke English, were also culturally inappropriate:

To the uninitiated using routine tests for defectives nearly all the peasants from certain European countries appear to be of the moron type; but of course this a fallacy. If these peasants are questioned about conditions existing in the land from which they come most of them will show average intelligence.<sup>100</sup>

Notwithstanding these cautions, Goddard applied IQ testing to immigrants arriving at Ellis Island, but only to those who were arriving in steerage; first and second class passengers were interviewed privately in their cabins. Goddard found that over eighty percent of the Russian, Hungarian, and Jewish immigrants tested at Ellis Island were mentally retarded.<sup>101</sup> Thousands of immigrants were refused entrance to the United States because of recommendations based on their IQ scores.

In a journal article summarizing his findings at Ellis Island, Goddard expressed surprise at the immigrants’ low scores, but his expectations can be amusing by contemporary standards. Reflecting on the inaccurate answers when immigrants were asked to tell the examiners the date, Goddard wrote,

What shall we say of his ignorance of the date? This does not mean the exact date, since a leeway of a few days is always allowed. Must we again conclude that the European peasant of the type that immigrates to America pays no attention to the passage of time? That the drudgery of life is so severe that he cares not whether it is January or July, whether it is 1912 or 1906? Is it possible that the person may be of considerable intelligence and yet, because of the peculiarity of his environment, not have acquired this ordinary bit of knowledge, even though the calendar is not in general use on the continent, or is somewhat complicated

---

99. *Id.* at 149.

100. Howard Andrew Knox, *The Moron and the Study of Alien Defectives*, 60 *J. Am. Med. Assn.* 105, 105 (Jan. 1913).

101. Leon Kamin, *The Science and Politics of IQ* 16 (Lawrence Erlbaum Assoc. 1974).

as in Russia? If so what an environment it must have been!<sup>102</sup>

Of course, viewed by contemporary standards, it seems reasonable that a person traveling from Russia in the early 1900s would have been traveling for such a long time, and for such a long distance, much of it in third-class accommodations at sea, that he or she might not know what month or even what year it was immediately upon arriving at Ellis Island.

In 1912, a German psychologist, while studying student scores on the Binet-Simon Scale, found that the “intellectual level” of the children corresponded to their chronological age. That psychologist, Wilhelm Stern, reformed the Binet-Simon Scale’s “intellectual level” measurement by dividing the child’s chronological age by the age at which the child reached her maximum achievement; the formula he used was  $IQ = \text{mental age} / \text{chronological age} \times 100$ . The resulting number was the mean IQ score, sometimes also referred to as a ratio IQ measurement. Lewis Terman of Stanford adopted Stern’s measurement system and named this ratio “intelligence quotient,” which was the origin of the term “IQ” that is used today. In 1916, Terman also published a revised Binet-Simon scale re-normed for American populations. This “Stanford Revision of the Binet-Simon Scale,” soon became known as the “Stanford-Binet” intelligence test that, with modern additions and revisions, is still in use today.

#### b. Testing of Recruits in World War I

World War I provided the impetus for the next groundbreaking event in the history of intelligence testing in the United States. Faced with assessing the abilities of thousands of new recruits, the army sought a way to classify soldiers according to their intellectual abilities. The then-president of the APA, Robert Yerkes, was asked to assemble a committee to devise an intelligence test that the army could administer to

---

102. Henry Goddard, *Mental Tests and the Immigrant*, 2 J. Delinquency 243 (Sept. 1917) (available at <http://www.arthurhu.com/99/09goddard.htm>) (accessed Dec. 19, 2003; copy on file with Journal of Appellate Practice and Process).

groups of new servicemen in an hour or less.<sup>103</sup> Yerkes and his team, relying on scales developed by experts such as Arthur Otis, created two scales, an “Alpha” scale for literate, English-speaking recruits and a “Beta” scale for those soldiers who were illiterate or non-English speakers.<sup>104</sup> Army records estimate that, from 1917 to 1919, the Alpha and Beta examinations were administered to 1,726,966 recruits.<sup>105</sup> In a reminder of Henry Goddard’s Ellis Island tests, Yerkes’s army-screening test proved to be highly influential to legislators seeking to establish quotas on the number of immigrants entering the United States. The quotas eventually barred six million European refugees from entering the country between 1924 and the outbreak of World War II. One author later theorized that this effort to reduce immigration may have prevented many of these refugees from escaping the Holocaust: “We know what happened to many who wished to leave but had nowhere to go. The paths to destruction are indirect, but ideas can be agents as sure as guns and bombs.”<sup>106</sup> The army tests, today referred to as the Otis Alpha/Beta scales, are now viewed with some disfavor,<sup>107</sup> but appellate courts should be aware that they are still sometimes used to assess prisoners.

During World War I, a psychological examiner named David Wechsler became convinced that the IQ tests then in use did not truly reflect the abilities of the soldiers to whom he administered the tests. He later recalled the case of a “a native, white, Oklahoman” who was tested at a mental age of eight on a Stanford-Binet scale, but who “before entering the Army. . . had gotten along very well, was supporting a family, [and] had been

---

103. Corwin Boake, *From the Binet-Simon to the Wechsler-Bellevue: Tracing the History of Intelligence Testing*, 24 *J. Clin. & Experimental Neuropsychology* 383, 390 (2002).

104. *Id.*

105. *Id.*

106. Gould, *supra* n. 89, at 263. Yerkes’s protégé, Carl C. Brigham, whose work helped pass the 1924 [Immigration] Restriction Act that led to the quotas, was later to be well-known himself as the secretary of the College Entrance Examination Board, where he designed the first S.A.T.

107. *See Invisible Defendant*, *supra* n. 51, at 536 (noting that tests administered in group settings are not appropriate for individualized mental-retardation evaluations).

working as a skilled-oil driller for several years.”<sup>108</sup> Wechsler observed that the Oklahoman he tested

systematically rates as a mental defective on mental tests, but, who can in no way be judged as such, when diagnosed on the basis of concrete social standards, i.e., in terms of capacity to adjust to the normal demands of his social and economic environment.<sup>109</sup>

After the war, Wechsler worked at New York’s Bellevue Psychiatric Hospital. While at Bellevue, he became convinced that the IQ tests then in use were inappropriate for adults because the ratio IQ measurement (calculated by using the formula  $IQ = \text{mental age} / \text{chronological age} \times 100$ ) was not applicable to adults, who were past the developmental stage and therefore, were not as easily compared with other adults as children can be compared with each other.<sup>110</sup> Wechsler also believed that some IQ tests relied too heavily on verbal scales and that low IQ measurements were more attributable to literacy levels than to actual intelligence.<sup>111</sup>

Wechsler designed the Wechsler-Bellevue Scale to address these concerns. The Wechsler-Bellevue, published in 1939, was an instrument with sub-tests to measure both verbal and nonverbal (performance) abilities that borrowed from many other tests, such as the army’s Alpha test.<sup>112</sup> He adopted a mean score of 100, following the Stanford-Binet metric, but replaced the ratio IQ with a deviation score. The deviation score determined how far above or below the mean (or point of average intelligence) the person tested was, and this measurement method is used today, with results expressed in “standard deviations from the mean.” In 1949, Wechsler produced the Wechsler Intelligence Scale for Children (WISC), and in 1955, he produced a revision of the adult scales named the Wechsler Adult Intelligence Scale (WAIS). Revised versions of both tests are used today, currently known as the WISC-III

---

108. Boake, *supra* n. 103, at 394.

109. *Id.*

110. *Id.* at 396.

111. *Id.*

112. *Id.* at 397.

and the WAIS-III. The WAIS-III is in fact the test of adult intelligence most widely used in clinical settings.<sup>113</sup>

c. Supreme Court Approval of IQ Testing: *Buck v. Bell*

In 1927, the Supreme Court indirectly affirmed the use of widespread IQ testing for social purposes in *Buck v. Bell*,<sup>114</sup> a case challenging a Virginia law mandating sterilization of persons with mental retardation. The petitioner, a young mother with a child of an allegedly feeble mind, was sterilized under the law after scoring a mental age of nine on the Stanford-Binet. Carrie Buck's mother, then fifty-two, had tested at a mental age of seven.<sup>115</sup> Oliver Wendell Holmes, Jr., delivered the Court's decision upholding the Virginia sterilization law, declaring,

We have seen more than once that the public welfare may call upon the best citizens for their lives. It would be strange if it could not call upon those who already sap the strength of the state for these lesser sacrifices, often not felt to be such by those concerned, in order to prevent our being swamped with incompetence. It is better for all the world, if instead of waiting to execute degenerate offspring for crime, or to let them starve for their imbecility, society can prevent those who are manifestly unfit from continuing their kind. Three generations of imbeciles are enough.<sup>116</sup>

*Buck v. Bell* opened the door to Virginia's forcible sterilization of 8,500 persons at institutions like the one where Carrie Buck was housed; the practice was stopped in the 1970s, although legislators did not call for repeal of the law until the 1980s.<sup>117</sup>

In an ironic postscript to *Buck v. Bell*, in the 1980s, the director of the institution where Carrie Buck was sterilized investigated her family history. The director discovered that she

---

113. Gordon E. Taub, *A Confirmatory Analysis of the Wechsler Adult Intelligence Scale—Third Edition: Is the Verbal/Performance Discrepancy Justified?* in *Prac. Assessment, Research & Evaluation* 7(22) (2001) (available at <http://pareonline.net/getvn.asp?v=7&n=22>) (accessed Nov. 26, 2003; copy on file with Journal of Appellate Practice and Process).

114. 274 U.S. 200 (1927).

115. Stephen Jay Gould, *Carrie Buck's Daughter: A Popular, Quasi-Scientific Idea Can Be a Powerful Tool for Injustice*, in *Natural History* 12, 14 (July/Aug. 2002).

116. *Buck*, 274 U.S. at 207.

117. Robert Reinhold, *Virginia Hospital's Chief Traces 50 Years of Sterilizing the "Retarded,"* 129 *New York Times* 6 (Feb. 23, 1980).

was apparently sent to an institution for the mentally retarded because she had been impregnated as a result of a rape and needed to go away to have the baby; the institution at which she ended up was the only place where there was room for her.<sup>118</sup> The director also discovered school records revealing that her child was considered bright and regarded as an adequate student.<sup>119</sup>

*F. The Legacy of IQ Testing:  
Questions for Today's Appellate Courts*

Fresh controversies that relate back to the origins of the tests have evolved in recent years. For decades, black children have been disproportionately represented in classes for mildly retarded students.<sup>120</sup> One example was the Chicago school system: In the 1978-79 school year, 13,225 children were enrolled in special education classes. Of these, 10,833, or eighty-two per cent, were black. In contrast, of the 106,581 white children enrolled in the system, only 1,404 were attending special education classes.<sup>121</sup> In California, black children represented ten percent of the student population in 1975, but made up twenty-five percent of the population enrolled in special-education classes.<sup>122</sup> Parents of black children nationwide sued their school systems, charging that the intelligence tests used to classify students were racially discriminatory.

In 1971, a school-aged plaintiff sued the San Francisco schools on behalf of a group of African-American children. He charged that the placement of black pupils in special programs for the "mildly retarded" was discriminatory and violated the Fourteenth Amendment, because discriminatory intelligence tests had been used to classify them as retarded. The use of such tests to assign minority children to special-education classes was

---

118. Gould, *supra* n. 115, at 16.

119. Reinhold, *supra* n. 118.

120. William G. Buss, *Intelligence Testing and Judicial Policy Making for Special Education*, in 2 *Psychology, Pub. Policy, & L.* 584, 584 (Sept./Dec. 1996).

121. *PASE*, 506 F. Supp. at 833 (finding that standard intelligence tests administered by the Chicago Board of Education were not, in the main, culturally biased against black children).

122. *Larry P.*, 495 F. Supp. at 931.

immediately prohibited by a federal court's injunction that prohibited the schools

from utilizing, permitting the use of, or approving the use of any standardized intelligence tests . . . for the identification of black E.M.R. [educable mentally retarded] children or their placement into EMR classes, without securing prior approval by this court.<sup>123</sup>

That decision was modified later by the Ninth Circuit, which affirmed only on the grounds brought under federal statutory law.<sup>124</sup>

In 1980, the *PASE* suit was filed in Chicago, and although the plaintiffs relied on the same legal arguments and experts used in the California case, the court discounted much of the expert testimony presented and examined the intelligence tests for itself. Its opinion lists all of the items and answers for three widely used tests, and the court eventually held

that the WISC, WISC-R, and Stanford-Binet tests, when used in conjunction with the statutorily mandated "other criteria for determining an appropriate educational program for a child" . . . do not discriminate against black children in the Chicago public schools.<sup>125</sup>

The different conclusions in these two cases may be traced back to the way in which the appellate courts involved viewed the effect of the test designers' known belief "that low test performance by Black test takers resulted from the genetic inferiority of Black people."<sup>126</sup> The California court focused on this "historical strain of racist beliefs,"<sup>127</sup> and found that in the absence of any other acceptable explanation for the disparity of test results between white and black children, the tests were culturally biased. The *PASE* court, on the other hand, discounted the significance of the history of the IQ tests because no attempt was made by the plaintiffs to connect these beliefs with the tests at issue. Instead, the *PASE* court held that it was the plaintiffs' burden to show that cultural bias was present in the test items themselves, a burden that the court found they did not meet.

---

123. *Larry P.*, 495 F. Supp. at 989.

124. *Larry P. v. Riles*, 793 F.2d 969 (9th Cir. 1984).

125. *PASE*, 506 F. Supp. at 883.

126. Buss, *supra* n. 120, at 595.

127. *Id.*

## IV. CONCLUSION

Viewed in light of the history of IQ testing in the United States, *Atkins* poses a significant challenge to the appellate courts. When called upon to evaluate the sentences of capital defendants who may suffer from mental retardation, they must ensure that experts who are specialists in mental retardation examine and evaluate those defendants. Equally important, they must resist the temptation to seek easy answers from IQ tests; history demonstrates that those tests are not designed to provide the sorts of assessments critical in the life-or-death circumstances of capital litigation.

It will also be important for appellate courts—and the lawyers who practice before them—to stay abreast of developments in this area of the law. The Supreme Court's decision in *Atkins* came less than fifteen years after its earlier decision in *Penry v. Lynaugh*,<sup>128</sup> which indicated that capital defendants with mental retardation could in fact be executed. This relatively rapid turn-around suggests that the law affecting the punishments appropriate for defendants with mental retardation will continue to evolve. The lawyers and judges responsible for capital appeals and post-conviction proceedings cannot afford to miss the inevitable developments as they occur.



---

128. 492 U.S. 302 (1989).

