



IN THE COURT OF CRIMINAL APPEALS OF TEXAS

NO. WR-55,161-02

Ex Parte ERIC DEWAYNE CATHEY, Applicant

**ON APPLICATION FOR A WRIT OF HABEAS CORPUS
IN CAUSE NO. 713189-B IN THE 176th DISTRICT COURT
HARRIS COUNTY**

COCHRAN, J., delivered the opinion of the Court in which KELLER, P.J., and MEYERS, WOMACK, JOHNSON, KEASLER, HERVEY, and ALCALA, JJ., joined. PRICE, J., joined Parts I and IIA and filed a concurring opinion.

OPINION

Applicant was convicted of capital murder and sentenced to death in 1997 for fatally shooting Cristina Castillo while kidnapping her. We affirmed his conviction and sentence in 1999,¹ and denied relief on his first application for a writ of habeas corpus in 2003.² On the day before his scheduled execution, applicant filed a subsequent writ alleging, for the first

¹ *Cathey v. State*, 992 S.W.2d 460 (Tex. Crim. App. 1999).

² *Ex parte Cathey*, No. WR-55,161-01 (Tex. Crim. App. April 2, 2003) (not designated for publication).

time, that he was mentally retarded and therefore exempt from the death penalty. The next day we stayed applicant's execution and issued an order finding that his claim satisfied the requirements of Article 11.071, § 5, and remanded the case to the trial court to conduct a hearing on his mental retardation claim.³ The trial judge conducted a five-day hearing that included testimony from numerous expert witnesses. Both the State and applicant filed proposed findings of fact and conclusions of law on February 21, 2011. On December 31, 2012, almost two years after the hearing and on the last day of her term of office, the trial judge signed applicant's proposed findings of fact and conclusions of law. We filed and set this case and ordered briefing by the parties.

We hold that applicant has not established, by a preponderance of the evidence, that he is mentally retarded⁴ under *Atkins v. Virginia*⁵ and *Ex parte Briseno*;⁶ therefore he is not exempt from the death penalty. We conclude that the record does not support the habeas judge's factual findings or legal conclusions. In short, the judge erred in finding,

³ *Ex parte Cathey*, No. WR-55-161-02, 2008 WL 4927446 (Tex. Crim. App. Nov. 18, 2008) (not designated for publication).

⁴ The term "mentally retarded" has been changed to "intellectually disabled," as mental health advocates decided that the former term had pejorative connotations. See Tomoe Kanaya, et al., *The Flynn Effect and U.S. Policies: The Impact of Rising IQ Scores on American Society Via Mental Retardation Diagnosis*, 58 AM. PSYCH. 778, 788 (2003) (noting that "the fact that the MR label carries with it an inherent negative stigma is no better illustrated than by the fact that a former label is continually supplanted by newer ones over time. For example, terms such as *imbecile* and *feeble-minded* were considered scientific and acceptable in the first quarter of the 20th century but were replaced after time with successive euphemisms."). The terms may be used interchangeably.

⁵ 536 U.S. 304 (2002).

⁶ 135 S.W.3d 1 (Tex. Crim. App. 2004).

- (1) The “Flynn Effect” authorized her to subtract 5.4 points from applicant’s IQ score of 77, and the standard measurement of error authorized her to subtract another 5 points from his IQ score, thus concluding that applicant’s “true” IQ score is as low as 66.6.
- (2) The State was not allowed to have applicant’s IQ retested with a more recently normed test when Dr. Flynn testified that his purpose in the “Flynn Effect” is to show that IQ tests should be normed and revised with greater frequency.⁷
- (3) The Vineland test answers given by applicant’s sister trying to retrospectively remember her brother’s behavior twenty-six years earlier and that of his former wife some eighteen years earlier were scientifically valid.
- (4) The Vineland test answers given by applicant’s sister and his former wife were reliable when, in fact, they contradicted their prior trial testimony at a time that they had no motive to exaggerate applicant’s poor adaptive behavior.
- (5) The applicant is mentally retarded or intellectually disabled, because we conclude that the evidence clearly demonstrates his intellectually competent adult behavior.

⁷ Applicant’s Proposed Findings at 38 (“Because Mr. Cathey’s experts relied on Dr. Yohman’s score [77 I.Q.] during the evidentiary hearing and did not present testimony based on a new intelligence test, retesting was not necessary.”). A footnote attached to this finding states that “[t]he State should be collaterally estopped from objecting now to Dr. Yohman’s testing and score because it failed to object on these grounds at Mr. Cathey’s trial.” This makes no sense. Collateral estoppel applies only when an elementary issue has been fully and finally litigated. There was no “element” of mental retardation at applicant’s trial, which took place before *Atkins* was decided. Furthermore, this footnote states that “[t]he issue of cognitive disability was placed before the jury during the punishment phase of trial, and the State had ample reason, at that time, to request testing of Mr. Cathey.” No such issue was “placed before the jury” at trial in 1997, and Dr. Yohman did not testify at trial that applicant was mentally retarded or intellectually disabled.

As we explained in *Ex parte Reed*, 271 S.W.3d 698, 727 (Tex. Crim. App. 2008), we normally defer to the habeas judge’s factual findings, but “[w]hen our independent review of the record reveals that the trial judge’s findings and conclusions are not supported by the record, we may exercise our authority to make contrary or alternative findings and conclusions.” And, “[w]hen our independent review of the record reveals findings and conclusions that are unsupported by the record, we will, understandably, become skeptical as to the reliability of the findings and conclusions as a whole. In such cases, we will proceed cautiously with a view toward exercising our own judgment.” *Id.*

Although we agree that factfinders may “consider” the concept of the “Flynn Effect” in assessing the validity of a WAIS or WAIS-R IQ test score, they may consider that effect only in the way that they consider an IQ examiner’s assessment of malingering, depression, lack of concentration, and so forth. It is a generalized consideration that could detract from the over-all validity of the score obtained. The preferred solution to an outdated IQ score is not to start subtracting from that score, it is to retest with a more recently normed IQ test.⁸

As Professor James Flynn⁹ stated at the writ hearing, “[T]here would be no competent

⁸ ALAN S. KAUFMAN, IQ TESTING 101, 203 (2009) (noting that publishers standardize IQ tests and determine the basis for calculating IQ scores at specific points in time; thus an IQ test grows outdated and its norms grow obsolete as time passes from the time when the publisher standardized the test, and obsolete norms inflate IQ scores because they measure test performances against the scores of test takers from the past, as opposed to the higher scores of test takers from the present); *see also* JAMES R. FLYNN, WHAT IS INTELLIGENCE? BEYOND THE FLYNN EFFECT 111-28 (Cambridge Univ. Press 2009). Professor Flynn notes that “the target percentage of about 2 percent” of children being diagnosed as mentally retarded “has been attained only fleetingly and then only by accident.” *Id.* at 128. *See also* James R. Flynn, *Individual Differences: Implications for Educational Policy: The Hidden History of IQ and Special Education: Can the Problem Be Solved?*, 6 PSYCH. PUB. POL’Y & L. 191, 191-98 (2000) (suggesting that a team of qualified psychologists gather a representative sample of MR children based on behavioral criteria and reform IQ tests every seven years); Tomoe Kanaya, et al., *The Flynn Effect and U.S. Policies*, 58 AM. PSYCHOL. 778, 780 (2003) (“As IQ norms age, fewer students receive MR services, but when a newly normed test is introduced, the number of students eligible for these services will suddenly increase.”).

⁹ Professor Flynn is a Professor Emeritus of Political Studies at the University of Otago in New Zealand who conducts research on intelligence testing. After noting “massive” IQ gains of 5 to 25 points in 14 different countries within a single generation, Professor Flynn posited, “The hypothesis that best fits the results is that IQ tests do not measure intelligence but rather correlate with a weak causal link to intelligence.” James R. Flynn, *Massive IQ Gains in 14 Nations: What IQ Tests Really Measure*, 101 PSYCH. BULL. 171, 171 (1987). That is, IQ tests measure abstract problem-solving ability (APA), but that abstract ability does not necessary correlate strongly to one’s competency to survive and succeed in the real world. *Id.* at 188. As Professor Flynn notes, if rising IQ scores really were an indication that Americans were getting significantly “smarter” with each generation, then their SAT scores should be increasing as well. But SAT scores have been declining over the past several generations. *Id.* at 189. He explained,

Thanks to gains on Wechsler-Benet tests, it seemed that those entering American

clinical psychologist today, if they inherited a score from a school psychologist that was ten years obsolete, any competent one would throw that out and regive a test. That I will say flatly.”

In sum, the trial judge’s finding that Dr. Yohman’s 1997 IQ test score was reliable after subtracting ten points was contradicted by the evidence and led to further factual-findings errors, including an error in the ultimate factual finding that applicant is intellectually disabled under *Atkins*.

I.

Applicant was charged with capital murder for fatally shooting twenty-year-old Cristina Castillo while kidnapping her on September 12, 1995. The evidence at trial showed that applicant, along with five other men, planned to rob Cristina and her boyfriend, Hector Alicia, because they thought the two had drugs and money in their apartment. According to one of the conspirators, applicant was the only person armed. He had a 9 mm pistol and

high schools were getting more and more intelligent, and yet they were leaving high school with worse and worse academic skills. Unless nonintellectual traits, such as motivation, study habits, and self-discipline were deteriorating at an incredible rate, how could more intelligent students be getting so much less education? Now the solution is apparent: High school students in 1981 did not necessarily have higher intelligence than their counterparts in 1963, they merely had higher APSA [abstract problem solving ability].

Id. One possible explanation of why IQ test scores rose immediately after WWI and then again during the post-WWII era is that, as nations moved from a relatively agrarian society into the industrial age and then from the industrial age into the technological age, the emphasis on abstract problem-solving increased, but over-all academic achievement, as measured by instruments such as the SAT, did not. As Professor Flynn notes, an IQ test score is probably less predictive of “success” in society than are other measurements of social and academic skills.

grabbed Cristina as she was getting out of her car at the apartment complex. Applicant held Cristina at gunpoint and forced her into a red car occupied by several of the conspirators, who then tied her up with duct tape. Applicant called the other conspirators, who were in a white car, and told them to meet at his mother's house on Palmer Street.

Once at the Palmer Street house, all six men questioned Cristina in an attempt to find the drugs and money. Even though they began to beat her, Cristina continued to deny any knowledge of drugs or money and told them that she was pregnant. Applicant and two others continued kicking and beating Cristina for about fifteen minutes. Finally, they took her to a remote location to abandon her. As one set of conspirators drove off, leaving Cristina with applicant, they heard a gunshot. Applicant later told his cohorts that he had shot her. Cristina's decomposed body was found almost two weeks later in a field. She had been shot three times in the head, and three 9-mm Luger casings were recovered from underneath her body. Police were able to match the shell casings to a 9 mm pistol that Mark Young had snatched from applicant over a month after the murder.

At the punishment phase, evidence of applicant's prior acts of violence was admitted, including evidence of the kidnapping of Mark Young and two little girls at a Chevron station. Evidence showed that applicant was accompanied by two other men, and he was armed and in charge. He made Mr. Young get into the back seat of his own car while applicant drove that car with the two little girls jammed in the front seat. He demanded money from Mr. Young and wanted to know where he lived, but, when the car stopped near some semi-

abandoned apartments, Mr. Young was able to snatch applicant's semi-automatic pistol away from him. Then applicant and his two cohorts ran off.

In a different incident, Frank Condley testified that he was walking from his apartment near the Sherwood Forest Apartments to a convenience store when he saw some men with cocked guns in a nearby parked van. Mr. Condley turned away, but applicant came after him, armed with a .38 or 9 mm gun in each hand. Applicant ordered Mr. Condley to lie down and then shot the prostrate man four times as he begged for his life. He still has three bullets in his body because they were lodged so close to his spine.

Antonio Glenn testified that he lived across the street from applicant during 1995 and sold cocaine to him in the Sherwood Forest Apartments. Applicant would then cut it and resell it for a 50% profit. One time applicant came to Glenn's apartment with a sawed-off shotgun, forced Glenn to undress, tied him up, and held his shotgun to Glenn's head, demanding drugs. When Glenn said that he didn't have any drugs right then, applicant beat him up with the stock of the shotgun.

Albert Garcia testified that applicant knocked on the door of his Sherwood Forest townhouse one night in October 1995 and demanded to be let in. Mr. Garcia refused to open his door and told applicant to leave. Applicant then began banging on the sliding glass patio door. The door broke while Mr. Garcia was calling 911, and applicant came into his bedroom, demanding to know where "the dope" was kept. He left through the front door with another man when Mr. Garcia told him that he was on the phone with the police.

Applicant's sister, Charlotte, testified that he went to Blackshear Elementary School, Brian Middle School, and Yates High School. He was "average" and played a little football and a little baseball while growing up. According to Charlotte, he was a "nerd" because he "read a lot of books, stayed to himself a lot, [and] did a lot of drawing." Applicant and his brother were kind of "spoiled," and "they never went without." Applicant was shy but "he opened up more to older people." As far as she knew, applicant did well in school, but he dropped out when he was seventeen to marry Noella. They had two children, but later divorced. While he was married, applicant sometimes worked for Charlotte's former husband, Luke Ezeh, at Dynamic Battery Exchange.

Mr. Ezeh testified that applicant worked for him "off and on" between 1991 and 1993, when applicant was twenty to twenty-three-years old. Mr. Ezeh said that applicant was a technician and a good, trustworthy worker who could also watch the shop when Mr. Ezeh made deliveries. Applicant was twenty-four when he committed this capital murder.

Applicant's school records showed that he was home schooled during most of third grade because he had tuberculosis, but he kept up with his class work.¹⁰ Applicant's former middle-school teacher, Anne Smith, testified that she taught him Texas history and she remembered him as "such a very well behaved, very nice, very sweet young man." He was shy, but well-liked by both boys and girls. He had "very good home training . . . he was a

¹⁰ This program was called "special ed," but it was based on a medical problem, not academics. He got 2 B's and 2 C's for his work in Math, Spelling, Language, and Reading during the first semester and all B's during the second semester. At the end of the year, his supervising teacher said that applicant "is a very good student. I feel he will do well in 4th grade."

very mannerable child.” In reviewing applicant’s school records, Ms. Smith noted that his conduct was always “[v]ery good to excellent.” She stated that applicant, like most of his schoolmates, “was functioning slightly below grade level.”¹¹ His high school records showed that he functioned at about the 30th/40th percentile in math; “[h]e passed all three sections of the math, the reading, and writing of the TEAMS Test, but he was still seriously below grade level.” Ms. Smith noted that when grades drop in the 9th or 10th grades, it is frequently because of the child’s poor adjustment from middle school to high school. Applicant’s grades dropped dramatically in 9th grade, and he quit school the following year to get married.

Applicant’s mother testified that his father was in construction work, but then turned to “selling drugs.” When applicant, his two sisters, and his brother were young, two men came to their home to rob their father of his money and drugs. The kids hid, but they saw the robbers and their guns. They took applicant’s father’s money and drugs. The kids were outside during a second home robbery with different gunmen looking for drugs and money. Applicant’s mother said that, after applicant was divorced, he started using drugs, mainly cocaine, because he was depressed.

Before trial, Dr. Robert Yohman, a clinical neuropsychologist, interviewed applicant for six hours in the Harris County Jail to evaluate his cognitive and emotional functioning. He was careful to ensure that applicant was not malingering or faking, so he gave him about

¹¹ Applicant’s grades in 6th grade were mainly B’s and C’s, but by 7th grade, most of those B’s had dropped to C’s and D’s. By 8th grade, most of applicant’s grades were D’s.

two dozen tests. Applicant scored a 77 IQ on the WAIS-R, which was “borderline intellectual functioning.”¹² In other achievement tests, applicant functioned in the borderline to mildly deficient range— about the 8th percentile. He did not have a specific learning disorder, but he was mildly deficient in most academic areas, and in the memory test, dealing with the ability to recall a short story, he was “low average to average.” On the word association test, applicant scored in the high average range of the 81st percentile. That is, 81% of the population would score lower than applicant. On the “Trails B” test, applicant scored in the 75th percentile.

Dr. Yohman gave applicant several personality tests, including the MMPI, which indicated that applicant was within normal limits for anxiety and depression, but was a “fairly naive individual, psychologically naive, unsophisticated.” Applicant “wanted to look good . . . wants to be well thought of, be liked.” Dr. Yohman did not, however, find anything in his testing that indicated “any impulse disorder, explosive disorder, anything of that nature.” Although applicant had had a couple of “blows to the head as a youngster,” nothing suggested any focal or localized brain damage. Dr. Yohman noted that applicant had a behavioral change after his wife left him. Overall, applicant fit in the borderline intelligence

¹² However, Dr. Yohman’s 1996 scoring sheet for applicant’s test contains a written notation concerning the Barona Index with an estimated Full Scale IQ of 83. The Barona Index estimates IQ taking into account various demographic and cultural features, including age, education, race, and occupational history.

function, a category that covers about 8% of the population.¹³

Dr. Walter Quijano, a clinical psychologist, also interviewed applicant for an hour and a half in the jail. He gave him the MCMI 2, a personality test, and determined that he was excessively dependent and compulsive. Dr. Quijano said that applicant did not meet the definition of “a full-blown antisocial personality,” but he exhibited some antisocial features. Dr. Quijano originally thought that applicant was “psychologically functioning okay,” but he had not known about the robberies, shootings, and murder that applicant had committed. If applicant had a history of those offenses, then Dr. Quijano believed that he would fit the “antisocial personality disorder” category.

No one at trial intimated that applicant was mentally retarded or intellectually disabled. No one suggested that he was mentally “slow” or had any adaptive deficiencies. His elementary school grades were entirely normal, even though he spent much of his 3rd grade being home-schooled because he had TB. His middle-school history teacher never suggested any intellectual disabilities; she attributed his plummeting grades to the difficulties of making the transition from middle school to high school. Still, applicant passed all three sections of the TEAMS Test in high school. Both applicant’s mother and sister thought he was entirely normal, if a bit “nerdy,” as a child. Applicant worked as a technician in a

¹³ According to the standardized IQ Bell Curve, only the lowest 2.2% of the population score at or below a 70-75 IQ and are considered mentally retarded or intellectually disabled. *See Atkins*, 536 U.S. at 309 n.5 (“It is estimated that between 1 and 3 percent of the population has an IQ between 70 and 75 or lower, which is typically considered the cutoff IQ score for the intellectual function prong of the mental retardation definition.”).

battery-replacement shop, and his ex-brother-in-law left him in charge while he made deliveries.

Neither applicant nor any mental health professional identified applicant as mentally retarded until ten years after he was sentenced to death for capital murder and six years after the *Atkins* decision exempted from execution those who are found to be mentally retarded.

II.

Applicant filed this subsequent writ application on November 17, 2008, the day before he was scheduled to be executed. Because the legal basis for his claim was unavailable on the date he filed his previous application, we granted his motion to stay the execution and remanded his application to the trial court for a live evidentiary hearing on his mental-retardation claim.¹⁴ Under Texas law, applicant is required to prove, by a preponderance of the evidence, that he is intellectually disabled under a three-pronged test: (1) “significantly subaverage general intellectual functioning,” (2) “that is concurrent with deficits in adaptive behavior,” and (3) “originates during the developmental period.”¹⁵ We conclude that the record does not support the trial judge’s factual findings¹⁶ that applicant has proven all three

¹⁴ The Texas Legislature has changed the applicable term from “mental retardation” to “intellectual disability” but the definition is still the same. TEX. HEALTH & SAFETY CODE § 591.003 (7-a), (13). *See supra* note 4.

¹⁵ *Briseno v. State*, 135 S.W.3d 1, 7 (Tex. Crim. App. 2004) (adopting AAMR and Texas Health and Safety Code definitions of intellectual disability).

¹⁶ The trial judge signed applicant’s Proposed Findings of Fact and Conclusions of Law on December 31, 2012, her last day in office. In this case, Applicant’s proposed findings are so adversarial and slanted that they are hard to credit. Many are not supported by the record.

prongs. He has not proven any of those three prongs by a preponderance of the evidence.

Although psychology and psychologists inform the factual decision, they do not determine whether an inmate is exempt from execution under *Atkins*.¹⁷ We must apply our own judgment on the “appropriate ways” to enforce the ultimately legal prohibition on executing mentally retarded offenders.¹⁸ *Atkins* did not conclude that there was a national consensus concerning the *definition* of mental retardation; rather, the Supreme Court concluded that there was a national consensus against execution of those offenders who fit within a given state’s definition of mental retardation, while permitting the states to continue to refine the contours of that definition in their own ways.¹⁹

With the understanding that juries and judges, not psychologists, decide the factual question of whether a particular person is “intellectually disabled” so as to be exempt from the death penalty, we turn to the Texas definition of “intellectual disability.”

¹⁷ See *Ortiz v. United States*, 664 F.3d 1151, 1168 (8th Cir. 2011) (“[P]sychology informs, but does not determinately decide, whether an inmate is exempt from execution.”); see also *Hooks v. Workman*, 689 F.3d 1148, 1172 (10th Cir. 2012) (*Atkins* could have adopted the clinical standard but explicitly declined to do so.”); *Clark v. Quarterman*, 457 F.3d 441, 445 (5th Cir. 2006) (*Atkins* “did not dictate that the approach” to defining mental retardation “must track the approach of the [AAIDD] or the APA exactly”); *United States v. Bourgeois*, Nos. C-02-CR-216 and C-07-223, 2011 WL 1930684, at *24 (S.D. Tex. May 19, 2011) (not designated for publication) (*Atkins* “left the contours of the constitutional protection to the courts”); *Briseno*, 135 S.W.3d at 9 (“Although experts may offer insightful opinions on the question of whether a particular person meets the psychological diagnostic criteria for mental retardation, the ultimate issue of whether this person is, in fact, mentally retarded for purposes of the Eighth Amendment ban on excessive punishment is one for the finder of fact, based upon all of the evidence and determinations of credibility.”).

¹⁸ *Atkins*, 536 U.S. at 317.

¹⁹ See *id.*; see also *United States v. Wilson*, 922 F. Supp. 2d 334, 340 (E.D. N.Y. 2013).

A. “Significantly subaverage general intellectual functioning.”

As we noted in *Ex parte Briseno*, “[s]ignificantly subaverage intellectual functioning is defined as an IQ of about 70 or below (approximately 2 standard deviations below the mean).”²⁰ As we explained, mental health professionals are flexible in their assessment of intellectual disability; sometimes a person whose IQ has tested above 70 may be diagnosed as intellectually disabled while a person whose IQ tests below 70 may not be disabled.²¹ In the new DSM-5, an IQ score is even vaguer and of less critical importance to the diagnosis than in earlier versions of the DSM,²² thus making the “intellectually disabled” diagnosis

²⁰ *Ex parte Briseno*, 135 S.W.3d at 7 n.24 (quoting DSM–IV at 39; see also AMERICAN ASSOCIATION ON MENTAL DEFICIENCY (AAMD), CLASSIFICATION IN MENTAL RETARDATION 1 (Grossman ed. 1983)).

²¹ *Id.* (quoting AAMD at 23); see also *Hall v. Florida*, 134 S.Ct. 1986, 1994-95 (2014) (rejecting State’s position that a firm cut-off IQ test score above 70 disqualifies a capital defendant from offering other evidence of possible intellectual disability; “Florida’s rule disregards established medical practice in two interrelated ways. It takes an IQ score as final and conclusive evidence of a defendant’s intellectual capacity, when experts in the field would consider other evidence. It also relies on a purportedly scientific measurement of the defendant’s abilities, his IQ score, while refusing to recognize that the score is, on its own terms, imprecise.”).

²² American Psychiatric Association, DSM-5 Intellectual Disability Fact Sheet, *available at* <http://www.dsm5.org/documents/intellectual%20disability%20fact%20sheet.pdf>. As the fact sheet explains,

DSM-5 emphasizes the need to use both clinical assessment and standardized testing of intelligence when diagnosing intellectual disability, with the severity of impairment based on adaptive functioning rather than IQ test scores alone. By removing IQ test scores from the diagnostic criteria, but still including them in the text description of intellectual disability, DSM-5 ensures that they are not overemphasized as the defining factor of a person’s overall ability, without adequately considering functioning levels. This is especially important in forensic cases.

It is important to note that IQ or similar standardized test scores should still be included in an individual’s assessment. In DSM-5, intellectual disability is considered to be approximately two standard deviations or more below the

even more of a subjective battle of the experts than it had been formerly.²³

To prove this first prong, applicant relied upon his 1996 WAIS-R IQ score of 77 to establish that he was intellectually disabled by arguing that (1) his score should be lowered five points to account for the SEM or standard error measurement, and (2) his score should be lowered another 5.4 points to account for “the Flynn Effect.” Therefore, what started as an IQ test of 77, with an SEM range of 72 to 82, well within the borderline intelligence category, but outside the mentally retarded or intellectually disabled category, became, according to applicant, an IQ score with a range of 66.6 to 76.6, which he argues satisfies the

population, which equals an IQ score of about 70 or below.

This change in the definition of intellectual disability turns an *Atkins* hearing into that much more of a subjective battle between dueling forensic experts. Thus, factfinders may choose to rely more upon the existence of objective, contemporaneous evidence of a person’s intellectual abilities to assess the reliability of conflicting psychological expert opinions.

This definitional subjectivity is the primary reason why we developed the seven, more objective, *Briseno* factors as a possible guide to assessing the type of intellectual-disability concerns raised by the *Atkins* Court. Of course, those factors are not part of the definition of “intellectual disability,” and trial and appellate courts may ignore some or all of them if they are not helpful in a particular case.

²³ See *Wiley v. Epps*, 625 F.3d 199, 215 (5th Cir. 2010) (noting that most *Atkins* cases involve “essentially a battle of the experts, who gave competing opinions as to [an inmate’s] IQ and intellectual functioning”). Part of the problem of relying too heavily upon the psychological community in determining whether an inmate is mentally retarded under *Atkins* is that the psychological community’s “understanding of mental retardation is evolving. The few short years since the *Atkins* decision has seen change in the definition of mental retardation, renovation of the name of the most prominent advocacy organization, and even abandonment of the very term mental retardation. Adoption of the phrase ‘intellectual disability’ is only the most-recent terminology in the psychological community’s developing understanding of mental retardation.” *United States v. Bourgeois*, 2011 WL 1930684, at *24 n.29 (S.D. Tex. May 19, 2011) (not designated for publication).

initial prong of intellectual disability.²⁴

When we remanded this case for an evidentiary hearing, we ordered the trial judge to evaluate evidence concerning the following four issues:

- (1) the scientific validity and reliability of the “Flynn Effect”;
- (2) whether clinical practitioners who are ordinarily called upon to diagnose mental retardation for purposes outside of the criminal justice system use and apply the “Flynn Effect” to I.Q. test results when making their particularized diagnoses of mental retardation;
- (3) whether the application of the “Flynn Effect” to individual test results is generally accepted scientific procedure in the pertinent professional community outside of the criminal justice system; and
- (4) the known or potential “error rate” of the “Flynn Effect” as it applies to a specific I.Q. test result.²⁵

1. *The “Flynn Effect” exists and is generally considered valid.*

The trial judge heard extensive evidence concerning the “Flynn Effect,” including testimony from Professor Flynn himself. It was generally agreed by all of the experts that the “Flynn Effect” does exist and is valid. Put simply, the “Flynn Effect” refers to the tendency for scores on an IQ test normed for one particular age group on one date to increase

²⁴ The trial judge’s finding number 205 states, in part, that “[w]ith correction for the Flynn Effect, Mr. Cathey’s score on the WAIS-R is a 71.6, and after applying the standard error of measurement, his corrected score falls within the range of mental retardation.” However, even under the most generous view of the SEM and even if courts were permitted to subtract points for the Flynn Effect, applicant’s IQ would fall within the range of 66.6 to 76.6; only 1/3 of this range falls two standard deviations below the average. Because there might be some possibility, however small, that applicant’s “true” IQ could fall below 70, the factfinder may consider other indicia of intellectual functioning, such as school records, in deciding whether applicant has proven this prong by a preponderance of the evidence.

²⁵ *Ex parte Cathey*, No. WR-55,161-02, 2008 WL 4927446, at *1 (Tex. Crim. App. Nov. 18, 2008) (not designated for publication).

when that same test is given to others many years later.²⁶ The aggregate average gain is approximately .3 IQ points per year from the time that an IQ test is originally normed. There is considerable debate as to precisely why such an effect occurs and equally robust debate as to whether that effect is increasing, decreasing, or changing in different populations.²⁷

Although we remanded this case in part to consider the known error rate of the Flynn Effect as applied to a specific test result, we agree with the testifying experts that this is not really an appropriate question because the Flynn Effect deals with IQ test score averages, not individualized scores.²⁸ Although the past average increase had been .3 IQ points per year after an IQ test is formed, there is considerable debate about the appropriateness of that number for all IQ tests (as opposed to simply Wechsler and Benet tests) and even greater debate concerning whether that effect exists at all in the WAIS-III or WAIS-IV versions.²⁹

²⁶ According to the writ-hearing expert witnesses, “norming a test” means that the test is given to a sample group that reflects the demographics of the population for which the test is intended (e.g., English-speaking Americans in 2010), such as age, gender, ethnicity, socioeconomic background, and geographical region, so that if 15% of the population has college degrees, then 15% of the same group used for norming would have college degrees.

²⁷ See generally THE RISING CURVE: LONG-TERM GAINS IN IQ AND RELATED MEASURES (Ulric Neisser ed., 1998) (collecting chapters by psychologists addressing the Flynn Effect and its possible causes). Taken all together, the experts seem to agree that nobody really knows what causes this phenomenon. *Id.*

²⁸ Professor Flynn acknowledged that he reached the .3 number in his research by averaging the rates of increase even though the rates of increase differ depending on the specific test, country, and year, and that .3 is not a static number. He admitted that the rate of increase appears to have slowed for the current generation.

²⁹ Because of that scientific uncertainty, we are highly skeptical that testimony about the Flynn Effect would be admissible when considering the WAIS-III or WAIS-IV IQ tests.

The general notion is that IQ scores on a specifically normed test tend to rise over time, at least in part, because modern societies and cultures have tended to emphasize abstract, problem-solving skills more with each passing generation over concrete, knowledge-based, skills. But test-takers should be normed against their own generational cohort, not against an earlier one.³⁰ Thus, the “Flynn Effect” does not mean that young people today are “smarter” than their parents who are, in turn, “smarter” than their grandparents, it simply means that the questions used on intelligence tests change from one

³⁰ As one expert has explained,

In discussions about FE [“Flynn Effect”] adjustments, the key issue centers on which generation constitutes an appropriate normative reference group for the individual being tested. A person who was born in 1978 and tested in 2010 at age 32 using a current IQ test will be compared with a normative reference group of 30-34-year-olds born between 1976 and 1980. In this case, the person is being compared with the generation to which he or she belongs. If the test used was 20 years old at the time the person was tested, then he or she would be compared with a group of 30-34-year-olds who were born between 1956 and 1960—clearly not the same generation. If generational effects exist—as all contributors to this special issue agree they do—then this is clearly not the optimal normative reference group for this individual. Consequently, an adjustment to the person’s score that takes into account changes in the normative reference group may be appropriate. This example makes clear that the FE is related to changes in the score distribution of the reference sample.

Lawrence G. Weiss, *Considerations on the Flynn Effect*, 28 J. PSYCHOEDUC. ASSESSMENT 482, 489 (2010). This article, along with ten others concerning the existence, significance, consideration, and use of the “Flynn Effect” were compiled in a special issue of the Journal of Psychoeducational Assessment. Some of these articles challenged Flynn’s theory, others agreed with it; some questioned whether the effect will continue into the future, others questioned whether IQ scores should be used at all to determine mental retardation. *See, e.g.*, Robert J. Sternberg, *The Flynn Effect: So What?*, 28 J. PSYCHOEDUC. ASSESSMENT 434 (2010) (concluding that the use of IQ scores for mental retardation determinations is limited and ignores ethical considerations because those scores measure only cognitive intelligence and not the more significant “ethical intelligence”). All of these articles were introduced into evidence at the writ hearing.

generation to another as does the testing environment and the instructions for the test.³¹ The “Flynn Effect” gains simply reflect the obsolete norms of outdated tests.³²

As the expert witnesses explained at the writ hearing,³³ IQ scores are, after all, relative, not absolute, and one’s IQ should be determined by using a scale based on the scores of other test-takers of similar age taking the test at approximately the same time.³⁴ After selecting and testing a representative standardization sample, test developers create a bell curve based on the scores of the representative sample with the average of the scores normed

³¹ The “Flynn Effect” seems to be much more apparent for “fluid” intelligence—abstract reasoning and problem solving—than it is for more concrete, knowledge-based intelligence. As one expert explained:

The third clue [in attempts to understand what cognitive ability is actually rising], which has been discussed above, consists of findings that the scores on culture reduced tests, or tests of fluid intelligence, show an increase twice as large as that observed for tests of learned information, or tests of crystallized intelligence. The increase represents largely an enhancement of people’s ability to solve certain kinds of problems rather than their acquisition of more information from the culture in which they live.

Merrill Hiscock, *The Flynn Effect and Its Relevance to Neuropsychology*, 29 J. CLINICAL & EXPER. NEUROPSYCH. 514, 517 (2007). The author notes that “IQ gains since World War II, according to Flynn, can be attributed to a shift of emphasis from reading, writing, arithmetic, and other ‘disciplined’ learning to ‘on-the-spot problem-solving skills.’ This educational shift seems to be associated with several demographic trends, such as greater urbanization and affluence, decreasing family size, changes in the kinds of work that people do, and the increasing importance of leisure activities.” *Id.* at 520.

³² KAUFMAN, *supra* note 5, at 203; James R. Flynn, *The Mean IQ of Americans: Massive Gains 1932 to 1978*, 95 PSYCHOL. BUL. 29, 32–34 (1984).

³³ Numerous expert treatises and journal articles were introduced into evidence at the writ hearing. They, rather than the experts’ courtroom testimony, are referred to whenever possible as the bench and bar may consult these published and widely available scientific articles without the need to find a copy of the writ hearing testimony in this particular case.

³⁴ KAUFMAN, *supra* note 8, at 130 (noting that the performance of others of the same age or age group on an IQ test at a specific point in time defines a person’s score on the same IQ test).

at 100, meaning that a score of 100 represents “average” performance on the IQ test. Additionally, IQ tests generally have a standard deviation of fifteen or sixteen points.³⁵ And the MR or Intellectual Disability category is approximately two standard deviations below the average, about an IQ score of 70. Approximately two percent of the population falls into this category and approximately two percent fall into the “gifted” category with an IQ score of about 130 or higher.³⁶

In sum, the Flynn Effect, its possible causes, and its meaning have been studied extensively since the 1980s, but it was not until the *Atkins* decision in 2002 that it took on practical significance in state and federal courts. The question for courts is whether psychologists or factfinders should adjust IQ scores for the Flynn Effect in making a determination of intellectual disability under *Atkins*. The answer to that question would seem to depend on whether clinicians adjust IQ scores in their normal working world outside the courtroom.

2. *There is insufficient evidence that clinical practitioners outside the criminal justice system normally use and apply the “Flynn Effect” to IQ test results.*

Although many psychologists agree that the historical data have shown that IQ test

³⁵ *Id.* at 119–23. While Wechsler IQ tests use fifteen points as the standard deviation, Stanford-Binet IQ tests have used sixteen points until recently. *Id.* at 107, 119–23. The Stanford-Binet Intelligence Scales, Fifth Edition now uses fifteen as its standard deviation. GALE H. ROID & ANDREW BARRARA, ESSENTIALS OF STANFORD-BINET INTELLIGENCE SCALES (SB5), Assessment 3 (2004).

³⁶ *What Is an IQ Test? What Is a High IQ Score?*
<http://www.i3mindware.com/what-is-an-iq-test-and-iq-score> (last visited Nov. 4, 2014).

scores on a given type of IQ test have risen on an average of .3 points a year between 1972 and 2002,³⁷ they disagree on whether clinicians normally do or should adjust individual IQ scores in their daily work. In making a determination of intellectual disability under *Atkins*, the factfinder should certainly be aware of how the clinical practitioner makes these determinations in the real world and may follow that procedure,³⁸ unless there are special reasons why that general routine should not be followed in a specific case.³⁹

³⁷ The experts disagree on whether the Flynn Effect continues to exist for the most recent IQ test revisions and, if it does, what effect it still has. See the ten articles contained in the 2010 symposium issue of the *Journal of Psychoeducational Assessment* referred to in note 30.

³⁸ See *Hall v. Florida*, 134 S. Ct. 1986, 1993 (2014) (noting that courts are informed by the work of medical experts and their “learning and skills to study and consider the consequences of the classification schemes they devise in the diagnosis of persons with mental or psychiatric disorders or disabilities. Society relies upon medical and professional expertise to define and explain how to diagnose the mental condition at issue.”). Those clinicians who actually administer IQ tests to a wide range of subjects for a wide variety of reasons are best positioned to know and apply the appropriate professional standards.

³⁹ See, e.g., *Coleman v. State*, 341 S.W.3d 221, 242 (Tenn. 2011) (stating that “[i]n formulating an opinion regarding a criminal defendant’s I.Q. at the time of the offense, experts may bring to bear and utilize reliable practices, methods, standards, and data that are relevant in their particular fields.”). The court explained,

[I]f the trial court determines that professionals who assess a person’s I.Q. customarily consider a particular test’s standard error of measurement, the Flynn Effect, the practice effect, or other factors affecting the accuracy, reliability, or fairness of the instrument or instruments used to assess or measure the defendant’s I.Q., an expert should be permitted to base his or her assessment of the defendant’s “functional intelligence quotient” on a consideration of those factors.

Id. at 242 n.55; see also *State v. Ball*, 2014 WL 2547721, at *36 (Tenn. Crim. App. May 30, 2014) (not designated for publication) (stating that it was following *Coleman* in concluding that a trial court “may reject the application of the Flynn Effect to adjust I.Q. scores based upon evidence of its lack of validity and consider the I.Q. score of 75 as the defendant’s functional I.Q.”) (internal quotation marks omitted); *Jahi v. State*, 2014 WL 1004502, at *106 (Tenn. Crim. App. March 13, 2014) (not designated for publication) (declining to apply Flynn Effect when defense expert “acknowledged that the application of the Flynn Effect was not considered an acceptable practice by either the APA or the AAIDD and that the Wechsler series did not allow for the results to be adjusted pursuant to the

The American Association on Intellectual and Developmental Disabilities (AAIDD) Manual states that “best practices” warrant recognition of the Flynn Effect when older versions of an IQ test are used.⁴⁰ It notes, “In cases of tests with multiple versions, the most recent version with the most current norms should be used at all times. In cases where a test with aging norms is used, a correction for the age of the norms is warranted.”⁴¹ The term used is “warranted,” not “required.” But applicant has failed to offer sufficient data to support a finding that ordinary clinicians in their normal work actually do subtract points from IQ scores to account for the Flynn Effect.⁴²

Flynn Effect. Dr. Bishop stated that, to her knowledge, capital litigation is the only area of the law addressing intellectual disability where the Flynn Effect was being applied.”); *Ledford v. Head*, ___ F. Supp. 2d ___, 2014 WL 793466, at *2-3 (N.D. Ga. 2014) (declining to “apply the Flynn Effect because the phenomenon is not used in clinical practice and the Court was ‘hesitant to apply a theory that is used solely for the purpose of lowering IQ scores in a death penalty context.’”).

⁴⁰ The AAIDD is a professional non-profit association, much like the American Bar Association or American Association of Retired Persons, that advocates for the rights of the mentally impaired and those with developmental disabilities. It does not develop, administer, or score IQ tests.

⁴¹ ROBERT L. SCHALOCK, ET AL., USER’S GUIDE 23 (11th ed., AAIDD 2012). The User’s Guide accompanies ROBERT L. SHALOCK, ET AL., INTELLECTUAL DISABILITY: DEFINITION, CLASSIFICATION, AND SYSTEMS OF SUPPORT (11th ed., AAIDD 2010) (AAIDD Manual).

⁴² Professor Flynn stated that clinicians need not adjust their IQ scores in routine examinations to decide if a child qualifies for extra tutoring or special education because there is no real concern with whether the score is 69, 70, or 71. But he advocates adjusting IQ scores when the death penalty is at stake because then an IQ score may be a matter of “life or death.” He admitted that he was adjusting the data to fit a desired result, but justified doing so because the consequence “might kill somebody.” He would adjust individual IQ scores if there is a benefit at stake “that will keep them alive.” According to one article introduced by applicant at the writ hearing, Professor Flynn advises psychologists to either “select the version of the IQ test that is more likely to yield the desired classification, or they can disregard IQ testing and classify individuals solely on the basis of adaptive functioning.” Merrill Hiscock, *The Flynn Effect and Its Relevance to Neuropsychology*, 29 J. CLINICAL & EXPER. NEUROPSYCH. 514, 525 (2009). This would appear to be result-oriented

Many experts disagree with Professor Flynn’s “correction” of IQ scores.⁴³ Indeed, Dr. Lawrence Weiss, senior psychologist of the Wechsler test group (the drafters of the WAIS-R, the WAIS-III, and the 2008 WAIS-IV), stated that “[a]s the publisher of the Wechsler series of tests, Harcourt Assessment does not endorse the recommendation by Flynn to adjust WAIS-III scores.”⁴⁴ The single most important question relative to “real reasoning at its apogee.

⁴³ Many courts also disagree with him. A few courts adjust IQ scores downward to account for the Flynn Effect, *see, e.g., United States v. Hardy*, 762 F. Supp. 2d 849, 866-68 (E.D. La. 2010) (noting that the Flynn Effect is “well established scientifically” and that adjusting for it is a “best practice”); *United States v. Lewis*, 2010 WL 5418901, at *11 (N.D. Ohio Dec. 23, 2010) (not designated for publication) (recognizing the Flynn Effect “as a best practice for an intellectual disability determination” and adjusting IQ score accordingly), but many more courts have declined to subtract points from IQ scores based on the Flynn Effect. *See, e.g., United States v. Candelario-Santana*, 916 F. Supp. 2d 191, 207-08 (D. P.R. 2013) (collecting cases and concluding that “the Flynn Effect remains highly controversial and many courts have declined to accept its application”); *United States v. Salad*, 959 F. Supp. 2d 865, 872 n.10 (E.D. Va. 2013) (“Because the Fourth Circuit does not necessarily instruct courts to apply an adjustment to account for the Flynn Effect, and because any such adjustments at this juncture would require unsubstantiated speculation, the court declines to apply any Flynn adjustments to the scores in this case.”); *United States v. Jimenez-Bencevi*, 934 F. Supp. 2d 360, 370 (D. P.R. 2013) (“The Flynn Effect remains controversial among scientific experts. Courts of law are not in the business of endorsing one side or the other in a scientific controversy. Instead, we look to ground our decisions on reliable sources. The Flynn Effect is sufficiently controversial as to be unreliable. Under such circumstances, the Flynn Effect has no relevance to our inquiry and we agree with the government’s experts that it should not apply here.”); *Hooks v. Workman*, 689 F.3d 1148, 1170 (10th Cir. 2012) (noting that “*Atkins* does not mandate an adjustment for the Flynn Effect” and that there is no uniform consensus concerning the application of the Flynn Effect in death penalty cases); *see generally*, Geraldine W. Young, Note, *A More Intelligent and Just Atkins: Adjusting for the Flynn Effect in Capital Determinations of Mental Retardation or Intellectual Disability*, 65 VAND. L. REV. 615, 630 (2012) (summarizing the inconsistent treatment of the Flynn Effect in *Atkins* cases).

⁴⁴ Lawrence G. Weiss, *WAIS-III Technical Report: Response to Flynn* (2007) available at: http://images.pearsonclinical.com/images/products/wais-iii/wais-iii_tr_lr.pdf. Although Dr. Weiss was speaking only of the WAIS-III, his thrust was that more modern IQ test development and norming may have slowed or stopped the “Flynn Effect” of rising IQ scores as the tests became “obsolete.” As Dr. Weiss explained, Professor Flynn’s only evidence to support his suggestion that

world” use of the Flynn Effect by ordinary clinicians is whether the IQ test manuals themselves require or recommend that every IQ test be adjusted downward by .3 points per year.⁴⁵ That would be the advice that the ordinary clinician is most likely to follow.

The authors of one recent psychology article in a professional symposium journal concerning the “Flynn Effect” stress that adjusting IQ scores based on the Flynn Effect “does not comport with the standard of forensic psychological practice.”⁴⁶ The authors cite a 2008

WAIS-III scores should be adjusted by 2.34 “is that WAIS-III scores do not fit expectations made based on the Flynn Effect. However, the progress of science demands that theories be modified based on new data. Adjusting data to fit theory is an inappropriate scientific method, regardless of how supported the theory may have been in previous studies.” *Id.* Dr. Weiss elaborated:

There are many reasons why the WAIS-III, SB-5 and DAS-II tests do not show the .3 point per year rise in IQ scores predicted by Flynn including a possible slowing of the effect, better representation of low SES subjects in more recent standardization projects, and construct changes in the newer versions of these tests. As Flynn observes, his effect is not consistent across all subtests. As test developers add or delete subtests when revising existing intelligence test batteries based on newer theories of cognition and brain functioning, the pattern of IQ increases across time will vary from expectations based on Flynn’s original data. Although such construct changes are necessary to advance the field of intellectual assessment, these same changes make it difficult to study changes in intelligence across the generations.

Id. In sum, although the Flynn Effect seems to have been valid, on average, for many prior IQ tests, beginning with the WAIS-III, its existence and dimension is considerably less certain.

⁴⁵ The AAIDD User’s Guide is not a manual for giving or scoring IQ tests, rather its purpose is “to provide that clear understanding of ID [Intellectual Disability] and summarize best practices in the field.” USER’S GUIDE, *supra* note 41, at 1. Like the American Bar Association (ABA), the American Association on Intellectual and Developmental Disabilities is a professional organization. Just as the ABA does not administer the bar exam, the AAIDD does not administer IQ tests.

⁴⁶ Leigh D. Hagan, et al., *IQ Scores Should Not Be Adjusted For the Flynn Effect in Capital Punishment Cases*, 28 J. PSYCHOEDUC. ASSESSMENT 474, 475 (2010); *see also* Roger B. Moore, Jr., Letter to the Editor, *Modification of Individual’s IQ Scores is Not Accepted Professional Practice*, PSYCHOLOGY IN MENTAL RETARDATION AND DEVELOPMENTAL DISABILITIES (American Psychological Association/ Division 33, Washington, D.C.) Fall 2006, at 11, 12 (“If there are factors that lead the psychologist to believe that the scores do not represent an accurate or reliable measure of the individual’s functioning, such issues are delineated in the discussion and interpretation of the

research article about a survey of program directors of APA-approved psychology programs, graduate faculty, and clinicians who were certified school psychologists.⁴⁷ These are the people who actually use IQ tests and score them as a part of their everyday line of work, and they do not adjust for the Flynn Effect in their practices.⁴⁸ The survey authors also found that IQ-test manuals, Social Security Administration reports and manuals, and APA ethical and testing guidelines did not refer to the Flynn Effect or suggest making any adjustments because of it.⁴⁹ Instead, all of these sources recommended that clinicians and psychologists—including forensic psychologists—rely on up-to-date test norms and use

scores; the scores themselves are not changed. Modification of individual scores is not accepted professional practice, for good reason, and should not be introduced into the court as such.”).

⁴⁷ Leigh D. Hagan, et al., *Adjusting IQ Scores for the Flynn Effect: Consistent with the Standard of Practice?*, 39 PRO. PSYCHOL.: RES. & PRAC. 619, 620-21 (2008). Dr. Hagan testified at the writ hearing about his national survey, and that his study reached the following conclusions: (1) adjusting obtained IQ scores and recalculating them on the basis of the Flynn Effect does not represent the conviction and custom in psychology; (2) recalculating an individual’s actual data likely violates the standardization procedures and departs from training practices, prevailing canons, guidelines, most treatises, and test instruction manuals; and (3) noting, in the narrative part of the report, any issues that may compromise the findings, is appropriate (including issues about out-of-date norms).

Dr. Hagan testified that, in a review of 5,000 special education IQ reports, only six mentioned the Flynn Effect and none of those six adjusted the IQ scores. This is potent real-world evidence that the Flynn Effect is an abstract intellectual concept that influences how frequently IQ tests should be renormed and redesigned, but that it is not to be used to “change” a specific person’s IQ test score. Similarly, Dr. Proctor testified that he has reviewed a large number of reports for the Social Security Administration and that he had never seen an individual IQ test report (except in the *Atkins* setting) that mentioned the Flynn Effect.

⁴⁸ *Id.*

⁴⁹ *Id.* at 622-23.

regularly updated IQ tests.⁵⁰ That is precisely what Professor Flynn said at applicant's writ hearing: Do not rely on outmoded IQ tests; instead, retest with the most recent version.⁵¹

There is, however, a certain tension, in death-penalty cases, between the reliability of using the most recently normed IQ test versus the reliability of using a pre-*Atkins*, pre-age-18 IQ test. The former may be discounted for potential malingering and the latter discounted for the "Flynn Effect."

When it is impossible to retest using the most current IQ test available, then factfinders may consider the Flynn Effect and its possible impact on IQ scores generally, just as they may consider the practice effect, potential malingering, the examiner's behavior, and so forth.⁵² These considerations should be noted in the interpretative narrative, but the IQ test score itself may not be changed.⁵³

⁵⁰ *Id.*

⁵¹ Professor Flynn's advice was echoed by applicant's other experts, Dr. Kaufman and Dr. Fletcher, as well as the State's experts, Dr. Proctor and Dr. Hagan.

⁵² AAIDD USER'S GUIDE, *supra* note 41, at 36 ("An IQ score is subject to variability as a function of a number of potential sources of error, including variations in test performance, examiner's behavior, cooperation of the test taker, and other personal and environmental factors.").

⁵³ Dr. Timothy Proctor testified that he recommends "considering the impact of the Flynn Effect in the framework of interpreting the score but not doing a correction or an adjustment to the score." He noted that there are "lots of other contaminants" that do not require an adjustment; in this case, for example, jail conditions and distractions might have artificially depressed his IQ score. Articles introduced at the writ hearing reiterate that a person's IQ score should not be changed to accommodate the Flynn Effect. *See* Robert J. Sternberg, *The Flynn Effect: So What?*, 28 J. PSYCHOEDUC. ASSESSMENT 434, 435 (2010) (the Flynn Effect "is not equally distributed across ability levels. If one were to try to adjust an individual's IQ level by the FE, one would be embarking on a hazardous mission, because the effect varies in magnitude across the distribution of IQs. . . . The FE seems to apply in the aggregate, but it is extremely difficult to apply it in individual

We therefore reject the habeas judge's finding that the evidence shows that the Flynn Effect is used in determining special education benefits and social-security benefits and that clinical practitioners use the Flynn Effect outside of the criminal-justice system. We conclude that the habeas judge erred in changing applicant's IQ score from 77 to 71.6 based on the Flynn Effect. We agree that, taking the SEM into account, applicant's IQ score range is between 72 and 82.

The fact that applicant took an outmoded⁵⁴ version of the WAIS-R in 1996 might tend to place his actual IQ in a somewhat lower portion of that 72-82 range, while the fact that he took the test under adverse circumstances, while in jail and awaiting trial in a capital murder case, might tend to place his actual IQ in a somewhat higher portion of that 72-82 range. Taken altogether, there is no reason to think that applicant's obtained IQ score of 77 is inaccurate or does not fairly represent his borderline intelligence during the developmental

cases."); Stephen J. Ceci & Tomoe Kanaya, *"Apples and Oranges Are Both Round:" Furthering the Discussion on the Flynn Effect*, 28 J. PSYCHEDUC. ASSESSMENT 441, 446 (2010) (concluding that "it is not appropriate to merely subtract 0.3 points for every year that a norm has aged until we know that everyone experiences the same gains on the same subtests and at the same time"); Leigh D. Hagan, et al., *IQ Scores Should Not Be Adjusted For the Flynn Effect in Capital Murder Cases*, 28 J. PSYCHOEDUC. ASSESSMENT 474, 475 (2010) ("Altering obtained IQ scores based on the FE does not comport with the standard of forensic psychological practice, . . . the current state of psychological science—particularly in light of the established variability of individual cases—does not support devising some other score based on the FE and then substituting that score for the one obtained.").

⁵⁴ "Outmoded" in this context means simply that the test was designed and normed several years earlier; it does not mean that there was a newer, "better" test available. In 1996, the WAIS-R was the "best" test available even though it was "outmoded."

stage.⁵⁵ Applicant could have taken the most recently revised and renormed IQ test (the WAIS-IV normed in 2008) if he had wanted to validate or dispute his 1996 IQ score, but he did not wish to, and he refused to allow the State's experts to do so.⁵⁶ Applicant failed to carry his burden to prove that he has "significantly subaverage" general intellectual functioning, the first prong of the three-part test for intellectual disability under *Atkins* and *Briseno*.

B. "Deficits in Adaptive Functioning."

The second prong of the intellectual disability definition is that of significant deficits or limitations in adaptive functioning. Adaptive behavior refers to the ordinary skills that are required for people to function in their everyday lives. Mental retardation or intellectual

⁵⁵ Applicant's school records and TEAMS testing appear to validate the accuracy of that score, while a TDCJ "Service Investigation Worksheet" for a 1998 prison disciplinary hearing indicates that one of the reasons a "Counsel Substitute" was appointed for that hearing was because of an "EA score below 5 and an IQ below 73," would dispute the accuracy of that score. We know nothing more about this TDCJ entry, however, and therefore, given its unknown reliability, will not consider it.

⁵⁶ A question that is not directly before us is whether a capital murder defendant or death row inmate who wishes to assert an *Atkins* claim may rely on expert testimony if he refuses to allow the State's experts to test his IQ and interview him concerning adaptive behavior. Normally, one ought not be able to use an IQ test or an adaptive functioning test as both a sword and a shield. The State argued that the *Lagrone* rule should apply to claims of mental retardation just as it applies to other psychiatric or mental-state defenses. *See Lagrone v. State*, 942 S.W.2d 602, 610 (Tex. Crim. App. 1997) (citing *Soria v. State*, 933 S.W.2d 46, 57–59 (Tex. Crim. App. 1996)) (when the defense plans to introduce testimony based on a psychiatric examination of defendant, the trial court may compel a psychiatric examination by a State's expert, and if the defense introduces expert testimony based on the defense expert's examination, the State may present expert rebuttal testimony); *see also Hernandez v. State*, 390 S.W.3d 310, 321–22 (Tex. Crim. App. 2012) ("When a defendant intends to present mental-health expert testimony, the State is entitled to compel the defendant to undergo examination by the State's expert for rebuttal purposes."). We need not resolve that issue, however, because the resolution of this case does not depend on that issue.

disability has been described as “the failure to carry out everyday activities at the level expected of adults.”⁵⁷ Similarly, the Texas Health and Safety Code defines adaptive behavior as “the effectiveness with or degree to which a person meets the standards of personal independence and social responsibility expected of the person’s age and cultural group.”⁵⁸

However, unlike medicine, education, or social services, criminal law is concerned with what *was* rather than what currently *is*. The point of an *Atkins* hearing is to determine whether a person was mentally retarded during his developmental period and at the time of the crime and therefore ineligible for the death penalty, not whether a person is currently mentally retarded and therefore in need of special services. Because of this, the determination of mental retardation in the *Atkins* context is always complicated by the problems associated with retrospective assessment and the well-known consequence of a diagnosis of mental retardation—exemption from the death penalty. Both experts and those answering questions about a person’s adaptive functioning may exhibit significant conscious or unconscious bias in addressing this issue.

The habeas judge found that applicant proved that he had significant deficits in adaptive behavior. The judge relied almost exclusively on a Vineland Adaptive Behavior

⁵⁷ Gregory Olley, *The Assessment of Adaptive Behavior in Adult Forensic Cases: Part 3, Sources of Adaptive Behavior Information*, PSYCHOLOGY IN MENTAL RETARDATION AND DEVELOPMENTAL DISABILITIES (American Psychological Association/Division 33, Washington, D.C.) Summer 2007, at 3-4.

⁵⁸ TEX. HEALTH & SAFETY CODE § 591.003(1).

Scales test that one of his experts, Dr. Fletcher, administered by telephone⁵⁹ to applicant's sister and his ex-wife. We cannot credit the results of this retrospective test because

- (1) the Vineland test was not designed to be administered retrospectively decades after the relevant time frame—here, when applicant was 18 or younger—and long after the reporters had significant daily contact with applicant;
- (2) the Vineland reporters—applicant's sister and his former wife—were highly motivated to misremember his adaptive abilities from some ten to twenty years earlier, knowing that a finding of intellectual disability would make him exempt from the death penalty;⁶⁰ and
- (3) the adaptive behavior applicant's sister reported to the expert as part of the Vineland test was contradicted by her trial testimony (before *Atkins* had been decided and any issue of mental retardation had arisen) that applicant was “average,” “nerdy,” and read books all the time.⁶¹

No one who testified at trial suggested that applicant was intellectually disabled or

⁵⁹ The Vineland is normally administered in a face-to-face interview with the reporters. In this case, applicant's expert admitted that the reporters knew that he would be calling them to conduct an interview about applicant's adaptive behavior while growing up because applicant's lawyers had called them and prepared them for his telephone interview.

⁶⁰ Applicant's sister married and moved out of the family home when applicant was just twelve. Applicant's former wife had married him when he was fifteen and left him about three or four years later. Neither had had significant personal contact with applicant in recent decades, although applicant corresponded regularly with his sister from prison.

⁶¹ Applicant's sister's trial testimony is consistent with applicant's present level of comprehension, reading, and writing abilities in prison. In her post-*Atkins* Vineland interview, applicant's sister says that applicant “believed anything that he was told and would do things—if he watched Spiderman, he believed that he could fly from buildings.” But applicant's handwritten prison letters show a sophisticated writer who certainly knows the difference between “play-acting” and reality, and indulges in “play-acting” as seduction. For example, in one letter to pen-pal Amanda Grant, applicant claims that he won \$240 million in the Texas Lottery, but then he tells her, “Alright, alright, I didn't win 240 million, it was only 15 million. © Seriously, though, if I did hit the Jackpot of \$240 million, I bet your pretty little ass would say something like ‘Wayne, Baby, you know it was always meant for me and you to be together’ . . . Yeah, yeah, I know, I need to stop being silly. But don't worry, if I ever win, I will most definitely take care of my girl.”

suffered from adaptive deficiencies. It is difficult to credit that a developmental intellectual disability can lie dormant and undiscovered for thirty-seven years and then spring full-grown, like Minerva from Zeus's forehead, only when that person would be exempted from the death penalty if found so disabled.

A 2008 affidavit filed by applicant's sister stated that she was nine years older than applicant and left home when he was about twelve. She stated that applicant did not get along with his father and, when his father asked applicant to do something, he "would often be very slow at doing it." Applicant never helped her with household chores unless she asked him to do so, and she had to teach him to use the microwave and clean the house.⁶²

By the time of the 2010 interview with Dr. Fletcher, she remembered that she had to tell applicant "over and over" to do something, that he was easily distracted, that he rarely initiated conversations (but his speech was clear and understandable), that he did not know his telephone number, and that she "thought" he had a sixth-grade reading level. She told Dr. Fletcher that her former husband never gave applicant any responsibility at the battery-replacement shop because he would "mess it up," but her husband had testified at trial that he often left applicant, his technician, in charge of the shop when he made deliveries because applicant was a good, trustworthy worker. Applicant's sister told Dr. Fletcher that applicant was "bullied" at school and had no friends, but that contradicted the trial testimony of

⁶² This behavior, although it might be indicative of an intellectual disability, is also consistent with that of twelve-year-old boys who are of average or above average intelligence. Pre-teens and teenagers do not like to be told to "take out the papers and the trash, yakety-yak."

applicant's teacher who said that he was well-liked by his classmates and got along with everyone.

Applicant's former wife told Dr. Fletcher in a 2010 interview that she had to show applicant how to wash clothes, cook, and do chores around the house. She was "still sort of angry" about how he wouldn't help her much and about the friends that he "hung out" with.

Based on his telephone interview with applicant's former wife, Dr. Fletcher scored applicant with a 61 in communication, 61 in daily living, and 60 in socialization. Based on his telephone conversation with applicant's sister, he scored applicant with a 69 in communications, 68 in daily living, and 66 in socialization. All of these scores are consistent with the presence of mild mental retardation.

Dr. Proctor, the State's expert, said that he would put very little stock in a retrospective Vineland test that asked applicant's family members to think back to his behavior eighteen to twenty-six years earlier. Furthermore, there were issues of potential bias in giving the Vineland test to applicant's family members who had a motive to underestimate his abilities and activities.⁶³ Further, Dr. Proctor said that clinicians question the validity of any retrospective use of a formal instrument such as the Vineland Scale

⁶³ The Fifth Circuit has commented on the potential bias of an inmate's relatives in attempting to make a retrospective behavior assessment. *Clark v. Quarterman*, 457 F.3d 441, 447 (5th Cir. 2006) (noting that state court had found an adaptive assessment based on the inmate's self-reporting coupled with his ex-wife's memories about what he could and could not do nine years earlier "unreliable because it did not account for the incentive of Clark and his ex-wife to misreport Clark's adaptive skills").

because the norms were not designed for doing this kind of backward-looking analysis and looking to behavior more than a decade earlier.⁶⁴ The record does not support the habeas judge’s uncritical acceptance of Dr. Fletcher’s opinion concerning applicant’s adaptive deficits based on the Vineland test.⁶⁵

⁶⁴ Experts in other *Atkins* cases have expressed the same concern. See, e.g., *United States v. Montgomery*, ___ F. Supp. 2d ___, 2014 WL 1516147, at *12, 52 (W.D. Tenn 2014) (“Dr. Marcopulos delivered a persuasive argument for why the Vineland Adaptive Behavior Scales (“VABS”) administered by Dr. Reschly in this matter are unreliable based on their discrepant scores and retrospective application, and thus, why the Court must examine other sources of evidence to consider Defendant’s adaptive functioning”; noting that expert concluded that “these Vineland ‘scores are not reliable, and I don’t feel that I can trust them as being reliable indices of [Defendant’s] adaptive functioning because they are not reliable within the person, they’re not reliable across the persons, and the test was administered in an unstandardized way using retrospective data.”); *United States v. Jiménez–Benceví*, 934 F. Supp. 2d 360, 372 (D. P.R. 2013) (criticizing a “fundamentally unreliable” VABS administration to the defendant’s sister, who “had a clear incentive to provide answers that were helpful to her brother” and derived from memories that “were at least ten years old, raising doubts about their reliability.”); *United States v. Candelario-Santana*, 916 F. Supp. 2d 191, 215-16 (D. P.R. 2013) (testifying expert described the use of retrospective use of Vineland test “controversial”); *Thorson v. State*, 76 So.3d 667, 673 (Miss. 2011) (trial court, in addressing *Atkins* mental-retardation case, found “the application of retrospective Vineland tests unreliable and unpersuasive”); Mark Tasse, *Adaptive Behavior Assessment and the Diagnosis of Mental Retardation in Capital Cases*, 16 APPLIED NEUROPSYCHOLOGY 114, 120 (2009) (“It should be noted that there is no research available examining the reliability or error rate of adaptive behavior assessments obtained retrospectively. At issue is the respondent’s ability to correctly recall from memory the assessed individual’s actual performance. Memory degradation is a real issue and we do not have any solid research regarding the forgetting curve regarding someone’s recollection of another person’s adaptive behavior.”). But see *Wiley v. Epps*, 625 F.3d 199, 216-18 (5th Cir. 2010) (recognizing that the authors of the Vineland test express that retrospective interviews are permissible in certain circumstances).

⁶⁵ One of the *Briseno* factors asks whether “those who knew the person best during the developmental stage—his family, friends, teachers, employers, authorities—think he was mentally retarded at that time, and, if so, act in accordance with that determination?” 135 S.W.3d at 8. Indeed, close family and friends, as well as teachers, are the most likely to contemporaneously spot a developmental disability, express concern about it, and to act upon that determination. What matters is what family and friends thought of the person *during* the developmental period, not what they “remember” when they know that their retrospective memories of disabilities and limitations may exempt their loved one from the death penalty. We would be expecting too much of human nature

Even if the Vineland had been administered with reliable subjects reporting on their contemporaneous knowledge of applicant's behavior, the Vineland would be only one part of a person's overall adaptive behavior profile. "[T]he process of assessing adaptive behavior, particularly on a retroactive sense, 'is a matter of drawing information from many sources, all of which are imperfect.'"⁶⁶ Given the vague and amorphous nature of the definition of adaptive behavior in the relevant statutes and treatises, courts have adhered to the "relative consensus that the best way to retroactively assess [an inmate's] adaptive functioning is to review the broadest set of data possible, and to look for consistency and convergence over time."⁶⁷ A significant impairment in adaptive behavior may be thought of

if we thought that a mother or other loved one, knowing that her memories suggesting intellectual disability would save her son from execution, would resolutely assert that he was perfectly normal in every respect. Clinical psychologists must take into account both "cognitive bias" and "confirmation bias." See R. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. OF GEN. PSYCH., 175, 177 (1998) (explaining that people tend to seek information that they consider supportive of favored hypotheses or existing beliefs and to interpret information in ways that are partial to those hypotheses or beliefs; conversely, they tend not to seek and perhaps even to avoid information that would be considered counterindicative with respect to those hypotheses or beliefs and supportive of alternative possibilities); see also Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U.L. REV. 632 (1999) (providing an overview of findings from cognitive psychologists and decision theorists suggesting that humans frequently behave in nonrational ways, and that these "cognitive biases" are largely incapable of being unlearned).

For these reasons, we cannot accept, at face value, the habeas judge's finding 126 that "the affidavits submitted by Mr. Cathey's family members [are] reliable and indicative of adaptive behavior deficits."

⁶⁶ *United States v. Candelario-Santana*, 916 F. Supp. 2d 191, 216 (D. P.R. 2013) (internal citation omitted) (quoting J. Gregory Olley, *The Assessment of Adaptive Behavior in Adult Forensic Cases: Part 2*, PSYCHOLOGY IN MENTAL RETARDATION AND DEVELOPMENTAL DISABILITIES (American Psychological Association/ Division 33, Washington, D.C.) (Fall 2006)).

⁶⁷ *Candelario-Santana*, 916 F. Supp. 2d at 216; see also *Ex parte Butler*, 416 S.W.3d 863, 874-75 (Tex. Crim. App. 2012) (Cochran, J., concurring) ("Assessing adaptive deficits in a

as “the extent to which the individual has required assistance to carry out age-appropriate activities.”⁶⁸

The best source of retrospective information concerning adaptive behavior during the developmental period is usually school records. Such records provide an objective, unbiased documentation of a person’s abilities at the most pertinent time—a time at which mental retardation or intellectual disability is most likely to be diagnosed if it exists.

Applicant’s school records show that he was performing above grade level during the third grade when he was home-schooled. His grades that year started with two B’s and two C’s, but he ended the year with straight B’s.

Applicant was always placed in regular classes and generally received passing grades. He made a B in reading lab in the 6th grade, a 72 in Algebra I in the 7th grade, a 72 in physical science, a 70 in history, an 83 in World History, and a 68 in English. In the 9th grade, he

retrospective *Atkins* hearing is an extremely difficult task. First, there is a tremendous incentive for those closest to the defendant to remember him as being deficient. Because a finding of mental retardation will prevent imposition of a death sentence, it is understandable that those who wish to spare the defendant’s life recall and focus on previously unnoted deficits or downplay competencies, consciously or otherwise. Second, the guidelines for assessing adaptive deficits are so vague and subjective that beauty frequently is in the eye of the beholder. In the context of *Atkins* hearings, experts routinely disagree about which behaviors to focus on and what significance different behaviors have. . . . It was partly for this reason that we adopted the *Briseno* factors to assist the factfinder—both the trial judge and this Court in the context of habeas cases—in considering pre-existing objective data that has not been collected for the sole purposes of deciding the question of mental retardation in the context of an *Atkins* hearing. Those factors focus on the defendant’s behavior and competency in ‘the real world’ before people are seeking specific evidence for (or against) a finding of mental retardation that would bar the defendant’s execution.”).

⁶⁸ Olley, *supra* note 57, at 3-4.

passed all three sections of the standardized TEAMS test (a test that mentally retarded students were usually exempt from taking in the late 1980's). Applicant's former middle school history and homeroom teacher saw him every day. She thought that he functioned "slightly" below grade level, but she never suggested that he was intellectually disabled. Applicant was well behaved, liked by other students, and got along well with everyone. She felt that applicant's falling grades (and his eventual dropping out) were the result of not making a smooth transition to high school. All in all, this is not the academic portrait of an intellectually disabled person.

And the inventory of applicant's death-row cell appears to validate his middle school teacher's assessment. Shortly after applicant filed his *Atkins* claim of mental retardation, the contents of his cell were photographed and inventoried. Those contents are not typical of a person who is intellectually disabled:

- Applicant's cell contained numerous books; a copy of *The Echelon Vendetta* by David Stone was open and face-down on applicant's bed; other books included *Tactics and Strategy of Chess*, *The Complete Jewish Bible* (including a bookmark with the word "redundant" written on it); Harper Collins *Spanish Dictionary*; *The Audacity of Hope* by Barack Obama; *AIDS in America*, by Susan Hunter; *Mein Kampf*, by Adolf Hitler; *The Pocket Oxford English Dictionary*; *The Source*; *Larousee Concise Dictionary*; *Great Speeches by African Americans*; *A Call to Spiritual Reformation*; and *Tom Clancy's Ghost Recon*, by David Michaels (with applicant's name and TDCJ number handwritten on its inside cover).
- Applicant also had an Amazon.com invoice addressed to him, listing the books *The Looking Glass Wars*, *The Looking Glass Wars—Book Two*, *Seeing Redd*, and *ArchEnemy*, all to be shipped to applicant at the Polunsky Unit.
- A composition book containing approximately 80 handwritten names and addresses of his pen pals and other correspondents.

- A TDCJ Offender Grievance Form containing applicant's handwritten name and TDCJ number with his handwritten grievance complaining that "within the last few years and essentially within the previous months the quality of food served has deteriorated drastically to a level on the verge of indecency."⁶⁹

An unrelated property inventory of applicant's cell on March 27, 2009, listed the following items: 55 magazines, 12 books, stamps, ink pens, tables, headphones, and a game board. Although some mentally retarded persons try to cover up their disabilities, the notion of a death-row inmate keeping 55 magazines and 12 books in his cell as "cover," as well as spending his scarce financial resources ordering more books from Amazon.com, is inconsistent with a mentally retarded person attempting to cover up his disability.

Applicant is not only a prison reader, he is a prison writer. One pen-pal letter, dated October 22, 2009, to a woman in Belgium states, "As for myself, well, yesterday after I found out that Bobby Woods had another execution date it really troubled my spirit because he and I basically have similar claims."⁷⁰ In a letter to Meg Harper in the United Kingdom, applicant writes, "Get together and draft up a letter addressing the injustice of the D/P, and lets send it to the U.S. attorney general Eric Holder and the president[.]" He also recounts the number of "blacks," "Mexicans," and "whites" who had been subject to "legal lynchings here in Texas," and states, "Now I elucidated this because Ruth felt like it would be a good idea

⁶⁹ Four other grievance forms, dated after applicant had filed his *Atkins* claim, contained the notation "assisted by" other inmates. It does not appear that applicant needed any such assistance in filling out forms until *after* filing his *Atkins* claim. Likewise, he apparently never needed assistance in writing letters to his pen pals discussing his thoughts on the death penalty and his pending *Atkins* claim.

⁷⁰ The evidence at the writ hearing showed that Bobby Wayne Woods had made an *Atkins* claim that also relied on the Flynn Effect.

to write the Obama administration to address the issue of the death penalty. And I agree. But the voices from the people on the outside will have a more powerful effect when injustice is declared, than when it comes from those who are incarcerated.” Another of his pen-pal letters inquires,

And speaking of news, what is your opinion of the racial incident that transpired with Professor Gates a few weeks ago? Now I did like the fact that the ole racist ass cop lied and falsified his police report. But I did find it kind of funny that President Obama offered to have a beer with both guys at the White House!⁷¹

In a letter to Sari Kauppinen in Finland, applicant gives a detailed description of reading *The Gates of Rome*, about Julius Caesar.

Dr. Proctor testified that applicant’s letter to Amanda Grant, instructing her on how to get an I-60 form for visitation, showed that applicant understood his environment and how to use forms, and that he could solve a problem using multiple steps. In another letter to Ms. Grant, applicant describes his upcoming January 2010 *Atkins* hearing and says, “So my lawyers are interviewing doctors, and others that may testify on my behalf as well as collecting medical and school records that are needed.” In a letter to *The Prison Journal*, applicant stated that he wanted to submit two poems and a drawing that he hoped the journal

⁷¹ Applicant is referring to the Boston incident in which Harvard Professor Henry Louis Gates was arrested at his home based on a 911 call that a burglar was breaking into it. That incident, in July 2009, received considerable media attention. See *Wikipedia, Henry Louis Gates Arrest Controversy*, http://en.wikipedia.org/wiki/Henry_Louis_Gates_arrest_controversy (last visited Oct. 6, 2014).

would publish.⁷²

After examining more than 100 letters written by applicant,⁷³ Dr. Proctor testified that these letters showed that applicant was aware of current events, capable of giving sound advice, capable of planning and abstract thinking, has political awareness, is concerned about how the death penalty is applied, and has ideas addressing the issue. According to Dr. Proctor, applicant uses humor, speaks in the abstract, talks about what he wants, expresses his feelings, and narrates events in his life. These letters demonstrate applicant's normal conceptual abilities and social interactions. We therefore cannot accept the habeas judge's findings that applicant had (1) "communicative deficits" and "difficulties expressing himself" based on his family members' recent recollections; (2) "failed to manage his money," in part because he overspent his inmate trust account at the commissary for "several purchases"; (3) "limited functioning in reading and writing," despite his vast wealth of reading materials and handwritten letters in his cell.

⁷² The P.U.R.E. Report Newsletter of June 2011, published one of applicant's four stanza poems, the first stanza of which reads as follows:

Bombarded by the cultivation
to ensnare a phantom destiny
of a parents dream lost
to the adversity of change.
Now Precious Angels of a cradle's caress
are forgotten, as their wrath of heaven
cast out its rebellious demons . . .

⁷³ These one hundred letters contain very few spelling errors, although the "punctuation police" might well suggest more commas and hyphens. All of the letters are intelligent, coherent, and consistent. This man intends to communicate with great grace, and he succeeds.

A TDCJ guard, Leah Madison, testified that applicant gave her a handwritten letter that began, "Hello Sunshine," described applicant's attraction to her, and included the following: "Because since the first several time[s] we initially came in contact with each other, I felt a sense of a kindred spirit between us. And I'm sure you can relate to what I speak of, simply because of the compassionate, gentle, loving, and caring attributes, that we both have in common." Ms. Madison reported the letter to the proper authorities and applicant was moved to a different pod and level. Applicant told her that he didn't think she would turn him in for writing the letter, but that he understood and knew the consequences. This letter demonstrates applicant's well-developed writing and reasoning abilities, although it also demonstrates his chutzpah and penchant for flouting the rules.

Speaking of flouting the rules, applicant participated in a notorious 1998 prison break-out shortly after he was sent to death row. Applicant was assigned as a sewing machine operator in the garment factory. He and several other inmates dyed some clothes to look like prison guard uniforms, left paper maché dummies in their cell bunks, and scaled the inner perimeter fence at the Ellis Unit. With the exception of the one inmate who got over the fence and drowned in a creek, they all stopped when guards began shooting at them. Dr. Proctor testified that the prison-escape plan contained some "elaborate" elements and that prisoners organizing a daring escape would not bring along a mentally retarded inmate.

Applicant is also an active member of P.U.R.E. (Panthers United for Revolutionary Education), a group associated with the Black Panthers. The P.U.R.E. Newsletter of

December 2010, contained an article written by applicant titled *The Echolon Privilege*, arguing that juries find police officers “not guilty” of murder or “felony brutality” because

[m]any of us in society have been indoctrinated with trusting those in authority and placing them on a high level of esteem. Therefore a common belief have been embedded in our subconscious that if we are good law abiding citizens, then we have nothing to fear from law enforcement officials. So when a jury encounters a situation where a police officer has used force (deadly or otherwise) their sympathy gravitates to the officer.

One may agree or disagree with applicant’s position, which he goes on to explain at great length, but it is surely cogently articulated. That newsletter also states, “Panthers United for Revolutionary Education, founded by Eric Cathey, a Texas death row Prisoner,” and contains a picture of applicant along with his TDCJ contact information.

Some psychologists argue that factfinders should not consider prison behavior in assessing whether a death row inmate is intellectually disabled because prison is such a highly regimented society in which inmates are required to perform rote and simple activities.⁷⁴ But courts should not become so entangled with the opinions of psychiatric

⁷⁴ AMERICAN PSYCHIATRIC ASSOCIATION, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 38 (5th ed. 2013) (“DSM-5”) (“Adaptive functioning may be difficult to assess in a controlled setting (*e.g.*, prisons, detention centers); if possible, corroborative information reflecting functioning outside those settings should be obtained.”). *See, e.g., Holladay v. Allen*, 555 F.3d 1346, 1358 n.16 (11th Cir. 2009) (“Both experts agreed that Holladay’s adaptive functioning cannot be accurately assessed now because he has spent over 17 years in prison, a highly restricted and restrictive environment.”); *Thomas v. Allen*, 614 F. Supp. 2d 1257, 1284 n.67 (N.D. Ala. 2009) (“The constraints of a maximum-security prison environment also limit the diagnostician’s ability to assess the subject’s adaptive skills consistently within the AAMR definition.”); *see also Thorson v. State*, 76 So.3d 667, 672 n.8 (Miss. 2011) (“Experts for each side agreed that being on death row for twenty years could have had an effect, either positively or negatively, on . . . adaptive functioning.”).

experts as to lose sight of the basic factual nature of the *Atkins* inquiry: Is this person capable of functioning adequately in his everyday world with intellectual understanding and moral appreciation of his behavior wherever he is? Or is he so intellectually disabled that he falls within that class of mentally retarded inmates who are exempt from the death penalty? In that inquiry, we should not turn a blind eye to the inmate's ability to use society and his environment to serve his own needs. And sound scientific principles require the factfinder to consider *all* possible data that sheds light on a person's adaptive functioning, including his conduct in a prison society, school setting, or "free world" community.⁷⁵

Some psychologists also say that factfinders should not consider a person's strengths,

⁷⁵ See *United States v. Montgomery*, ___ F. Supp. 2d ___, 2014 WL 1516147, at *49 (W.D. Tenn. 2014) (noting that some psychologists decline to give weight to an inmate's behavior in jail or prison in assessing mental retardation, but concluding that "[t]he fact that post-incarceration observers of Defendant's adaptive behavior would be inadequate reporters for a standardized adaptive behavior scale does not mean that all information regarding Defendant's post-incarceration behavior should be ignored entirely."). In *Montgomery*, the federal district judge noted that one expert in that case

disagrees with the statements in the AAIDD User's Guide instructing examiners not to consider past criminal behavior in their assessment of adaptive functioning. According to Dr. Welner, "the essence of an ethical practice of forensic psychiatry is that you don't pick and choose your data. You rely on all available sources of data, ... the idea of just ignoring behavior altogether is something that has no foundation in the practice of forensic psychiatry." He further testified that he disagrees with the User's Guide's statement that diagnosis of MR/ID is not based on a person's street smarts, behavior in jail, or criminal adaptive functioning.

Id. at *49 (record citations omitted). Thus, the district court refused to disregard the inmate's "criminal and post-incarceration behavior that may lend support one way or another to Defendant's adaptive functioning profile." *Id.* at *50. The *Montgomery* judge noted that this was the approach of some federal courts as well, including the Fifth Circuit. *Id.*; see *Clark v. Quarterman*, 457 F.3d 441, 447 (5th Cir. 2006) (relying on evidence that inmate's "behavior in prison casts serious doubts on his claims of adaptive limitations as evidence collected from his cell" showed handwritten letters, complaints, diet plans, notes about the effects of various chemicals, handwritten puzzles "including the decipherment of several extremely complicated codes").

but only his weaknesses, when deciding the question of intellectual disability.⁷⁶ Most courts, however, consider *all* of the person’s functional abilities, those that show strength as well as those that show weakness.⁷⁷ For example, it would seem foolhardy to say that a person who has obtained a graduate law degree (demonstrating his conceptual abilities), who is a television talk-show host (demonstrating his social skills), but who simply cannot learn to drive properly and has multiple automobile accidents (demonstrating a limitation in practical skills), meets the adaptive-deficits prong of intellectual disability by ignoring all of his educational and social strengths and focusing exclusively on his deficiencies.

Given the entire body of evidence taken from the trial and the habeas hearing, including applicant’s school records and the death-row cell exhibits of his pen-pal letters and P.U.R.E. articles and poems, we conclude that applicant has failed to prove, by a preponderance of the evidence, that he suffers from significant adaptive deficits or limitations. We must therefore also conclude that applicant did not establish the third and final prong of intellectual disability—its onset during the developmental period. If applicant

⁷⁶ See AAIDD Manual, *supra* note 41, at 94 (advocating an approach that “focuses on the individual’s limitations”).

⁷⁷ See *Hooks v. Workman*, 689 F.3d 1148, 1172 (10th Cir. 2012) (rejecting defendant’s contention that *Atkins* requires courts to focus solely on a person’s limitations, and concluding that adaptive functioning means, “What is a given defendant able and unable to do? Both strengths and deficiencies enter into this equation because they make up the universe of facts tending to establish that a defendant either has ‘significant limitations’ or does not. Not only does *Murphy* not require the [state court] to focus on deficiencies to the exclusion of strengths but—most relevant to our inquiry here—neither does *Atkins*”; relying, in part, on defendant’s prison letters in concluding that he did not suffer adaptive deficits under *Atkins*).

has failed to prove that he is intellectually disabled, he clearly did not prove that he was intellectually disabled before the age of approximately eighteen. For these reasons we reject applicant's *Atkins* claim and deny relief on his subsequent application for a writ of habeas corpus.

Delivered: November 5, 2014

Publish