No. 24-872
CAPITAL CASE

IN THE

Supreme Court of the United States

JOHN Q. HAMM, COMMISSIONER
OF THE ALABAMA DEPARTMENT
OF CORRECTIONS,
Petitioner,

v.

JOSEPH CLIFTON SMITH,
Respondent.

On Writ of Certiorari to the United States Court of Appeals for the Eleventh Circuit

BRIEF OF *AMICI CURIAE* THE AMERICAN ASSOCIATION ON INTELLECTUAL AND DEVELOPMENTAL DISABILITIES (AAIDD), THE ARC OF THE UNITED STATES, THE BAZELON CENTER FOR MENTAL HEALTH LAW, AND THE NATIONAL DISABILITY RIGHTS NETWORK (NDRN) IN SUPPORT OF RESPONDENT

1117 Stanford, N.E.
Albuquerque, NM 87131
delpha@law.unm.edu
(505) 277-2146

ANN M. DELPHA
Counsel of Record
JAMES W. ELLIS
Counsel for Amici Curiae

September 24, 2025

WILSON-EPES PRINTING CO., INC. – (202) 789-0096 – WASHINGTON, D.C. 20002

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Federal Cases

State Cases

Other Authorities

INTERESTS OF *AMICI*[1]

THE AMERICAN ASSOCIATION ON INTELLECTUAL AND DEVELOPMENTAL DISABILITIES ("AAIDD"), founded in 1876, is the nation's oldest and largest organization of professionals in the field of intellectual disability. Through its professional journals, conferences, and book publishing, AAIDD works diligently to advance scientific understanding of intellectual disability. Primarily focused on clinical, psychological, scientific, educational, and habilitative issues, the Association also has a longstanding interest in legal issues that affect the lives of people with intellectual disability. AAIDD has appeared as *amicus curiae* in this Court in a variety of cases involving mental disability, including cases as diverse as *City of Cleburne v. Cleburne Living Center, Inc.*, 473 U.S. 432 (1985), and *Atkins v. Virginia*, 536 U.S. 304 (2002). AAIDD has formulated the most widely accepted clinical definition of intellectual disability. Both as the formulator of the clinical definition of intellectual disability, and as an interdisciplinary membership organization concerned with maintaining appropriate professional standards in the diagnosis of intellectual disability, AAIDD and its members have a strong

---

[1] This brief was written entirely by counsel for *amici*, as listed on the cover. No counsel for a party authored this brief in whole or in part, and neither counsel for a party nor any party made a monetary contribution intended to fund the preparation or submission of this brief. No person other than *amici curiae*, the members of the organizational *amicus*, or their counsel made a monetary contribution to the preparation or submission of this brief.

interest in the manner in which *Atkins* claims are evaluated by courts.

THE ARC OF THE UNITED STATES ("The Arc"), is the largest national community-based organization advocating for and serving persons with intellectual and developmental disabilities (IDD) and their families. Founded in 1950, The Arc has nearly 600 state and local chapters. The Arc seeks to promote and protect the civil and human rights of people with intellectual and developmental disabilities and to actively support their full inclusion and participation in the community.

THE BAZELON CENTER FOR MENTAL HEALTH LAW is a national public interest organization founded in 1972 to advocate for the rights of individuals with mental disabilities. Through litigation, legislative and administrative advocacy, and public education, the Bazelon Center advances equal opportunity for individuals with mental disabilities in all aspects of life, including in employment, education, housing, health care, community living, voting, and family rights, as well as in the criminal justice system.

THE NATIONAL DISABILITY RIGHTS NETWORK ("NDRN") is the non-profit membership organization of the federally mandated Protection and Advocacy (P&A) and Client Assistance Program (CAP) agencies for individuals with disabilities. The P&A and CAP agencies were established by the United States Congress to protect the rights of people with disabilities and their families through legal support, advocacy, referral, and education. There are P&As and CAPs in all 50 states, the District of Columbia, Puerto Rico, and the U.S. Territories

(American Samoa, Guam, Northern Mariana Islands, and the US Virgin Islands), and a P&A and CAP affiliated with the Native American Consortium which includes the Hopi, Navajo, and San Juan Southern Paiute Nations in the Four Corners region of the Southwest. Collectively, the P&A and CAP agencies are the largest provider of legally-based advocacy services to people with disabilities in the United States.

SUMMARY OF ARGUMENT

Intellectual disability is a condition that is diagnosed using both quantitative and qualitative data, as the diagnosis of the condition "is intended to reflect a clinical judgment rather than an actuarial determination." American Association on Intellectual and Developmental Disabilities, *Intellectual Disability: Definition, Classification, and Systems of Supports* 40 (11th ed. 2010) [hereinafter *AAIDD-11*]. Clinicians with in-depth knowledge of the range of intellectual and adaptive functioning manifested by human beings have been carefully trained to conduct evaluations for this condition. Intellectual disability is defined using three diagnostic criteria or 'prongs': significant limitations in a person's intellectual functioning (Prong 1), and significant limitations in adaptive behavior (Prong 2), both of which began prior to adulthood (Prong 3). American Association on Intellectual and Developmental Disabilities, *Intellectual Disability: Definition, Diagnosis, Classification, and Systems of Supports* 1 (12th ed. 2021) [hereinafter *AAIDD-12*].

Over many decades, clinical experts have developed a robust clinical framework that must be used in the diagnostic process. *See, e.g.*, *AAIDD-12*;

American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders* 37 (5th ed. text rev. 2021) [hereinafter *DSM-5-TR*]; *APA* [American Psychological Association] *Handbook of Intellectual and Developmental Disabilities* (Laraine Masters Glidden et al. eds., 2021). Clinical judgment is an essential part of that framework, and it involves a detailed series of standards and guidelines that experts must follow when evaluating an individual. Robert L. Schalock & Ruth Luckasson, *Clinical Judgment* (2d ed. 2014) [hereinafter *Clinical Judgment*].

Using that clinical framework, both intellectual functioning and adaptive behavior must be measured using objective, standardized instruments, *AAIDD-12* at 35; however, qualitative information concerning these two prongs must also be identified and evaluated. *See id.* at 26 tbl.3.1, 30 tbl.3.3 (listing examples of limitations in intellectual functioning and in adaptive behavior). As a result, a diagnosis of intellectual disability necessarily involves a comprehensive assessment of evidence that addresses the criteria for all three prongs. *See Clinical Judgment* at 14.

Intelligence tests quantify intellectual capacities and illuminate limitations in intellectual functioning. *See AAIDD-12* at 28-29. Qualitative data regarding intellectual functioning can be gleaned from the person's history of demonstrated abilities by considering hallmarks of intelligence in daily life, such as reasoning, planning, and learning from experience. *Id.* at 26 tbl.3.1. Likewise, as adaptive behavior scales quantify limitations in real-world functional capacities, qualitative information can be found in a person's abilities in such areas as

problem-solving, understanding money, and making social judgments. *Id.* at 30 tbl.3.3.

As an objective assessment, intelligence quotient (IQ) testing can provide a good estimate of intellectual functioning. *See AAIDD-11* at 31. However, IQ tests (and the scores obtained from them) have limitations. As this Court has recognized, IQ tests have inherent imprecision. *Hall v. Florida*, 572 U.S. 701, 713 (2014). *See also AAIDD-11* at 36. In addition, clinicians may produce erroneous scores if they do not select, administer, and score tests correctly. *See* American Educational Research Association, American Psychological Association & National Council on Measurement in Education, *Standards for Educational and Psychological Testing*, 114-18, 142-44 (2014) (discussing Standards 6.1 through 6.9: Standards for Test Administration, Scoring, Reporting and Interpretation, and Standards 9.0 through 9.10: Standards for Test Users' Rights and Responsibilities).

Standardized tests alone cannot substitute for a complete analysis of intellectual functioning. Clinicians must always review qualitative information about the individual's functioning as well. *AAIDD-12* at 41 (clinicians should conduct a social, medical, and educational history). When IQ test results are not near the boundary of intellectual disability, and the qualitative information about the person's functioning agrees with the quantitative IQ testing results, a diagnostic determination may be relatively straightforward. However, a complete review of all quantitative and qualitative data is imperative when IQ scores are ambiguous or when their accuracy is in doubt.

The existence of multiple IQ scores for an individual does not change the requirement for a complete analysis. Clinicians can make use of the information from multiple test scores in a variety of ways (after conducting an analysis of each test to consider the accuracy of each score). *AAIDD-11* at 35-41. Additional scores do provide more information, but their presence cannot completely cure imperfections in IQ score accuracy, nor do they provide qualitative information about the person's intellectual functioning. Neither clinicians nor courts should treat the existence of multiple scores as an excuse to avoid the need for a clinically valid, complete analysis of those scores and of all the other information about the person's intellectual functioning and adaptive behavior.

ARGUMENT

I. INTRODUCTION

In *Atkins v. Virginia*, this Court recognized the public and legislative consensus that executing people with intellectual disability serves none of the purposes of the death penalty. 536 U.S. 304, 316 (2002). The *Atkins* Court referred to the AAIDD (then AAMR[2]) and the American Psychiatric Association definitions of intellectual disability. These definitions

[2] The previous name of the organization was the American Association on Mental Retardation (AAMR). The organization changed its name to the American Association on Intellectual and Developmental Disabilities in 2007. *See* Robert L. Schalock et al., *The Renaming of Mental Retardation: Understanding the Change to the Term Intellectual Disability*, 45 *Intell. & Developmental Disabilities* 116, 116 (2007).

were (and remain) essentially identical to one another. *Id.* at 308 n.3.

Although some of the terminology has been revised over time, the core definition of intellectual disability used in *Atkins* remains unchanged. *See AAIDD-12* at 20. For well over half a century, intellectual disability has been understood as a lifelong condition that manifests before adulthood and is characterized by substantial difficulties in thinking and daily functioning in comparison to the general population. *Id.* at 16 tbl.2.1. The clinical definition of intellectual disability consists of significant limitations in intellectual functioning (Prong 1), accompanied by significant limitations in adaptive behavior (Prong 2), and the requirement that these limitations first occur during the developmental period (Prong 3). *Id.* at 1. These criteria had been essentially unchanged for over 40 years before *Atkins* was decided, and they have remained so since that time. *Id.*

This definition is the foundation of the clinical diagnostic framework for intellectual disability. *See id.* at 23. The standards, procedures, and clinical judgment that comprise the diagnostic process are central to that framework. *Clinical Judgment* at 7-8 (discussing the definition, purpose, and importance of clinical judgment and its related processes and standards). This framework is the product of hundreds of thousands of hours of scientific research, review, clinical observation, and practical experience. *See, e.g.*, R.C. Scheerenberger, *A History of Mental Retardation: A Quarter Century of Promise* (1987) (discussing the many scientific and clinical advances regarding our understanding of intellectual disability made between 1960 and 1985). Scholars, clinicians,

and physicians have developed and refined this understanding of intellectual disability for the last 150 years. *Id.*; R.C. Scheerenberger, *A History of Mental Retardation* 109-254 (1983) (discussing the history of scientific and clinical advances regarding intellectual disability that occurred during the late-nineteenth through the mid-twentieth centuries).

Courts and other decision-makers have long relied on this diagnostic framework as the established method for obtaining an accurate diagnosis in order to address a number of different clinical and legal issues. *See AAIDD-12* at 40-41 (discussing contexts in which a diagnosis might be necessary). Because people who have intellectual disability generally need ongoing supports throughout their lives, a clinical diagnosis is of particular importance in accessing early intervention, special education, home and community-based services, Supplemental Security Income, and other such supports. *Id.* at 23 (noting that one of the primary purposes of a diagnosis of intellectual disability is to "establish eligibility for relevant benefits, supports and services"); *AAIDD-11* at 11 (discussing the role of a diagnosis of intellectual disability in eligibility for services).

The clinical framework for diagnosing intellectual disability was largely developed long before, and independently of, the criminal courts' interest in diagnosing intellectual disability. *See* R.C. Scheerenberger, *A History of Mental Retardation: A Quarter Century of Promise* (1987). Since *Atkins*, this Court has determined that the clinical diagnostic framework should also inform intellectual disability

determinations in *Atkins* claims.[3] *See Hall v. Florida*, 572 U.S. 701, 721 (2014) ("The legal determination of intellectual disability . . . is informed by the medical community's diagnostic framework."); *Moore v. Texas*, 581 U.S. 1, 13 (2017) (*Moore I*) (restating the requirement that an *Atkins* evaluation be informed by the medical community's framework (*citing Hall*, 572 U.S. at 721)); and *Moore v. Texas*, 586 U.S. 133, 139 (2019) (*Moore II*) (reiterating that any evaluation of an *Atkins* claim using non-clinical factors is not acceptable). Using this framework and its clinical definitions, criteria, and standards to diagnose intellectual disability is the only reliable method of ensuring the necessary level of accuracy and validity in both evaluative procedures and diagnostic results. *See Hall*, 572 U.S. at 704 (using a non-clinical standard for evaluating an *Atkins* claim "creates an unacceptable risk" of executing a person with intellectual disability).

Regardless of the circumstances, a diagnostic evaluation for intellectual disability requires a comprehensive, broad-based assessment of intellectual functioning and adaptive behavior using standardized instruments: intelligence tests and adaptive behavior scales. *Clinical Judgment* at 27-29. However, the clinical diagnostic framework also requires examining qualitative evidence. The areas for examination typically include educational and occupational history, academic ability (reading, writing, math), relevant family history, significant health history, mental health history, developmental history (including events that may have a bearing on

---

[3] However, *Atkins* cases comprise only a tiny fraction of the universe of intellectual disability assessments.

educational attainment or psychological adjustment), social interactions and peer relationships, teacher observations, family observations, and, of course, any prior test results. *See AAIDD-12* at 41 (explaining that clinicians must gather and review a thorough social, medical, and educational history). In many cases, the information correlates well and the diagnostic decision is clear. However, there are times when a more intensive review is required.

II. THE DEFINITION AND DIAGNOSIS OF INTELLECTUAL DISABILITY

A. The Definition of Intellectual Disability

Intellectual disability is defined as "significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills.[4] This disability originates during the developmental period."[5]

---

[4] States that enacted legislative or judicial definitions of intellectual disability in response to *Atkins* often adopted the terminology used at the time. For example, Alabama used and still refers to the term "significantly subaverage intellectual functioning" to describe the limitations measured in Prong 1 of the definition. *Ex parte Perkins*, 851 So.2d 453, 456 (Ala. 2002). This is the term from AAIDD's 1992 manual. American Association on Mental Retardation, *Mental Retardation: Definition, Classification, and Systems of Supports* 1 (9th ed. 1992) [hereinafter *AAMR-9*]. While some of the terms have been updated during the last 30 years, the meaning of those terms has not changed. 'Significantly subaverage' intellectual functioning has become 'significant limitations' in intellectual functioning, *see AAIDD-12* at 13, but Prong 1 has always required an IQ score "of approximately 70 to 75 or below." *AAMR-9* at 5.

[5] In its most recent diagnostic manual, AAIDD described the end of the developmental period as occurring at age 22 where it was previously defined as ending at age 18. *AAIDD-12* at 32-33. This change incorporates new scientific information and multiple

AAIDD-12 at 1. The American Psychiatric Association's definition is essentially identical. Intellectual disability has an "onset during the developmental period that includes both intellectual and adaptive functioning deficits in conceptual, social, and practical domains." *DSM-5-TR* at 37. (There have been no substantive changes to the definition of intellectual disability in over 60 years. *See AAIDD-12* at 16 tbl.2.1, 18-19 tbl.2.2.)

B. Using Clinical Judgment in Diagnosing Intellectual Disability

'Clinical judgment' refers to the processes and strategies necessary to reach a high-quality, valid, and accurate determination. *Clinical Judgment* at 6-7. The principles of clinical judgment require a clinician to use specialized training and experience to gather and review extensive data, employ clinical strategies and guidelines in reviewing and assessing information, and use professional analytical skills to make a determination.[6] *Id.* at 26-39 (discussing these

---

perspectives on development and the developmental period. *Id.* at 32. This modification is not expected to (and was not intended to) result in any appreciable expansion of the number of people diagnosed with intellectual disability. *Id.* at 33. The prevalence of intellectual disability in the general population has remained essentially constant and is estimated to be around 1 to 1.5% of the U.S. population. *See* Pallab K. Maulik et al., *Prevalence of Intellectual Disability: A Meta-analysis of Population-based Studies*, 32 *Research in Developmental Disabilities* 419, 431 (2011) (estimating the prevalence of intellectual disability at 1.037%).

[6] The term 'clinical judgment' is not an invitation for clinicians to make casual or impulsive diagnoses. *Clinical Judgment* at 15. In fact, the principles of clinical judgment specifically prohibit undocumented determinations, decisions based on poorly-administered or inadequately-documented assessments,

strategies and guidelines in detail). These requirements are designed to produce an organized, logical evaluation based on clinical knowledge and culminating in a well-reasoned, clear, and fully documented analysis. *Id.* at 7; *see* American Association on Intellectual and Developmental Disabilities, *User's Guide: Intellectual Disability: Definition, Classification, and Systems of Supports* 4-5 (2012) [hereinafter *AAIDD User's Guide*].

C. The Diagnostic Process

In planning and conducting a diagnostic evaluation, clinicians begin with the understanding that the adaptive (day-to-day) behavior limitations of Prong 2 are just as important to the diagnosis of intellectual disability as are the limitations in intellectual functioning required by Prong 1. *AAIDD-12* at 33-34. A clinical evaluation of intellectual functioning (Prong 1) requires an objective, standardized assessment in the form of an intelligence test. *Id.* at 27-28. Similarly, the assessment of adaptive behavior (Prong 2) requires objective, standardized testing with an adaptive behavior scale, *id.* at 35, such as the 'ABAS-3,' Patti L. Harrison & Thomas Oakland, *ABAS-3: Adaptive Behavior Assessment System, Third Edition* (3d ed. 2015), or the 'DABS,' Marc Tassé et al., *Diagnostic Adaptive Behavior Scale* (2017).[7]

---

decisions ignoring measurement error, and any reliance on stereotypes or personal bias. *Id.*

[7] Instruments such as these have been designed to provide an objective measure of a person's ability to perform conceptual, social, and practical functions in home or community

Empirical data is a necessary component of an evaluation, but intellectual disability cannot be assessed by standardized measurements alone. In addition to these objective measures, an accurate diagnosis of intellectual disability requires a clinician to gather, review, and analyze multiple sources of information relevant to each prong of the definition. *AAIDD-12* at 41-42; *Clinical Judgment* at 14 ("For diagnosis, data will be obtained from intelligence tests, adaptive behavior scales, developmental measures, and social history, medical, and educational records."); *see also DSM-5-TR* at 38 (explaining that the diagnosis of intellectual disability is based on both clinical assessment and standardized testing). When a clinician is evaluating an adult for intellectual disability, much of the needed information may concern events that occurred and were documented when the person was a child or adolescent. *See AAIDD-12* at 41-42 (explaining that such a "retrospective diagnosis" requires a "thorough social, medical, and educational history" that establishes "multiple valid data points").

Information collected for this detailed history will often be relevant to at least two, and sometimes all three, prongs of the diagnosis of intellectual disability. Direct observations may reflect limitations in both intellectual functioning and adaptive behavior. Additionally, the observation of such limitations in childhood addresses the third prong of the diagnostic criteria. For example, school records, which often describe limitations in adaptive behavior, may also include information describing the

---

environments, and in comparison to same-aged peers. *See AAIDD-12* at 29-30 tbl.3.3, 31 tbl.3.4.

individual's limitations (in comparison to peers) in reasoning, thinking abstractly, and understanding complex ideas, which are elements of intellectual functioning. *See id.* at 26. School records from the developmental period may address both intellectual functioning and adaptive behavior, for example, performance on standardized tests; reading level; grasp of numbers, time, and money; social judgment and gullibility; and capacity to follow rules. *See Clinical Judgment* at 28.

Once all the information has been gathered, clinicians must employ the principles of clinical judgment to objectively assess the quality and reliability of each piece of information, both quantitative and qualitative, in order to determine whether, and how well, it supports[8] (or diverges from) the rest of the evidence. *See id.* at 32 tbl.4.2 (listing requirements for integrating and assessing the information gathered). Among other tasks, they must verify reports of limitations in adaptive behavior and intellectual functioning and be wary of stereotypes that may have played a role in evaluations, records,

---

[8] Courts will sometimes see the term 'convergence' in clinical reports or scholarly works discussing the diagnostic process. This term describes the extent of the consistency of the information provided by different sources. As clinicians review the information related to an individual's intellectual functioning, they will note how well information from intelligence tests converges with direct reports of the individual's behavior, and whether all the data points to the same conclusion. Marc J. Tassé, *Adaptive Behavior Assessment and the Diagnosis of Mental Retardation in Capital Cases*, 16 *Applied Neuropsychology* 114, 118 (2009); *see AAIDD-12* at 40 (explaining the requirement to synthesize and corroborate information from multiple sources including a thorough social, educational, and medical history).

and observations. *Id.* at 33-39 (discussing the steps of a thorough collection and review of relevant materials). The clinician's objective is to evaluate all the information and avoid both false positives and false negatives when making a final determination as to whether the individual has limitations that meet or exceed the requirements for each prong of the definition. *AAIDD-12* at 39-40.

III. EVALUATING LIMITATIONS IN INTELLECTUAL FUNCTIONING (PRONG 1)

A. The Definition of Intellectual Functioning

Intellectual functioning is a broad concept that also includes the common characteristics of intelligence, such as reasoning, planning, problem-solving, thinking abstractly, comprehending complex ideas, learning quickly, and learning from experience. *AAIDD-12* at 25; *DSM-5-TR* at 38. Furthermore, intellectual functioning affects and is affected by other dimensions of human functioning in relation to one's environment. *AAIDD-11* at 31.

IQ scores can be, and often are, a good representation of intellectual functioning, but a full assessment also considers information that is not captured by IQ testing. *See id.* (noting that limitations in intelligence should be considered in light of other dimensions of human functioning, and while IQ scores are the best single representation of intellectual functioning, they are "far from perfect"). In cases where an individual's IQ scores are close to the boundary of an intellectual disability diagnosis, or where a clinician suspects that scores may be inaccurate, clinical judgment requires further inquiry. *See id.* at 41 ("[T]he assessment of

intellectual functioning must be based on sound procedures and may, at times, require information from multiple sources."); *DSM-5-TR* at 42.

B. An Overview of Intelligence Tests

IQ tests are designed to gauge intelligence by measuring individual capacities that make up general intelligence using numerous, smaller subtests. A full-scale IQ score is then derived from the combined, scaled subtest scores. *See, e.g.*, Elizabeth O. Lichtenberger & Alan S. Kaufman, *Essentials of WAIS-IV Assessment* 99-118 (2d ed. 2013) (discussing the entire scoring procedure for this test). Intelligence tests are also comparative measures, which means their scores indicate how well the individual's performance compares to a representative group of peers in the general population. *Id.* at 99. Moreover, IQ tests are not direct measurements like height, weight, or body temperature. They are indirect measures that quantify different aspects of intelligence through their subtests, each of which are designed to gauge specific cognitive abilities related to overall intelligence. *Id.* at 25-26, 31-36. In order to interpret IQ scores correctly, clinicians must fully understand the comparisons being made between the individual and the general population, as well as the purpose and nature of each subtest. American Educational Research Association, American Psychological Association & National Council on Measurement in Education, *Standards for Educational and Psychological Testing* 142-43 (2014) (discussing Standard 9.2 which states that the test user should know and understand the testing manual and the studies validating the test).

Clinicians must also be conscious of the inherent limitations of IQ tests and scores. Despite their careful construction, intelligence tests are not simple measures and are not perfectly precise. *AAIDD-11* at 36. Experts in intellectual disability have long understood this concern and, as a result, have recognized that a single, 'bright-line' cutoff score for intellectual disability "is not psychometrically justifiable." *Id.* at 40. Therefore, the boundary of intellectual disability is defined as being *approximately* two standard deviations below the population mean[9] of the test. *AAIDD-12* at 130 (emphasis added); *DSM-5-TR* at 38. On many of the frequently used tests, this is a range of approximately 65 to 75 points. *DSM-5-TR* at 38; *see AAIDD-12* at 36.

The degree of unavoidable imprecision in intelligence tests can be mathematically estimated. *AAIDD-12* at 130. The result of that calculation is the 'standard error of measurement,' or SEM. *Id.* at 130-131. The SEM is used to produce a 'confidence interval' ranging above and below the reported score. *Id.* An IQ test score should always be reported with the SEM. Taken together, they indicate, with a 95 percent degree of confidence, where the individual's

---

[9] The 'mean' score of an intelligence test is the score that the largest group of people in the population would achieve if everyone were to take the test. In other words, 100 is the 'most common' score. On most intelligence tests, the mean is a score of 100 points.

The 'standard deviation' of an intelligence test is a measure of how dispersed the actual scores of the general population are in relation to the mean. If scores are widely spread, the standard deviation is higher. If scores are mostly closer to the mean, the standard deviation is lower. *See AAIDD-12* at 130.

so-called 'true'[10] score is likely to fall. In practical terms, this often means the range is about 5 points above and below the reported score.[11] *DSM-5-TR* at 38; Floyd et al., *Theories and Measurement of Intelligence* at 411; *see AAIDD-12* at 35-36, 119.

C. Assessing Scores from Earlier Intelligence Testing

Many adults who are evaluated for an intellectual disability diagnosis have been given one or more intelligence tests at earlier points in their lives.[12] While scores from prior testing may be helpful

---

[10] A 'true' IQ score is a term of art and a purely hypothetical construct. *See AAIDD-12* at 133.

[11] Test manuals and computerized test-scoring programs sometimes provide clinicians with confidence intervals that are fewer than 5 points. However, experts have warned against over-reliance on these smaller intervals when diagnosing intellectual disability for two reasons. First, these smaller intervals are often based on a tiny subset of the overall norming data and, therefore, may not have the precision required for an intellectual disability diagnosis. Randy G. Floyd et al., *Theories and Measurement of Intelligence*, in 1 *APA Handbook of Intellectual and Developmental Disabilities* 385, 411 (Laraine Masters Glidden et al. eds., 2021) [hereinafter Floyd et al., *Theories and Measurement of Intelligence*]. Second, the amount of error in IQ test results increases when IQ scores are very high or very low. *Id.*; *see also AAIDD-11* at 39 (discussing the potential for additional error with 'extreme scores' and the fact that relatively few people who have scores below 70 are in the norming data of many tests). The smaller confidence intervals that may be provided by the test's manual are calculated using only the age of the test subject and do not reflect the additional error that may occur when scores are very high or low. Floyd et al., *Theories and Measurement of Intelligence* at 411; *AAIDD-11* at 39.

[12] When a child is referred for testing it often indicates that parents, teachers, or other adults have noticed that the child appears to have deficits in intellectual functioning or adaptive

in evaluating intellectual functioning, clinicians must assess them with great caution. Clinical judgment is absolutely essential in investigating the quality of the test, its administration, and its scoring, as these factors each have a potent influence on the accuracy of a reported score.

1) Evaluating Intelligence Testing Instruments

To be acceptable for use in diagnosing (or ruling out) intellectual disability, a test must be specifically designed to be a comprehensive measure of intelligence. *See* American Educational Research Association, American Psychological Association & National Council on Measurement in Education, *Standards for Educational and Psychological Testing* 142-43 (2014); *AAIDD-12* at 29. The test must also be standardized, comprehensive, and individually administered. *AAIDD-12* at 38; *DSM-5-TR* at 38. Additionally, it should have strong statistical validity, reliability, sensitivity, and specificity[13] ratings.

---

behavior as compared to peers. *See Clinical Judgment* at 28 (discussing information that may be found in educational records, including testing). (In addition to any Prong 1 or Prong 2 relevance, this sort of evidence is also relevant to the Prong 3 requirement of manifestation during the developmental period.)

[13] 'Validity,' 'reliability,' 'sensitivity,' and 'specificity' are all psychometric terms of art. Validity is a complicated topic, but a test's validity concerns how well the test measures what it is intended to measure. Anne Anastasi & Susana Urbina, *Psychological Testing* 8 (7th ed. 1997). Reliability measures how often the same test would produce the same or similar scores if it could (hypothetically) be given to the same individual repeatedly, each for the first time. *Id.* Sensitivity refers to how well the test will detect true positives and will avoid false

Clinicians should also review the norming[14] procedures for that specific edition of the test. The test should have been normed using a large and representative sample of people, including people with intellectual disability. *See* Richard W. Woodcock, *Norms*, in 2 *Encyclopedia of Intelligence* 770, 772 (Robert J. Sternberg et al. eds., 1994). The test norms must have been current at the time the test was given because researchers have determined that intelligence test norms inevitably become less reliable over time. *See AAIDD-12* at 42.

Finally, clinicians should consider the appropriateness of the test used with the particular individual. *AAIDD-11* at 36. For example, an individual with a severely limited verbal ability should obviously not be given a test that relies on verbal responses. *Id.* Excluding consideration of a score from such an inappropriate testing instrument is also an example of clinical judgment, as it is likely to be an inaccurate indication of the individual's intellectual functioning.

---

negatives. *AAIDD-12* at 40. Specificity refers to how well the test will detect true negatives and will avoid false positives. *Id.*

[14] 'Norming' is the process of determining what the average person's performance would be on a newly designed or revised test. *See* Richard W. Woodcock, *Norms*, in 2 *Encyclopedia of Intelligence* 770, 770-71 (Robert J. Sternberg, et al. eds., 1994). A test is 'normed' by administering it to a representative sample of individuals of different ages, backgrounds, and intellectual abilities. *Id.* at 770. These norms are the general population's scores against which an individual's score could be compared. *Id.* at 771.

2) Intelligence Test Administration

A clinician reviewing an individual's prior IQ scores should gather all the available information about the administration of each prior test, such as the test administrator's qualifications and whether the test was administered according to the instructions in the test's manual. Anne Anastasi & Susana Urbina, *Psychological Testing* 10 (7th ed. 1997). The clinician should also take note of any concerns the administrator expressed about testing conditions, *id.* at 14-15, or about the individual's effort on the test. *See AAIDD-12* at 40-41 (discussing strategies to resolve any concerns about effort).

Another concern can arise when the same or a similar test is given to the same individual more than once in relatively close succession. The administration of two or more identical (or similar) tests close in time may cause an artificially inflated score on the second and any subsequent test. Scores inflated due to this 'practice effect' do not reflect an actual change in intellectual functioning but are simply the result of exposure to the same test material. Alan S. Kaufman & Elizabeth O. Lichtenberger, *Assessing Adolescent and Adult Intelligence* 163-65 (3d ed. 2006). This effect can occur when a similar test is given months, or even a few years, after the first administration of the test. *Id.* at 164.

3) Scoring Errors and Bias in Scoring

Scoring sheets, notes, and any other available materials from the administration of previous intelligence tests should be examined closely because they may provide useful information about the mindset and expectations of the individual who administered and scored the test. Even the most

careful test administrator may make one or more errors in scoring intelligence tests.[15] John R. Slate et al., *Practitioners' Administration and Scoring of the WISC-R: Evidence That We Do Err*, 30 *J. School Psychology* 77, 81 (1992) (noting that practitioner errors deviated as much as 4 points from the correct score). These errors can range from simple arithmetical mistakes to more pervasive issues. Paul A. McDermott et al., *Whose IQ Is It?—Assessor Bias Variance in High-Stakes Psychological Assessment*, 26 *Psychological Assessment* 207, 208 (2014) (explaining that drifting from the standardized protocol for test administration and erroneously scoring test responses are common sources of error).

Reviewing information from the test's administration may also reveal any conscious or unconscious bias in the administration of the test. Anne Anastasi & Susana Urbina, *Psychological Testing* 18 (7th ed. 1997) (noting that examiners may affect the test taker's responses through the examiner's own expectations); John R. Slate et al., *Practitioners' Administration and Scoring of the WISC-R: Evidence That We Do Err*, 30 *J. School Psychology* 77, 81 (1992) (explaining that 'generosity' in assigning points may reflect a sincere desire to help a child but erroneously elevates scores). For example, notes from testing could reveal that a school psychologist had already reached a conclusion about the child's abilities before testing the child.

---

[15] Scoring errors may occur even when computerized scoring programs are used. *AAMR-9* at 48 (explaining that clinicians should determine whether scoring software controls for missing data and data entry errors, either of which could create calculation errors).

4) Synthesis of Individual Scores with Other Information

After gathering all the available information for each test and reviewing it for test quality, administration, and scoring errors, the clinician must use the established principles of clinical judgment to determine whether the score for each test should be included in the evaluation of intellectual functioning, and, if so, how much weight should be placed on it. *See Clinical Judgment* at 9 tbl.2.1 (explaining that synthesis includes consideration of the relative weight and possible combination of information as a basis for decisions); *AAIDD-12* at 40 (synthesis involves corroborating information from multiple sources); see also *DSM-5-TR* at 42. Some test scores may be so suspect that the score should be eliminated altogether from further consideration in the diagnostic process.

The final step is synthesis of the information. *Clinical Judgment* at 31. This holistic analysis includes integrating all the information from different sources, reviewing the likelihood of false positives and negatives, exploring any possible reasons for inconsistency (divergence) in the information, and recognizing any factors that could have affected the test results. *Id.* at 32 tbl.4.2. In many cases, scores and qualitative information converge, and the process of synthesis is not a complicated endeavor. However, when the information does not clearly converge, additional examination is required.

D. Consideration of Multiple Intelligence Scores

Multiple IQ test scores for the same individual can provide additional information for an evaluation of intellectual disability. When these scores converge

and are accurate, they can be very informative. However, the existence of multiple scores does not lessen or eliminate the need for clinical expertise. More scores often elevate the level of analysis and clinical judgment required, in part because they may actually increase the risk of an incorrect diagnosis if they are not carefully reviewed for indicators of inaccuracy. *See AAIDD User's Guide* at 19 tbl.3.3 (explaining that when scores differ, clinicians must thoroughly explore the reasons for differences in this data). Also, the availability of additional scores does not completely resolve the problem of the inherent imprecision in IQ scores. Most importantly, the availability of additional IQ scores will not provide a perfect measure of intellectual functioning that eliminates the need for other types of information.

Any number and variety of IQ tests might be found in a given individual's records. Other than requiring the careful use of clinical judgment and professional responsibility in performing a complete analysis, there can be no single, mandatory empirical method for clinicians to use in considering multiple scores. There are too many variables to consider in determining the validity of each test score and the convergence of that score with other scores and qualitative information to rationally support the use of only one analytic procedure. Clinicians do often examine how well multiple scores converge by determining whether and by how much the confidence interval of each test overlaps with the intervals of the other tests. This information sometimes may indicate the relative accuracy of the individual and collective

scores.[16] Dale G. Watson, *Intelligence Testing*, in *The Death Penalty and Intellectual Disability* 113, 124 (Edward A. Polloway ed., 2015) (scores should be examined to determine whether the confidence intervals overlap).

It is possible to use a statistical calculation to create a single composite score from multiple scores.[17] *See, e.g.*, Floyd et al., *Theories and Measurement of Intelligence* at 415-17. This is a "complicated endeavor," *Hall*, 572 U.S. at 713, but in some circumstances clinicians might find it helpful to consider scores in this way. Clinicians and courts should note, however, that such a calculation presupposes that all the scores used to create it are accurate. Unintentionally including inaccurate scores in a calculation creates additional error rather than improving clarity.

Courts should be wary of purported methods of combining multiple IQ scores that might seem obviously correct, but which will not produce a valid

---

[16] However, as this Court has recognized, convergence of scores alone is not a guarantee of accuracy. "[B]ecause the test itself may be flawed, or administered in a consistently flawed manner, multiple examinations may result in repeated similar scores, so that even a consistent score is not conclusive evidence of intellectual functioning." *Hall*, 572 U.S. at 714.

[17] This method requires information that clinicians may not have, that might not be easily accessible, or that simply may not exist. Dale G. Watson, *Intelligence Testing*, in *The Death Penalty and Intellectual Disability* 113, 124 (Edward A. Polloway ed., 2015). Estimates can be calculated to approximate some of that missing information, but using those estimates to create a composite score could create yet another layer of uncertainty and potential error. *See id.*

result. For example, IQ scores cannot be averaged[18] in the same way that direct measurements of height, weight, or body temperature can be averaged. *See* Floyd et al., *Theories and Measurement of Intelligence* at 414-15.

Similarly, when an individual's IQ scores are clustered relatively close together, it might seem reasonable to assume that the 'true' score must be somewhere *within* that grouping. However, statistically, it is more likely than not that the 'true' score is *beyond* the grouping in a direction *away from* the mean. When an individual's scores are all exceptionally low (or all exceptionally high)—that is, if they are grouped far away from the mean score of 100 used on most IQ tests—it is a statistical indicator that the individual's 'true' score is likely to be *even farther away from the mean*. In other words, when IQ scores are all close to the boundary of intellectual disability, the 'true' score may actually be even lower than the reported scores.[19]

---

[18] Due to the construction of the tests that produce them, averaging multiple IQ scores is statistically (and clinically) invalid. The result of such an averaging calculation would no longer be based on the same scale as the scores that were used to create it and would not have the same meaning as the measurements from either scale. *See* Floyd et al., *Theories and Measurement of Intelligence* at 414-15. In the same way, this so-called 'average' test score would not be a more accurate representation of the likely 'true' score and, in fact, would be misleadingly high when scores are low and misleadingly low when scores are high. *Id.* at 415.

[19] This concept may seem counterintuitive, but it is correct. Scores tend to 'regress to the mean': any given score will tend to fall closer toward the point at which most scores will fall, that is, closer to the mean. *See* W. Joel Schneider, *Why Are WJ IV Cluster Scores More Extreme Than the Average of Their Parts?*, *Woodcock*

IV. CLINICIANS MUST CONSIDER ALL THE RELEVANT INFORMATION ABOUT THE INDIVIDUAL

An IQ score cannot be the "final and conclusive evidence of a defendant's intellectual capacity, . . . experts in the field would consider other evidence." *Hall* at 712. A clinically sound evaluation of intellectual disability will always entail a complete review of all the relevant information of the person's functioning, not just test scores. *AAIDD-12* at 38 (obtaining a thorough history, conducting a broad-based standardized assessment, and synthesizing the obtained information are strategies used in every diagnosis); *Clinical Judgment* at 9 tbl.2.1 (synthesis provides data for generating and testing hypotheses, considering the relative weight and possible combination of information as a basis for decisions). Qualitative information provides a breadth and depth of understanding to the analysis beyond IQ scores, and it can serve as a check against imprecise or suspect test results. "Intellectual disability is a condition, not a number." *Hall*, 572 U.S. at 723.

When test scores converge with each other, as well as with the quantitative reports of the

---

Johnson IV Assessment Service Bulletin Number 7 (Riverside Assessments, LLC, Itasca, Ill.) (2016) at 8 https://info.riversideinsights.com/hubfs/ASBs/WJIV_ASB_7_FINAL.pdf.

In other words, if a low score is simply due to chance, one would expect that future scores from the same individual would fall closer to the mean. However, if several scores from the same individual cluster near a point farther away from the mean, it is likely that the individual's hypothetical true score is even lower than the reported scores, because the reported scores are going to regress towards the mean. *Id.*

individual's functioning, this can be a clear and uncomplicated process. However, when IQ scores are close to the boundary of intellectual disability, or when a clinician has concerns that one or more IQ tests may have produced unreliable scores, the analysis becomes more complex. *See Clinical Judgment* at 32 tbl.4.2, 39 (clinicians should explore reasons for inconsistency in the obtained information); *AAIDD User's Guide* at 19 tbl.3.3.

In such situations, the clinician must conduct a detailed review of qualitative reports of the individual's skills and behavior "focusing heavily on the functional assessment of what the person actually does" and on how the person interacts with the environment. *Clinical Judgment* at 29. The purpose of this inquiry is to create the most accurate picture of the person's day-to-day intellectual functioning by reviewing observations from different people in different contexts regarding the individual's actual capacity to reason, plan, solve problems, and other similar abilities. *Id.*; *see also AAIDD-12* at 26 tbl.3.1 (listing examples of significant limitations in intellectual functioning). For instance, reports of an individual's inability to comprehend simple instructions or to learn the rules of a childhood game are potential examples of real-life limitations in intellectual functioning.

An evaluation of this type requires a professional's clinical expertise and judgment, but correct diagnoses using these methods have been made for decades in a variety of contexts. *AAIDD-11* at 41 (assessment of intellectual functioning may require information from multiple sources); *AAIDD User's Guide* at 19 tbl.3.3 (diagnosis requires integrating data from both standardized assessments

and a thorough history); *AAMR-9* at 15 (consideration of adaptive skills may confirm functional limitations when the validity of IQ scores is in question or scores are equivocal); American Association on Mental Retardation, *Mental Retardation: Definition, Classification, and Systems of Supports* 57 (10th ed. 2002) (the existence of the adaptive deficits requirement may operate in part to prevent a false diagnosis); *DSM-5-TR* at 42 (substantial adaptive behavior problems in social functioning or in other areas may warrant a diagnosis of intellectual disability when IQ scores are somewhat above 65-75); American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorder*s 37 (5th ed. 2013) (a diagnosis of intellectual disability may be warranted when an IQ score is above 70 when social or other adaptive behavior problems are severe).

Using this clinical framework for diagnosing intellectual disability is the only reliable method of ensuring the necessary level of accuracy and validity in both evaluative procedures and diagnostic results.

CONCLUSION

For the foregoing reasons, *amici* urge the affirmance of the judgment below.

Respectfully submitted,

ANN M. DELPHA
Counsel of Record

JAMES W. ELLIS
Counsel for Amici Curiae

September 24, 2025