Testing for Intellectual Disability: The When, the How, and the What

Dale G. Watson, Ph.D.

August 29, 2025

2025 Making the Case for Life Seminar

Las Vegas, NV

In 2002 the United States Supreme Court, in the case of Daryl Renard Atkins (*Atkins v. Virginia, 536 US 304 (2002)*), determined that the execution of those with mental retardation (now intellectual disability) was a violation of the Eighth Amendment's prohibition of cruel and unusual punishments. However, the Court left it to the states to determine how the condition should be determined. For example, Florida established a "bright line" cutoff such that intellectual disability could only be established if an individual's Intelligence Quotient (IQ) were below 70, i.e., 2 standard deviations below the population mean of 100. However, in *Hall v. Florida, 572 US 701 (2014)*, the Court ruled, recognizing that IQ tests are imperfect, that using such a fixed cutoff for determining intellectual disability was an unconstitutional implementation of *Atkins*. Such a "rigid rule" created an unacceptable risk that individuals with intellectual disability would be executed. Rather, the Court, relying upon professional standards, recognized that any given IQ score obtained represented a range of potential scores rather than a fixed and immutable score. Thus, in *Atkins* and *Hall* the Court made clear that understanding the essentials of the diagnosis of intellectual disability was foundational in capital litigation.

To understand intellectual disability, we must first understand what intelligence is. Though there is no single definition of intelligence, most lay people and intelligence researchers recognize components of what we understand it to be. In general, intelligence is thought to be some combination of acquired information, facility with language, problem solving capacity, processing speed, and reasoning abilities. A common aphorism is that

"Intelligence is like gravity – you can't touch it, but you know it's there." Intelligence is in fact a latent variable. Intelligence is not a score on an IQ test. Rather, it is an emergent feature of infinite combinations of brain processes that evolved because of their survival value.

David Wechsler, the developer of the Wechsler IQ tests, told us that "What we measure with tests is not what tests measure – not information, not spatial perception, not reasoning ability. These are only a means to an end. What intelligence tests measure is something much more important: the capacity of an individual to understand the world about him and his resourcefulness to cope with its challenges."[1] Intelligence is a resource, and those with lower levels of intellectual capacity are much more vulnerable to the vicissitudes of life.

The two most prominent current theories of intelligence focus, on the one hand, on the psychometric properties or structure of intelligence and on the other, on the neurological basis for intelligence. From the psychometric perspective, the Cattell-Horn-Carroll theory (Schneder & McGrew, 2018), which is derived from factor analysis, posits that intelligence is made of many specific skills subsumed under a set of broad cognitive abilities that are in turn subsumed under a general intelligence factor or "*g"* at the apex. Those broad cognitive abilities most commonly measured by IQ tests include Fluid Reasoning (*Gf),* Comprehension-Knowledge (*Gc),* Short-term Working Memory Capacity (*Gwm),* Visual-Spatial Processing (*Gv),* Learning Efficiency (*Gl)*, Retrieval Fluency (*Gr)*,

[1] Wechsler, 1975 cited by Wechsler, D., Raiford, S. E., & Presnell, K. (2024). *Wechsler Adult Intelligence Scale: Technical and interpretive manual* (5th ed.). NCS Pearson, p. 4).

Auditory Processing (*Ga),* and Processing Speed (*Gs).*[2] Fluid Reasoning (*Gf*) and Comprehension-Knowledge (*Gc*) are those broad ability domains most closely related to general intelligence (*g).*[3] The former represents problem solving in novel circumstances that require "on the spot" responses. In contrast, the latter measures the capacity to develop and utilize knowledge gained through experience.

In contrast to the psychometric model represented by CHC Theory, the Parietal-Frontal Integration Theory (P-FIT) approaches intelligence more directly as a function of the integration of brain networks. The greater the efficiency of these networks – the higher the level of intelligence that will be displayed. Based in neuroimaging studies evaluating intelligence, Jung and Haier (2007) proposed that intelligence arises out of the integration of grey matter areas of the frontal and parietal regions and the white matter connections between these regions.[4] Grey matter is comprised of cell bodies where processing takes place and white matter, composed of myelinated (or insulated) axons, provides rapid transfer of information between networks. Vakhtin and colleagues (2014), with some specificity, reported that "The regions involved in the networks that were… consistent with the P-FIT, localiz[ed] to the bilateral medial frontal and parietal regions, right superior

---

[2] Floyd, R. G., Farmer, R. L., Schneider, W. J., & McGrew, K. S. (2021a). Theories and measurement of intelligence. In L. M. Glidden, L. Abbeduto, L. L. McIntyre, & M. J. Tasse' (Eds.), *APA Handbook of Intellectual and Developmental Disabilities*. *vol. 1, Foundations* (Vol. 1, pp. 385–424). American Psychological Association.

[3] American Association on Intellectual and Developmental Disabilities. (2021). *Intellectual disability: Definition, classification, and systems of supports* (12th ed.), p. 27. (Hereafter, AAIDD).

[4] Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135–154. https://doi.org/10.1017/s0140525x07001185

frontal lobule, and the right cingulate gyrus."[5] Adding to the understanding of the P-FIT model, Hearne et al. (2016) reported that their results "revealed a novel contribution of across-network interactions between default-mode and fronto-parietal networks to individual differences in intelligence at rest. Specifically, we found that greater connectivity in the resting state was associated with higher intelligence scores."[6] [7] Moreover, these studies support the connection between psychometric *g* and the integration of these regions of the brain in that activation and structural features of these regions are associated with scores on IQ tests.

Given this understanding of the nature of intelligence we can now turn our attention to understanding the nature of intellectual disability (ID). ID has been termed a "status condition," that is, it is defined by levels of functioning in the areas of intellectual ability and adaptive behaviors with onset during the developmental period. As such, ID does not imply a specific etiology and in fact most cases of mild ID do not necessarily have an identifiable cause. Nonetheless, both biological and environmental factors have been shown to be risk factors for ID.[8]

---

[5] Vakhtin, A. A., Ryman, S. G., Flores, R. A., & Jung, R. E. (2014). Functional Brain Networks contributing to the Parieto-frontal integration theory of Intelligence. *NeuroImage*, *103*, 349–354. https://doi.org/10.1016/j.neuroimage.2014.09.055 . p. 1.

[6] Hearne, L. J., Mattingley, J. B., & Cocchi, L. (2016). Functional brain networks related to individual differences in human intelligence at rest. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep32328

[7] The Default Mode Network (DMN) is a network of brain regions active when one is not actively engaged in a task (resting), such as when one's mind is "wandering."

[8] Magana, S. & Vanegas, S.B. (2021). Culture, race, and ethnicity and intellectual developmental disabilities. In L. M. Glidden, L. Abbeduto, L. L. McIntyre, & M. J. Tasse' (Eds.), *APA Handbook of Intellectual and Developmental Disabilities. vol. 1, Foundations* (Vol. 1, pp. 355-382). American Psychological Association. See also Floyd et al.

Current criteria for the diagnosis of ID are provided by the American Association on Intellectual and Developmental Disabilities (AAIDD), in the manual *Intellectual Disability: Definition, Diagnosis, Classification, and Systems of Support, 12th Edition,* the American Psychiatric Association (APA) in the *Diagnostic and Statistical Manual of Mental Disorder, 5th Edition, Test Revision (DSM-5-TR),* and, though less widely used, the World Health Organization (WHO) in the *International Classification of Diseases, 11th edition (ICD-11; 2018)*. Importantly, ID is also commonly defined by state statutes that often parallel the diagnostic criteria of AAIDD and APA. The AAIDD and APA diagnostic criteria are quite similar using a three prong definition of ID. For example, AAIDD reports, "Intellectual disability (ID) is characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills. This disability originates during the developmental period, which is defined operationally as before the individual attains age 22."[9] APA adds more detail to Prongs 1 and 2, describing the deficits in intellectual functions to include "…reasoning, problem solving, planning, abstract thinking, judgment, academic learning, and learning from experience…"[10] which are confirmed by individualized, standardized IQ testing. APA further reports that "Deficits in adaptive functioning …result in failure to meet developmental and socio-cultural standards for personal independence and social responsibility. Without ongoing support, the adaptive deficits limit functioning in one or more activities of daily life such as

[9] AAIDD, p. 1.
[10] American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders: DSM-5-TR*. American Psychiatric Association Publishing, p. 37. (Hereafter, DSM-5-TR).

communication, social participation, and independent living, across multiple environments, such as home, school, work, and community."[11] Conversely, APA does not specify when the developmental period ends as AAIDD does.

Operationally, both organizations define deficits in intellectual functioning as an IQ on an individually-administered IQ test of approximately two standard deviations below the population mean, considering the margin of error. Virtually all IQ tests have a mean of 100 and a standard deviation of 15 such that two standard deviations below the mean would be an IQ of 70 ± 5 points considering the margin of error. This margin of error is based on the Standard Error of Measurement (SEM) for the overall IQ score (which in turn is derived from the internal consistency reliability of the test).[12]

The most current versions of widely used, comprehensive, individually administered IQ tests include the Woodcock Johnson, 5th Edition (WJ V), published in 2025, the Wechsler Adult Intelligence Scale, 5th Edition (WAIS-5), published in 2024, and the Stanford-Binet, 5th Edition, published in 2003. Each of these are consistent with the recommendation of AAIDD which endorsed using a comprehensive test of general intelligence which "should (a) include at least six subtests, and (b) sample at least three (preferably more) CHC broad-strata abilities."[13]

---

[11] *Id*.

[12] The SEM (Standard Error of Measurement) estimates the consistency of observed scores if the person took the test repeatedly (without practice effects). As such it provides an estimate of the amount of error in a person's observed test score. It is calculated as follows:

$$SEM = \text{SD} \times \sqrt{1 - r_{xx}}$$

Where SD = Standard Deviation (i.e., 15) and r_{xx} = Reliability coefficient of the scale (WAIS 5 = 0.97).

[13] AAIDD, p. 29.

The Woodcock Johnson Tests of Cognitive Abilities, Fifth Edition, is the most recently published IQ test. It has 20 subtests that can be used to derive 12 different indices. Importantly, the WJ V measures 7 of the primary CHC factors described above (missing only Auditory Processing (*Ga*)). As such, it measures more factors of intelligence than any other battery. This edition of the Woodcock Johnson is an entirely digital assessment instrument, requiring a laptop and iPad for test administration. Moreover, the WJ 5 requires an active internet connection for administration and thus may not be appropriate for use in a prison setting. **ADMINISTER ONLINE?**

The WAIS-5 is the latest iteration of the Wechsler scales first developed by David Wechsler. The WAIS-5 full scale IQ (FSIQ) represents the best measure of general intelligence (*g*) derived from this instrument. The FSIQ is derived from the scores on just 7 of the subtests – Similarities, Block Design, Matrix Reasoning, Digit Sequencing, Coding, Vocabulary, and Figure Weights. This is a change from the WAIS-IV which required 10 subtests to derive the full scale IQ. Given this change, the WAIS-5 FSIQ puts less emphasis on measures of cognitive efficiency (working memory & processing speed) and more emphasis on fluid reasoning than did the WAIS-IV. Given this change, it is possible that there might be changes in IQ scores based on the changes in content and structure for the WAIS-5. However, according to the WAIS-5 Technical and Interpretive Manual, at the group level, “All the index score means are lower on the WAIS-5 than on the WAIS-IV, consistent with the Flynn effect (see below)…At the subtest level, standard differences exceeding .20

are found for Similarities, Matrix Reasoning, Figure Weights, and Symbol Search"[14] with the WAIS-5, on average, having lower scores.

The WAIS-5 has a total of 20 subtests that can be used to derive 22 indices. The primary composite scales include the following indices: Verbal Comprehension (VCI), Visual Spatial (VSI), Fluid Reasoning (FRI), Working Memory (WMI), and Processing Speed (PSI) – assessing five primary CHC domains. Ancillary Summary Indices include those expanding each of the primary indices but also those for comparisons between general ability (General Ability Index (GAI)) representing core intellectual abilities and cognitive proficiency, focused on working memory and processing speed. Other ancillary indices include measures of nonverbal ability and nonmotor abilities. The WAIS-5 can be administered as a paper and pencil test, in a digital format using two iPads (reducing the possibility of scoring errors), or remotely with some adaptations. It does not require an active internet connection for administration. Studies have shown that the paper/pencil version and the digital version are equivalent.

The Stanford Binet, published in 2003, should be considered, at this point, obsolete for capital litigation given the impact of norm obsolescence (the Flynn effect – see below) on obtained IQ scores.

In the context of *Atkins* litigation, a number of statistical concepts are important to understand. The first of these is the reliability of the test. Reliability measures the extent to

[14] Wechsler, D., Raiford, S. E., & Presnell, K. (2024). *Wechsler Adult Intelligence Scale: Technical and interpretive manual* (5th ed.). NCS Pearson, p. 86.

which the test measures a construct consistently. Reliability is measured as either internal consistency reliability or test-retest reliability. Most modern IQ tests have internal consistency reliability coefficients of 0.95 or higher where 0.00 would be entirely unreliable and 1.00 would be perfectly reliable. Reliability is most often an issue when considering tests that were given years prior and may have not been exceptionally reliable tests, which significantly expands any confidence interval.

It is also important to understand the Standard Error of Measurement (SEM), described above, but also the Standard Error of Estimate (SEE). These indices allow for developing confidence intervals around the obtained score, in which the "true score" is likely to fall. While the literature on ID is entirely focused on the role of the SEM in the identification of ID, test publishers generally use the SEE as the means to create confidence intervals around an obtained IQ score. It then must be reiterated that the score on an IQ test is likely to contain error variance due to differences in test environments, test administration procedures, and examiner and examinee characteristics.[15] As a result, the IQ score obtained is bracketed by a confidence interval. The use of a 95% confidence interval is standard within the field of assessment. It is generally accepted that confidence intervals of ± 5 points provide a reasonable degree of certainty as to the underlying (or

[15] For the assessment of ID, it is critical that tests with high reliability are used to determine IQ. With lower levels of reliability, as might be seen on a group test or a brief IQ test, the band of error increases dramatically making it difficult to determine with confidence the actual level of intellectual ability. Kranzler and Floyd (2013) cited by Floyd et al. (2021), p. 398 recommended an internal consistency coefficient of at least 0.95 for IQ tests.

latent) "true score."[16] The use of the SEM provides equal intervals above and below the obtained score in which the true score is believed to fall. On the other hand, the SEE, bracketing scores within the ID range, sometimes provides unequal intervals around the obtained IQ – resulting in a greater probability that the true score will be closer to the mean of the population rather than below the obtained score. This is because of regression to the mean effects, that is, given how extreme the score is, it is more probable that the true score is closer to the mean of the population than below the obtained score. The practical consequences of this difference between using the SEM and SEE are usually minimal but can become important in specific circumstances. For example, assume an individual obtains an IQ of 75 on the WAIS-5. The 95% confidence interval around that score would be 75 $\pm$ 5 – approximately two standard deviations below the mean considering the SEM and sufficient for meeting the Prong 1 requirement. However, if instead, the true score is estimated to be 75.75 and using the SEE, the 95% confidence interval would be 71$\pm$ 5, or between 71 and 81 – just missing the two standard deviations below the mean mark.[17] In

[16] Floyd, R. G., Farmer, R. L., Schneider, W. J., & McGrew, K. S. (2021). Theories of measurement of intelligence. In L. M. Glidden, L. Abbeduto, L. L. McIntyre, & M. J. Tassé (Eds.), *APA Handbook of Intellectual and Developmental Disabilities* (Vol. 1, pp. 385–424). American Psychological Association. , p. 397.

[17] The formula for the Estimated True Score is:

$$Estimated\ True\ Score = SD\ (r_{xx})(X - 100),$$

where SD is the standard deviation (15) and (r_{xx}) is the reliability coefficient (0.97 for the WAIS-5).

The formula for the SEM is:

$$SEM = SD\sqrt{1 - r_{xx}}$$

The formula for the SEE is:

$$SEE = SD\ (r_{xx})\sqrt{1 - r_{xx}}$$

addition, it can always be argued, when using the SEE, that it is more likely that the true score is higher than that which was obtained.

It is also of importance to understand the nature of practice effects on IQ testing. Practice effects refer to the phenomenon of increasing scores with each new administration of the same or similar tests. Clients who have taken a Wechsler scale will almost always remember the blocks used for the Block Design subtest and recalling items or procedures will likely improve their performance on re-testing. In the case where the same IQ test has been administered multiple times, later tests will likely have inflated obtained test scores. Scharfen et al. (2018) completed a meta-analysis (a study of studies) on the impact of repeat testing on cognitive ability test scores. This research found "A gain in test scores of almost a third of a standard deviation for a mere repetition of a cognitive ability test…"[18] This would mean that if an individual obtained a WAIS 5 FSIQ of 71 on one occasion, the score would likely be 76 on a subsequent administration. Moreover, "An increase in test scores of half a standard deviation for the third and fourth test were [also] revealed."[19] As an antidote to combat this phenomenon, the authors recommended administration of a different IQ test for subsequent testing. In addition, longer retest intervals partially combat this problem though these authors found that even waiting nearly a year decreased the practice effect by just a single point. Statistical software provided by

---

[18] Scharfen, J., Peters, J. M., & Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence*, *67*, 44–66. https://doi.org/10.1016/j.intell.2018.01.003
[19] *Id*.

the publisher can sometimes be used to calculate whether there has been any real change in the IQ scores or whether the improvement is simply expected by the practice effect.[20]

Next, we must address the issue of norm obsolescence, commonly referred to as the Flynn effect, named after the researcher James Flynn, who wrote extensively about the subject. The Flynn effect refers to the progressive inflation of IQ scores over time as norms for a test become outdated. The implication is that successive generations are getting smarter. Two separate meta-analyses published in 2014 and 2015 established that, within the United States, IQ scores have been increasing at a rate of approximately 3.0 points per decade throughout the 20th century.[21] This prompted AAIDD to stipulate that "Current best practice guidelines recommend that in cases in which an IQ test with aged norms is used as part of a diagnosis of ID, a correction of the full-scale IQ score of 0.3 points per year since the test norms were collected is warranted."[22] However, there has been a suspicion that the Flynn effect may have been moderating, or even reversing, especially within European nations.[23] The recent publication of the WAIS-5, added substantially to the view

---

[20] The WAIS-IV scoring software provided analyses for serial assessments.

[21] Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin*, *140*(5), 1332–1360. https://doi.org/10.1037/a0037173 ; Pietschnig, J., & Voracek, M. (2015). One Century of global IQ gains: A formal meta-analysis of the Flynn effect (1909-2010). *Perspectives on Psychological Science*, *10*(3), 282–306. https://doi.org/10.1177/1745691615577701

[22] Schalock, R. L., Luckasson, R., & Tassé, M. J. (2021). *Intellectual disability: Definition, diagnosis, classification, and systems of supports* (12th ed.). American Association on Intellectual and Developmental Disabilities, p. 42.

[23] Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences*, *115*(26), 6674–6678. https://doi.org/10.1073/pnas.1718793115 ; Dutton, E., van der Linden, D., & Lynn, R. (2016). The negative Flynn effect: A systematic literature review. *Intelligence*, *59*, 163–169. https://doi.org/10.1016/j.intell.2016.10.002 ; Teasdale, T. W., & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence*, *36*(2), 121–126. https://doi.org/10.1016/j.intell.2007.01.007

that the Flynn effect is also moderating in the United States.[24] Using the results of individuals tested on both the WAIS-IV and the WAIS-5, the Flynn effect was estimated to be only 1.3 points per decade for the Full Scale IQ, the VCI, and the FRI – considerably lower than the 3.0 points previously identified for the Flynn effect. In addition, some have raised concerns that capital clients are scoring higher on the WAIS-5 than on the WAIS-IV. This is certainly possible in individual cases, however, the corrected correlation for the FSIQ between the instruments was .92 – essentially equal to the WAIS-5 test-retest correlation of .93. In a sample of 186 individuals tested on each instrument in a counterbalanced fashion, the mean FSIQ for the WAIS-5 was 99.7 (SD = 14.5) and for the WAIS-IV it was 100.9 (14.3) – consistent with the relatively modest Flynn effect.[25]

REVIEW THE RECENT WAIS-5 ARTICLE FOR INCLUSION

Given the data on practice effects and the apparently moderating Flynn effect, questions are raised as to when and if to retest a client, and what instrument to use. The practice effect obtained when repeating the same or similar test may result in a higher obtained IQ score, though statistical analysis can sometimes determine if any change is a real change in the underlying level of intellectual functioning. Nonetheless, in such cases it may be more appropriate to use an alternative test – though given the limitations in current instruments that may be difficult. On the other hand, faced with the choice of administering the WAIS-5 or the WAIS-IV, then answer is likely to use the newer test. If one

[24] Wechsler et al., p. 86.
[25] *Id.,* p. 85.

were to administer the WAIS-IV at this time, there would likely be heated argument over the current value of the Flynn effect.[26] Should 0.3 points or 0.13 points per year be used in the calculation of a Flynn-corrected score? Such a fight would muddy any determination of the actual IQ value.

The second prong of the diagnostic criteria for the diagnosis of ID requires deficits in adaptive functioning or behaviors. According to DSM-5-TR:

> Adaptive functioning involves adaptive reasoning in three domains: conceptual, social, and practical. The *conceptual (academic) domain* involves competence in memory, language, reading, writing, math reasoning, acquisition of practical knowledge, problem solving, and judgment in novel situations, among others. The *social domain* involves awareness of others' thoughts, feelings, and experiences, empathy; interpersonal communication skills; friendship abilities; and social judgment, among others. The *practical domain* involves learning and self-management, recreation, self-management of behavior, and school and work task organization, among others.[27]

DSM-5-TR does not specify that the deficits need to be "significant" but just that there is sufficient impairment in one of the domains of adaptive functioning as to require ongoing support. In contrast, AAIDD does require significant deficits in adaptive behavior, which is operationalized as "an adaptive behavior score that is approximately 2 standard deviations or more below the mean in at least one of the three adaptive behavior domains:

---

[26] The standard within the field, with some exceptions, is that examiners should shift to a newer test within one year of publication.
[27] DSM-5-TR, p. 42.

conceptual, social, or practical, considering the standard error of measurement for the specific, individually administered instrument used."[28] DSM-5-TR does specifically acknowledge that, "Standardized measure are used with knowledgeable informants (e.g., parent or other family member; teacher; counselor; care provider) and the individual to the extent possible (emphasis added)."[29] The evaluation can be completed using both clinical evaluation and standardized testing. However, given that the diagnostic criteria require that the deficits in intellectual and adaptive behaviors be present during the developmental period, it is commonly necessary to do a retrospective evaluation of adaptive behaviors. This is made difficult in that none of the adaptive functioning measures were developed or normed for this type of evaluation. AAIDD does indicate that such a retrospective analysis is possible and that the diagnosis need not have been made during the developmental period. In such circumstances, clinical judgment is a critical component in assessing adaptive behaviors and in determining ID, based on a thorough bio-psycho-social history.

The development of that bio-psycho-social history is a critical component in the assessment of adaptive behaviors. This task often involves initial interviews by a mitigation investigator – identifying those individuals who have sufficient knowledge of the individual as to provide valid information regarding their adaptive behaviors. It is important to recognize that the focus of such an investigation in cases of ID is on what adaptive behaviors the individual actually performed. This is distinct from many such investigations

[28] AAIDD, p. 31.
[29] DSM-5-TR, p. 42.

that focus on the trauma the individual might have experienced during their development. A thorough review of records establishes a foundation for the subsequent stages of the evaluation. These records should provide information regarding educational, medical, psychiatric, work, legal, family, and substance abuse histories.

The next phase of an adaptive behavior evaluation will involve collateral interviews of those who knew the client well. Again, the focus of the interview will include what the individual actually did in their daily life in terms of adaptive behaviors. Understanding of the neighborhood and community in which they grew up may also provide context for the evaluation. Interviews with teachers and school administrators are particularly valuable. This will allow for understanding the specific practices in those school districts which the student attended – were children who were ID routinely instead identified as SLD? The next phase may involve administration of adaptive behavior instruments. Currently used measures of adaptive behaviors include the Adaptive Behavior Assessment System – 3 (ABAS-3), the Adaptive Behavior Diagnostic Scale (ABDS), the Diagnostic Adaptive Behavior Scale (DABS), and the Vineland Adaptive Behavior Scale – 3 (VABS-3). These require 3rd party reporters who were familiar with the individual during the developmental period. The procedures for administration vary, with the ABAS-3 recommended as an inventory that the reporter completes. The ABAS-3 can be administered in person or via an online platform. In contrast, the Vineland is commonly completed as a semi-structured interview that allows the examiner to complete ratings of the individual based upon their judgment. The Vineland requires greater skill on the part of the examiner than does the ABAS-3. On the other hand,

those completing the ABAS-3, who are often family members, can be accused of bias in their ratings. The ABDS is unique in that it allows for the use of supplemental raters should the primary rater not have had the opportunity to observe particular behaviors, though each must be rating the individual based upon the same age.

Given the complexities of the knowledge bases required for the investigation of intellectual and adaptive behavior functions, separate experts may be required to complete the evaluation of prong 1 and prong 2. Close collaboration between these two experts is ideal. This is particularly important when each expert is relying on the information provided by the other to formulate their opinion. In such cases, prosecutors will commonly attempt to undermine each expert's opinion through criticism of the other expert's technique or qualifications. In practice, it can be helpful for the prong 1 expert to participate in the adaptive function interviews conducted by the prong 2 expert – inoculating to some extent against any such criticism.

The final prong involved in the diagnosis of intellectual disability is the requirement that the deficits in intellectual and adaptive behaviors had their onset during the developmental period. Neuroscience research has established that brain development continues into the mid-20s. However, the developmental period is often defined by state statutes. Moreover, DSM-5-TR does not specify an age when the developmental period ends though noting that this requirement "refers to recognition that intellectual and

adaptive deficits are present during childhood and adolescence."[30] Conversely, AAIDD does specify that the onset of ID occurs "before the individual attains age 22."[31] In cases in which a diagnosis of ID was not made during the developmental period, clinical judgment is essential in making a current diagnosis. In evaluating both prong 1 and prong 2, some judgment must be made as to whether a current IQ test score and level of adaptive function represents that which would have been present during the developmental period. Factors that might be considered in these judgments would be the timing of the onset of severe psychiatric illness, the presence of substance abuse, any history of traumatic brain injury or other neurological condition, the individual's medical history including such conditions as diabetes, seizures, strokes, which can affect cognitive functioning, etc. A review of educational records can also be important – for example, was the individual in Special Education and was there an individualized education program (IEP) and under what condition did the individual qualify for Special Education. In California, because of the *Larry P. v. Riles* (495 F. Supp. 926 (N.D. Cal. 1979)) case, African American school children could not be given IQ tests to determine eligibility for Special Education. These children would often be identified as having a Specific Learning Disability (SLD) rather than ID. It is important to be aware of these specific state factors. In addition, the value of prior evaluations conducted during the developmental period cannot be overstated – these give direct information regarding the individual's functioning during this time. Further, every

[30] *Id.,* p. 42.
[31] AAIDD, p. 1.

effort should be made to secure any actual test protocols that might still exist as this would allow for re-scoring of the test to confirm the validity of the results.

In summary, *Atkins* and its progeny provide guidance as to the nature of the evaluation that must be performed in cases in which intellectual disability (ID) is suspected. This involves investigation and creation of a detailed bio-psychosocial history that includes information about the individual, their family, their neighborhood, and their educational system. Requirements for establishing the diagnosis of ID include valid, comprehensive, individualized assessment of intellectual functioning, a valid and comprehensive evaluation of the individual's adaptive behavior/functioning, and confirmation that any deficits in intellectual and adaptive behaviors had their onset during the developmental period. This is a time consuming and complicated process, and it behooves attorneys to have a basic understanding of the parameters of such an undertaking.

Dale G. Watson, Ph.D.
Forensic Neuropsychologist
3377 Deer Valley Road, PMB 310
Antioch, CA 94531
watson.dale@comcast.net
866-536-5301 (Service)
925-989-8874 (mobile)