



Perspective Paper

From stigma to science: Rethinking “Intelligence” in the neuropsychology of epilepsy

David W. Loring^{1,2}  and Bruce P. Hermann³

¹Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA, ²Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA and ³Department of Neurology, University of Wisconsin School of Medicine & Public Health, Madison, WI, USA

Abstract

Objective: To review the historical, conceptual, and ethical foundations of intelligence testing in neuropsychology and to consider whether alternative cognitive performance labels offer greater conceptual precision while reducing stigma. **Method:** We conducted a narrative review of early twentieth century cognitive assessments, tracing the evolution of intelligence testing and its intersections with eugenic ideology. Key examples include the Army Alpha and Beta tests administered during World War I and Ellis Island immigration assessments, which were frequently interpreted without consideration of cultural or educational influences. We examine how these practices informed early interpretations of neuropsychological performance, particularly in individuals with epilepsy, and shaped initial characterizations of neurologically based cognitive abilities. **Results:** Early intelligence testing was grounded in the belief that intelligence was a fixed and directly genetically determined trait. Test performance was interpreted as an index of biological superiority, lending scientific legitimacy to eugenic ideologies and reinforcing stigma toward individuals with epilepsy. Although modern frameworks emphasize multidimensional cognitive abilities, intelligence-based characterization persists and continues to be frequently reported as a primary outcome of neuropsychological testing. **Conclusions:** In contexts that require a single summary indicator of cognitive performance, labels such as Total Cognitive Composite are recommended since they avoid implying a fixed or unitary capacity. Continued reliance on the construct of “intelligence” is inconsistent with contemporary models of cognition, reflects outdated theoretical assumptions, and carries enduring psychosocial stigma. Moreover, its circular and internally inconsistent definitions substantially limit its validity and appropriateness within contemporary adult clinical neuropsychological practice.

Keywords: Cognitive assessment; epilepsy; eugenics; intelligence testing; IQ limitations; social determinants of neuropsychology

(Received 15 August 2025; final revision 6 January 2026; accepted 7 January 2026; First Published online 19 February 2026)

Statement of Research Significance

Research Question(s) or Topic(s): This paper explores how historical misuse of intelligence testing contributed to stigma and biased interpretations of cognitive function, particularly among marginalized groups and individuals with epilepsy. It examines the early connection between intelligence testing and eugenic policies and how these influences persist in modern neuropsychology. **Main Findings:** Intelligence tests were historically used to justify segregation, sterilization, and immigration restrictions, reinforcing social hierarchies and clinical bias. Although contemporary frameworks recognize cultural and environmental factors, using the term “intelligence” still suggests that cognitive ability is fixed and innate. Primarily reliance on IQ scores obscures domain-specific strengths and underestimates variability in cognitive performance across contexts and cultures. **Study Contributions:** Replacing “intelligence” with labels such as “Total Cognition Composite” reflects a multidimensional, ethically grounded approach to neuropsychological assessment, avoids implying a unitary cognitive capacity, and promotes reduction of stigma.

Introduction

Since the emergence of psychology as a distinct scientific and clinical discipline in the late nineteenth and early twentieth centuries, intelligence testing has been a primary component of psychological assessment, serving as one of the earliest standardized,

quantitative methods for characterizing individual differences. This paper examines the historical, conceptual, and ethical foundations of intelligence testing in neuropsychology, with a particular focus on its application to individuals with epilepsy. The primary goals are to trace how early twentieth century assumptions

Corresponding author: David W. Loring; Email: dloring@emory.edu

Cite this article: Loring DW and Hermann BP. (2026) From stigma to science: Rethinking “Intelligence” in the neuropsychology of epilepsy. *Journal of the International Neuropsychological Society* 32: 213–220, <https://doi.org/10.1017/S1355617726101842>

© The Author(s), 2026. Published by Cambridge University Press on behalf of International Neuropsychological Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

about intelligence, including its reification as a genetically fixed and unitary trait, caused harms and provided scientific support for eugenics.

Because the term intelligence remains embedded in contemporary neuropsychological practice, we examine whether continued use of “intelligence” is conceptually justified in light of modern multidimensional models of cognition, while acknowledging that global summary scores can serve as pragmatic tools for summarizing cognitive performance in specific contexts without implying a fixed or unitary construct. Finally, we argue for more precise, descriptive alternatives to intelligence based terminology and discuss the implications of terminology reform for assessment, research, and clinical decision making.

From good to bad intentions: Origins and misappropriation of intelligence testing

The foundation for viewing intelligence as a fixed and heritable trait associated with direct Mendelian inheritance can be traced to Francis Galton, who coined the term “eugenics” in the late nineteenth century and argued that mental ability was largely inherited, drawing on the evolutionary principles developed by his half-cousin, Charles Darwin (Galton, 1883; Paul & Moore, 2012). Galton generalized the concept of evolution to include social policy, promoting the selective breeding of individuals believed to possess superior traits (i.e., positive eugenics). In *Hereditary Genius*, Galton defined eugenics as “the study of all agencies under human control which can improve or impair the racial quality of future generations” (Galton, 1883). Galton sought to operationalize these ideas by promoting anthropometric measurement practices aimed at ranking individuals and populations, which he argued provided a scientific basis for social stratification and reproductive control.

Alfred Binet and Théodore Simon designed the first formal cognitive assessment with the explicitly designed aim to assess functional performance. Commissioned by the French Ministry of Education, the Binet-Simon scale was designed to identify school aged children who were performing poorly academically and who could benefit from targeted educational support. Binet and Simon emphasized that their test was “to determine, in a practical manner, the degree to which a child is able to profit from instruction. . . . It is not a question of measuring innate intelligence, but of evaluating the child’s performance” (Binet & Simon, 1908/1916).

This distinction was abandoned with the scale’s introduction to the United States, where Henry Goddard’s 1910 translation and adaptation reframed it as a “scale of intelligence” (Goddard, 1910, 1913). Through his work at the Vineland Training School and later at the Ohio Committee on the Sterilization of the Feeble Minded, Goddard used test results to classify individuals into formal categories using now offensive labels (e.g., “moron,” “imbecile”). While Binet developed the test to identify students who would benefit from additional educational support, Terman applied it to identify high performing individuals and to characterize “superior intelligence” (Terman, 1925). As a result, the focus on high performing individuals reinforced the belief that IQ scores provided a definitive and comprehensive representation of intellectual capacity at all levels of ability.

The seduction of quantification

The emergence of IQ testing in the early 20th century coincided with major advances in statistical methods, including correlation and factor analysis. Figures such as Karl Pearson, Charles

Spearman, and Ronald Fisher developed techniques that later became foundational to modern data science. Applied to intelligence scores, these methods conveyed an appearance of scientific rigor while, within the social context of the time, reinforcing existing patterns of social stratification.

Karl Pearson, the first Galton Chair of Eugenics at the University of London, applied correlation to study heredity and human characteristics, asserting that intelligence and other biological attributes were inherited and statistically predictable (Pearson, 1907). Charles Spearman observed the “positive manifold,” the pattern whereby individuals’ scores across different cognitive tests tend to correlate, which prompted him to propose a general intelligence factor, or “g,” as part of his Two-Factor Theory of intelligence (Spearman, 1927).

Ronald A. Fisher succeeded Karl Pearson’s as Chair of Eugenics at the University of London, where he also served as editor of the *Annals of Eugenics* and developed statistical techniques that became central to eugenic research (Fisher, 1930). Recognized primarily for creating analysis of variance, with the F being his initial, Fisher actively championed eugenic causes throughout his career. As a member of the British Eugenics Society’s executive committee, he advocated for policies designed to encourage reproduction among individuals he deemed genetically “superior” (he had eight children) and opposed efforts to separate the movement from its discriminatory foundations.

Eugenics: The scientific rationalization of social hierarchy

During the early 1900s, the eugenics movement expanded throughout the United States with the stated goal of “enhancing” the nation’s genetic quality through controlled reproduction and the systematic exclusion of individuals labeled as biologically deficient. Proponents drew parallels to agricultural practices, contending that just as selective breeding enhanced crops and livestock, similar methods could be applied to human populations. Affiliations with universities and emerging disciplines such as statistics, psychology, and cognitive assessment provided the movement with the appearance of scientific legitimacy, extending its influence to education, medical practice, and governmental policy (Figure 1).

The Eugenics Record Office (ERO) at Cold Spring Harbor, New York, established by biologist Charles Davenport and directed by Harry H. Laughlin, was at the center of the American eugenics movement (Allen, 1986). The ERO gathered and examined family genealogies (Figure 2), institutional records, and intelligence test results to develop accounts of inherited “deficiency,” thereby reinforcing hierarchies based on race, class, and disability status. Its findings influenced public health and educational policies in ways now recognized as discriminatory.

The Second International Congress of Eugenics, held in 1921, symbolized the movement’s mainstream acceptance and international ambitions (DeSalle, 2021; Kevles, 1985). Henry Fairfield Osborn, president of the American Museum of Natural History, opened the Congress remarking “the right of the state to safeguard the character and integrity of the race or races on which the future depends is, to my mind, as incontestable as the right of the state to safeguard the health and morals of the people” (Osborn, 1921). He argued that science must guide government in preventing the “multiplication of worthless members of society.”

As eugenic principles gained wider implementation, compulsory sterilization became a central policy. Beginning with Indiana’s 1907 law, more than 30 states enacted statutes targeting those

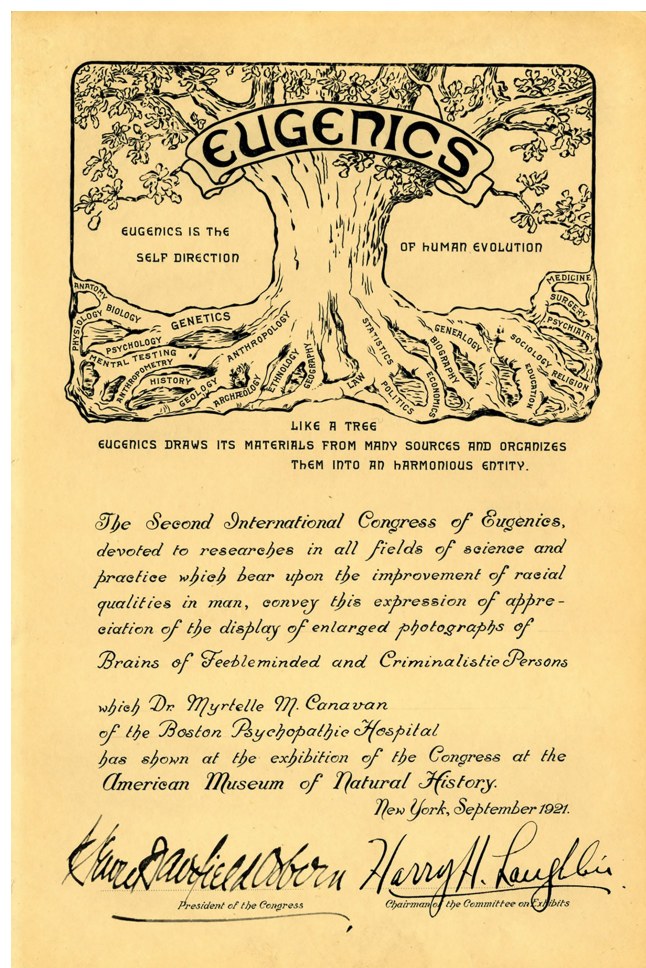


Figure 1. Program from the Second International Congress of Eugenics (1921) proclaiming that “eugenics is the self-direction of human evolution.” The tree’s roots depict disciplines deemed foundational to the eugenics movement including psychology, mental testing, and statistics highlighting the perceived scientific legitimacy drawn from these emerging fields to justify eugenic ideology and policy “which bear upon the improvement of racial qualities in man.” Public domain.

deemed unfit, guided by the ERO’s Model Eugenical Sterilization Law (Laughlin, 1922). These measures were widely endorsed by medical and public health authorities and later influenced Nazi Germany’s 1933 sterilization law (Sofair & Kaldjian, 2000). American eugenics thus shaped both domestic policy and international practice, with the belief in sterilization as a public good persisting until the last U.S. statute was repealed in 2013 (West Virginia Legislature, 2013).

From army testing to classification bias

Intelligence testing extended beyond institutional settings to the military, where Robert Yerkes designed the Army Alpha test for literate recruits and the Army Beta test, a nonverbal, picture-based measure, for illiterate or non-English-speaking recruits (Yerkes, 1921). Developed for rapid wartime assessment, these tests were administered to 1,726,966 soldiers between 1917 and 1919. Although the Army Beta acknowledged the impact of education and sociocultural background, scores from both tests were treated as equivalent and combined to generate overall intelligence rankings.

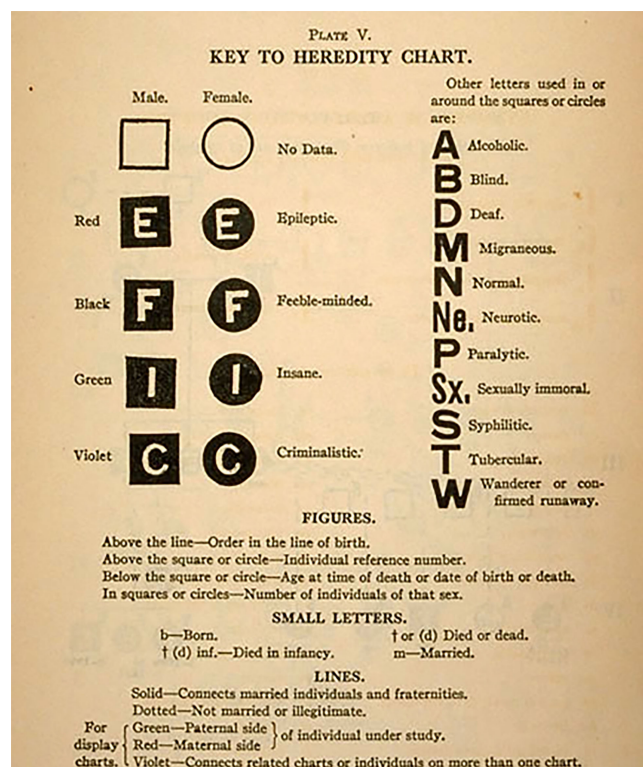


Figure 2. Heredity chart from early 20th century categorizing individuals by sex, ancestry, and alleged hereditary traits. Conditions marked as inheritable include “feeble-mindedness,” “criminalistic” behavior, “insanity,” and “epilepsy,” which was red symbols labeled “E.” These notations were used in the construction of family pedigrees. Reproduced from Davenport, C. B. et al. *The Study of Human Heredity: Methods of Collecting, Charting and Analyzing Data*. Eugenics Record Office, Cold Spring Harbor, NY (1911). Public domain.

Princeton psychologist Carl Brigham used test results to rank immigrant groups by IQ and promote hierarchical views of intelligence by nationality (Brigham, 1923; see Figure 3). In his original analysis, native-born white Americans of Northern and Western European descent scored highest, whereas immigrants from Southern and Eastern Europe scored lower and were labeled intellectually inferior. Brigham later retracted these conclusions, acknowledging the cultural biases inherent in the tests that were not previously considered, explicitly stating “test scores may not represent unitary things” (Brigham, 1930).

Intelligence testing in the U.S. played a key role in shaping immigration policy, reinforcing discriminatory ideas about race, intelligence, and national identity. Laughlin’s influence helped embed these views in the Johnson-Reed Immigration Act of 1924, which barred entry to groups labeled “intellectually and racially inferior” (Lombardo, 2008). The eugenics movement’s promotion of IQ testing provided supposed scientific justification for a heredity-focused, meritocratic vision, mirroring the arguments of Madison Grant’s *The Passing of the Great Race* (Grant, 1916).

Intelligence testing in the early twentieth century served as a mechanism of immigration control and exclusion at Ellis Island and other U.S. ports of entry. Under Howard A. Knox, cognitive assessments, including nonverbal, performance-based measures, were implemented to manage immigrant intake and reduce apparent cultural and linguistic bias. Nevertheless, these approaches remained embedded in hereditarian assumptions, and, as Richardson (2011) documented, routinely misclassified individuals in ways that reinforced existing social hierarchies.

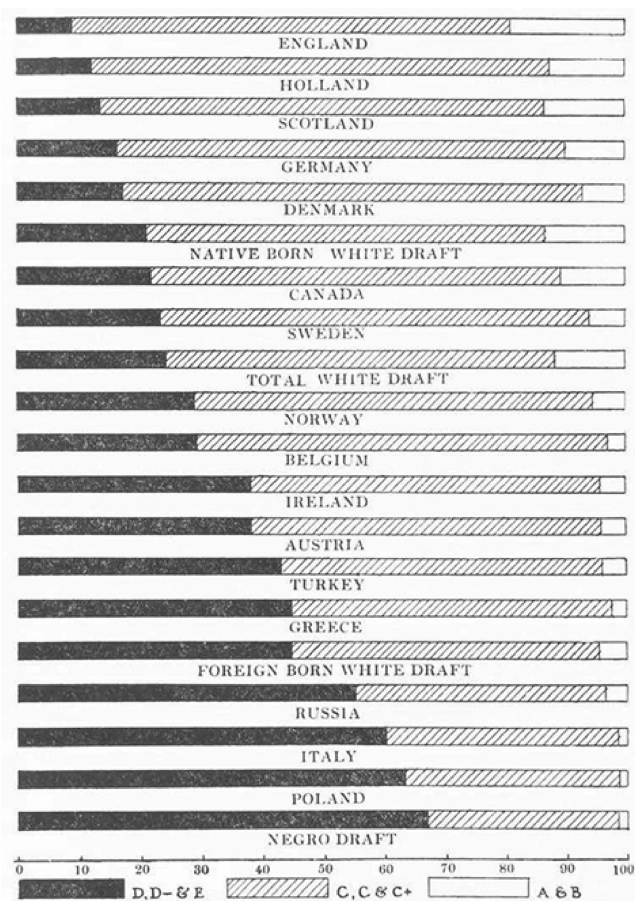


Figure 3. Ranking of immigrant groups by average IQ scores based on Army Alpha and Beta test results. These results were widely circulated to support eugenic arguments and immigration restrictions despite being heavily biased by language, education, and cultural familiarity. Reproduced from Carl C. Brigham, *A Study of American Intelligence* (Princeton University Press, 1923). Public domain.

Intelligence, eugenics, and the epileptic other

Religious and charitable organizations initially established epilepsy colonies to provide residents with specialized care, education, and vocational training. These communities were designed as supportive, self-contained environments where individuals with epilepsy could live, work, and receive treatment. Over time, however, several colonies were co-opted by eugenicists, who framed epilepsy as a hereditary defect threatening the “germ plasm” and used these institutions to justify segregation and restrict reproduction among those deemed genetically “unfit.” Notable examples include the *Monson State Hospital for Epileptics* in Massachusetts and the *New Jersey State Village for Epileptics* in Skillman, which embodied both the therapeutic and coercive dimensions of these early institutional approaches (Loring & Hermann, 2017).

Wallin (1912) used the Binet-Simon intelligence test to assess 333 patients in the *New Jersey State Village for Epileptics* between 1910 and 1911, providing early empirical data and helping integrate psychological testing into neurological settings. Subsequently, Fox (1924) showed that academic underachievement among children with epilepsy excluded from public education was not fully explained by Binet-Simon IQ scores, underscoring both the value and limitations of early intelligence testing in epilepsy (Fox, 1924; Hermann, 2010).

As eugenics gained scientific legitimacy, epilepsy colonies increasingly incorporated intelligence testing. At that time, epilepsy was frequently conflated with cognitive disability and mental illness, and IQ scores became a central tool in labeling residents as permanently “deficient.” Eugenic field workers created detailed family pedigrees to trace perceived hereditary “taint,” regularly identifying individuals as socially unfit based on seizure history, low test scores, and assumptions of intellectual inferiority. Intelligence testing was interpreted without considering seizure activity, medication effects, or broader social disadvantages, and low IQ scores were presented as objective proof of intrinsic deficiency. Consequently, some residents of epilepsy colonies were subjected to forced sterilization including surgical castration under the belief that they posed a threat to the genetic health of future generations (Kelley, 2006).

Eugenic ideology represented what was regarded at the time as progressive medical and scientific thinking. William J. Mayo, founder of the Mayo Clinic, used a 1923 address published in the *Boston Medical and Surgical Journal* to endorse sterilization and immigration restriction as means of “public health improvement.” Mayo argued that poverty and illness reflected “hereditary inferiority” and advocated reducing the number of people cared for “at taxpayer expense” through eugenic measures (Lombardo, 2024).

Endorsements of eugenic policies by medical authorities supplied the scientific and moral justification for subsequent legal measures. The most consequential of these was *Buck v. Bell* (1927), a landmark U.S. Supreme Court case that upheld the constitutionality of Virginia’s eugenics-based sterilization law. The case involved Carrie Buck, institutionalized at the *Virginia Colony for the Epileptic and Feeble-minded* and labeled “feeble-minded,” a designation also applied to her mother and daughter. *Buck v. Bell* served as the test case for whether the state could involuntarily sterilize individuals deemed unfit to reproduce. In an 8–1 decision, Justice Oliver Wendell Holmes Jr. famously declared, “Three generations of imbeciles are enough.” This ruling legitimized compulsory sterilization under the guise of public health and the purported goal of genetic improvement.

Early IQ studies unfortunately suffered from ascertainment bias. Because intelligence test data were largely derived from residents of institutional epilepsy colonies, early reports frequently showed low median IQs around 70 (Hermann, 2019), reflecting not only epilepsy severity and underlying neurologic etiologies but also treatment effects such as bromides and the cognitive and social consequences of prolonged exclusion from formal education. In the 1940s, however, epileptologist William Lennox, despite his support of eugenics (Offen, 2003), and psychologist A. Louise Collins studied patients in private practice and demonstrated that average IQ scores were common, revealing bias in earlier institutional samples and challenging prevailing assumptions about cognition in epilepsy (Collins & Lennox, 1947; Collins, 1951). By then, however, misuse of IQ testing had already contributed to enduring stigma shaping clinical and societal attitudes.

Early neuropsychology in epilepsy heavily relied on IQ-based patient classification. Donald Hebb (1939a, 1939b) observed that patients undergoing extensive epilepsy surgeries, including unilateral and bilateral frontal or unilateral temporal resections, often showed stable, unimpaired, or even improved postoperative IQ scores. These findings led him to emphasize the assessment of specific cognitive function. Subsequent work by Meyer and Yates

(1955) and Scoville and Milner (1957) reinforced this perspective, laying the groundwork for the structure-functions approach in epilepsy neuropsychology.

Cognitive functioning in epilepsy exhibits substantial heterogeneity both between and within syndromes, rendering global IQ scores insufficient for accurately characterizing domain-specific impairments. In many cases, particularly for self-limited epilepsies, overall IQ is typically average, yet individuals may exhibit selective deficits in specific cognitive domains. Individuals with temporal lobe epilepsy, for example, often experience selective memory and language impairments that can be obscured by average or above-average overall IQ, without consistent verbal versus performance IQ differences corresponding to seizure onset laterality (Hermann et al., 1995; Loring et al., 2008). By integrating neuropsychology early into epilepsy surgery outcomes research, and aligning cognitive findings with EEG and neuroimaging data, domain specific impairments rather than global IQ scores emerged as essential for determining surgical candidacy and evaluating treatment outcomes (Chelune et al., 1998; Loring & Hermann, 2017; Loring, 2010). These findings underscore the importance of moving beyond IQ as a primary measure of cognitive ability and highlight the need for more domain-specific assessments to accurately identify clinically meaningful cognitive deficits associated with epilepsy (IC Code) (Norman et al., 2021).

Contemporary neuropsychology research in epilepsy increasingly emphasizes cognitive phenotypes as a framework for characterizing both seizure syndromes and variability within syndromes, moving beyond reliance on global IQ indices. Studies in adults with frontal lobe epilepsy demonstrate discrete cognitive phenotypes defined by differential involvement of language, attention, executive functioning, processing speed, and learning, highlighting substantial heterogeneity that is obscured by summary measures of overall ability (Arrotta et al., 2022). Parallel work examining subjective and objective cognition shows domain specific correspondence between patient reported difficulties and measured performance in areas such as attention, memory, and executive functioning, supporting the clinical utility of standardized multi domain assessment approaches (Hohmann et al., 2023). In pediatric epilepsy, comparisons between localization related and absence syndromes reveal syndrome specific cognitive vulnerabilities, including focal verbal memory weaknesses, despite broadly preserved intellectual functioning (Kernan et al., 2012). Similarly, studies comparing childhood epilepsy syndromes identify both shared and distinctive patterns of impairment, such as differential effects on visual attention versus spatial abilities, further reinforcing the value of domain based cognitive characterization (Cheng et al., 2017). Collectively, this literature reflects a shift toward cognitive phenotyping as a means of refining syndrome classification, capturing within syndrome variability, and identifying clinically meaningful cognitive risk profiles in epilepsy.

From fixed trait to dimensional constructs: Contemporary characterization

Binet recognized a fundamental limitation of intelligence tests, noting that “the scale does not permit the measure of intelligence, because intellectual qualities are not superposable, and therefore cannot be measured as linear surfaces are measured” (1909, as cited in Zenderland, 1998). Such tests capture performance on specific tasks rather than a fixed underlying intelligence. Sternberg (1999) extended this view by defining “successful intelligence” as the

ability to balance the needs to adapt to, shape and select environments in order to attain success, emphasizing the capacity to navigate real-world, context-dependent, and ambiguous problems and illustrating intelligence as a multifaceted, contextually shaped ability rather than a linear trait.

Modern approaches to cognitive assessment, exemplified by Process Overlap Theory (POT; Kovacs & Conway, 2016), conceptualize intelligence as an emergent outcome of interacting executive and domain general processes rather than a single causal factor. By treating “g” as a statistical product of overlapping processes instead of a unitary construct, POT provides a contemporary framework that better accounts for cultural and contextual influences on cognitive performance.

Cognitive abilities and traits show substantial correlation heritability and longitudinal stability (Breit et al., 2024; Polderman et al., 2015); however, this does not demonstrate that cognitive test scores provide a direct measure of intelligence. Cattell and Horn explicitly rejected Spearman’s notion of a unitary general intelligence factor “g,” arguing that it reflects a statistical artifact of correlations among heterogeneous cognitive tasks rather than a single underlying mental capacity (Cattell, 1963; Horn, 1991). The common practice of interpreting “g” as evidence of global intelligence, particularly in the absence of a universally accepted definition of intelligence, reflects the subjective assumptions and biases of the researchers choosing that specific descriptive label.

The Cattell–Horn–Carroll (CHC) model is a hierarchical, empirically derived framework that organizes cognitive abilities across multiple levels of specificity. At its broadest level, the model includes a general factor that may emerge statistically from correlations among cognitive tasks but is not assumed to represent a unitary or fixed biological construct. Beneath this level, the CHC model specifies multiple broad cognitive abilities that are relatively distinct yet interrelated, including fluid reasoning, crystallized knowledge, processing speed, visual and auditory processing, short term and working memory, long term storage and retrieval, and quantitative knowledge. Each broad ability is further decomposed into narrower, more specific skills that correspond closely to performance on individual neuropsychological and psychometric tests. The CHC framework provides a structured approach for understanding cognitive strengths and weaknesses across domains rather than reducing performance to a single global score. Despite its prominence in cognitive and educational assessment, and its frequent use in test development and research, neuropsychological test interpretation in routine clinical practice rarely adopts the CHC framework as an organizing model (Hermann et al., 2023; Jewsbury et al., 2017).

Contemporary implications and ongoing inequities

Despite advances in theory and measurement, the legacy of intelligence testing, including its constructs, terminology, and interpretive frameworks, continues to shape neuropsychological practice and policy, sometimes reinforcing inequities. Within large-scale datasets such as the Adolescent Brain Cognitive Development study, broad cognitive composites derived from factor-analytic or predictive modeling approaches are described as indexing *General Cognitive Ability* (GCA) (Sripada et al., 2021). GCA does not denote a single, unitary ability, but rather an empirically derived dimension capturing shared variance across diverse cognitive tasks, reflecting the well-documented tendency for performance across cognitive domains to covary. Research with

this approach indicates that these general factors are associated with a range of educational, occupational, and health-related outcomes and are robustly linked to socioeconomic conditions, while also reflecting highly distributed, small-effect contributions across neural systems rather than localized or modular substrates.

Consistent with this empirical approach, terms such as *Total Cognition Composite*, introduced with the NIH Toolbox (Heaton et al., 2014), provide descriptive, data-driven summaries of broad cognitive performance by averaging standardized scores across neuropsychological batteries to capture overall efficiency relative to normative expectations. Characterizing group performance averages offers a practical and transparent summary across multiple measures without implying a fixed or essentialized construct of intelligence. Such composites, widely used in clinical trials to track cognitive status, disease progression, and treatment effects in conditions including HIV (Carey et al., 2004), multiple sclerosis (Erlanger et al., 2014), Alzheimer's disease (Donohue et al., 2014), traumatic brain injury (Schneider et al., 2024), and epilepsy (Meador et al., 2025), represent efficient tools for summarizing and monitoring cognitive functioning over time rather than as direct measures of intelligence or formal models of discrete cognitive abilities.

Clarifying terminology is necessary but insufficient for meaningful reform. The continued use of intelligence based language reflects entrenched practices in how cognitive performance is measured, normed, interpreted, and applied across educational, medical, and legal contexts. Promoting equity in neuropsychological evaluation requires pairing linguistic revision with procedural modernization, including diversified normative samples, socioculturally informed interpretation, standardized language access practices, and reduced reliance on rigid cutoff scores that obscure individual variability. Aligning conceptual precision with systemic reform strengthens both scientific validity and ethical integrity.

Equitable assessment also depends on structural change, including multi-method, context sensitive evaluation, elimination of fixed IQ thresholds in legal settings, and training in the historical and sociocultural foundations of testing. *Larry P. v. Riles* (1979) exposed racial bias in school placement; *Atkins v. Virginia* (2002) and *Hall v. Florida* (2014) demonstrated the harms of inflexible IQ cutoffs in capital cases; and the NFL Players' Concussion Injury Litigation (2015) showed that race-based baseline assumptions produced systematically lower compensation for Black players. Together, these cases underscore that meaningful reform in neuropsychology requires coordinated procedural and structural change to ensure fairness and transparency. Despite well-documented methodological and sociocultural limitations, some secondary analyses of large datasets continue to misinterpret summary cognitive scores as indices of intelligence to rank racial and ethnic groups (Human Varieties, 2023). Additional examples of the continued misuse of IQ scores are provided in the Supplementary File.

Beyond terminology: Toward systemic reform and equity

Although traditionally referred to as "intelligence scales," instruments such as the Wechsler Scales are more accurately understood as assessing a range of distinct cognitive processes that differ in their relevance to everyday functioning. Continued use of terms like "IQ" and "intelligence" in research, or the inclusion of "Intelligence" as a section header in clinical reports, mischaracterizes what these measures actually capture and risks reinforcing reductive and stigmatizing views of cognition.

Reporting "intelligence levels" is particularly problematic given the absence of a universally accepted definition of intelligence and the documented potential for such labels to carry pejorative implications. Adopting terminology that specifies discrete cognitive functions rather than invoking an assumed global construct promotes more accurate interpretation and aligns with principles of nonmaleficence by minimizing harm and distancing clinical practice from the cultural and historical baggage associated with outdated conceptions of intelligence.

Psychological science has repeatedly revised terminology to reflect current conceptual accuracy, as seen in the transition from "mental retardation" to "intellectual disability" in both the DSM and ICD frameworks (Harris, 2013). Similarly, research by Fletcher et al. (1994) and Siegel (1989) demonstrated how empirical studies of cognitive and developmental functioning prompted revisions to learning disorder and childhood psychopathology classifications. These examples illustrate the shift from outdated constructs to evidence-based, data-driven diagnostic frameworks.

Epilepsy provides a clear precedent for reevaluating terminology in neuropsychology. In South Korea, the replacement of the term for "epilepsy," which carried pejorative connotations, with a neutral, biologically grounded alternative (i.e., *cerebroelectric disorder*) led to measurable improvements in public awareness, reduced stigma, and more patient-centered framing (Koh et al., 2024; Shorvon, 2024). This demonstrates how diagnostic language shapes social meaning and clinical expectations.

The continued use of "intelligence" in neuropsychology mirrors this problem. Like the historical term for epilepsy, intelligence is an overutilized construct increasingly at odds with contemporary scientific models and, as evidence shows, directly linked to stigma and prejudice. Persisting in its use reinforces reductive interpretations and reflects an unwarranted reluctance to adopt modern conceptual frameworks. *Epilepsia*, the official journal of the *International League Against Epilepsy*, now discourages the term "intelligence" in its research reports, highlighting the field's growing awareness of the potentially stigmatizing effects of specific language.

Abandoning historically burdened labels, as seen with the epilepsy name change in South Korea, can shift public and patient perceptions, strengthen clinical care, and reduce misinterpretation. Revising "intelligence" and "IQ" aligns with this scientific imperative. Over 35 years ago, Muriel Lezak declared IQ "dead" in her 1988 INS presidential address, calling for a move beyond reductive measures. Advancing neuropsychology requires both multidimensional cognitive models and precise language to uphold scientific accuracy and mitigate stigmatizing implications.

Language change alone is insufficient; meaningful reform requires systemic action. Transitioning away from the term intelligence will be more challenging in pediatrics, where it is embedded in formal diagnostic labels such as intellectual disability and in educational systems that use it for student placement into gifted programs. Even in these contexts, framing performance in terms of abilities rather than intelligence remains appropriate and preserves diagnostic specificity. Importantly, there are growing efforts to integrate historical and cultural literacy into training, such as those advanced through the Minnesota Conference (Stringer et al., 2025), with relevance to the interpretation of cognitive performance across all assessment approaches. Updating terminology reflects the natural evolution of scientific language toward greater precision and contextual validity. As neuropsychology advances, its vocabulary should reflect, rather than constrain, our understanding of cognitive diversity.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1355617726101842>.

The Supplemental File was generated with the assistance of ChatGPT (OpenAI, GPT-5 mini, accessed January 2026). This was in response to reviewer requests to provide contemporary examples of harms resulting from the inappropriate use of IQ metrics. ChatGPT was employed to identify ten recent (within the past 10 years) instances reported in major U.S. national and regional news outlets describing misuse of IQ. Specific inclusion criteria were applied (see Supplemental File), and multiple iterations were conducted to prevent duplication, overlapping individuals, proceedings, or enforcement actions, as well as to verify the accuracy of press report links. Both individual-level cases and system-level documented abuses were considered provided they were analytically distinct.

Notes

This article is based on the INS Presidential Address “Views of Brain Functions Through an Historical Lens: Anatomy, Physiology, and Prejudice” delivered by David Loring at the Global Neuropsychology Congress (July 2024) in Porto, Portugal. Bruce Hermann was an advisor for the Presidential Address and is an active co-author of the present manuscript.

Funding statement. None.

Competing interests. None.

References

- Allen, G. E. (1986). The eugenics record office at Cold Spring Harbor, 1910–1940: An essay in institutional history. *Osiris*, 2(1), 225–264.
- Arrotta, K., Reyes, A., Kaestner, E., McDonald, C. R., Hermann, B. P., Barr, W. B., Sarmey, N., Sundar, S., Kondylis, E., Najm, I., Bingaman, W., & Busch, R. M. (2022). Cognitive phenotypes in frontal lobe epilepsy. *Epilepsia*, 63(7), 1671–1681.
- Atkins v. Virginia. (2002). 536 U.S. 304.
- Binet, A., & Simon, T. (1908). Le développement de l'intelligence chez les enfants. *L'Année Psychologique*, 14, 1–94.
- Binet, A., & Simon, T. (1916). *The development of intelligence in children* (E. S. Kite, Trans.). Williams & Wilkins.
- Breit, M., Scherrer, V., Tucker-Drob, E. M., & Preckel, F. (2024). The stability of cognitive abilities: A meta-analytic review of longitudinal studies. *Psychological Bulletin*, 150(4), 399–439.
- Brigham, C. C. (1923). *A study of American intelligence*. Princeton University Press.
- Brigham, C. C. (1930). Intelligence tests of immigrant groups. *The Psychological Review*, 37(2), 158–165.
- Buck v. Bell. (1927). 274 U.S. 200.
- Carey, C. L., Woods, S. P., Gonzalez, R., Conover, E., Marcotte, T. D., Grant, I., & Heaton, R. K. (2004). Predictive validity of global deficit scores in detecting neuropsychological impairment in HIV infection. *Journal of Clinical & Experimental Neuropsychology*, 26(3), 307–319.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1–22.
- Chelune, G. J., Naugle, R. I., Hermann, B. P., Barr, W. B., Trenerry, M. R., Loring, D. W., Perrine, K., Strauss, E., & Westerveld, M. (1998). Does presurgical IQ predict seizure outcome after temporal lobectomy? Evidence from the Bozeman Epilepsy Consortium. *Epilepsia*, 39(3), 314–318.
- Cheng, D., Yan, X., Gao, Z., Xu, K., Zhou, X., & Chen, Q. (2017). Common and distinctive patterns of cognitive dysfunction in children with benign epilepsy syndromes. *Pediatric Neurology*, 72, 36–41.e1.
- Collins, A. L. (1951). Epileptic intelligence. *Journal of Consulting Psychology*, 15(5), 392–399.
- Collins, A. L., & Lennox, W. G. (1947). The intelligence of 300 private epileptic patients. *Research Publications: Association for Research in Nervous and Mental Disease*, 26, 586–603.
- DeSalle, R. (2021). The eugenics movement in retrospect. *Natural History Magazine*. Retrieved from <https://www.naturalhistorymag.com/features/093896/the-eugenics-movement-in-retrospect>
- Donohue, M. C., Sperling, R. A., Salmon, D. P., Rentz, D. M., Raman, R., Thomas, R. G., Weiner, M., & Aisen, P. S. (2014). The preclinical Alzheimer cognitive composite: Measuring amyloid-related decline. *JAMA Neurology*, 71(8), 961–970.
- Erlanger, D. M., Kaushik, T., Caruso, L. S., Benedict, R. H., Foley, F. W., Wilken, J., Cadavid, D., & Deluca, J. (2014). Reliability of a cognitive endpoint for use in a multiple sclerosis pharmaceutical trial. *Journal of the Neurological Sciences*, 340(1–2), 123–129.
- Fisher, R. A. (1930). *The genetical theory of natural selection*. Clarendon Press.
- Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (1994). Learning disabilities: From identification to intervention. *Journal of Clinical & Experimental Neuropsychology*, 16(6), 901–911.
- Fox, J. T. (1924). The response of epileptic children to mental and educational tests. *British Journal of Medical Psychology*, 4(3), 235–248.
- Galton, F. (1883). *Hereditary genius: An inquiry into its laws and consequences* (2nd ed.). Macmillan.
- Goddard, H. H. (1910). Four hundred feeble-minded children classified by the Binet method. *Pedagogical Seminary*, 17, 387–397.
- Goddard, H. H. (1913). *Sterilization and segregation*. Department of Child Helping of the Russell Sage Foundation.
- Grant, M. (1916). *The passing of the great race; or, The racial basis of European history*. Charles Scribner's Sons.
- Hall v. Florida. (2014). 572 U.S. 701.
- Harris, J. C. (2013). New terminology for mental retardation in DSM-5 and ICD-11. *Current Opinion in Psychiatry*, 26(3), 260–262.
- Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., Beaumont, J., Casaletto, K. B., Conway, K., Slotkin, J., & Gershon, R. (2014). Reliability and validity of composite scores from the NIH toolbox cognition battery in adults. *Journal of the International Neuropsychological Society*, 20(6), 588–598.
- Hebb, D. O. (1939a). Intelligence in man after large removal of cerebral tissue: Report of four left frontal lobe cases. *Journal of General Psychology*, 21, 73–87.
- Hebb, D. O. (1939b). Intelligence in man after large removals of cerebral tissue: Defects following right temporal lobectomy. *Journal of General Psychology*, 21(2), 437–446.
- Hermann, B. (2010). 100 years of Epilepsia: Landmark papers and their influence in neuropsychology and neuropsychiatry. *Epilepsia*, 51(7), 1107–1119. <https://doi.org/10.1111/j.1528-1167.2009.02486.x>
- Hermann, B. (2019). Intelligence and epilepsy: The early era. *Epilepsy & Behavior*, 101(Pt A), 106597. <https://doi.org/10.1016/j.yebeh.2019.106597>
- Hermann, B. P., Gold, J., Pusakulich, R., Wyler, A. R., Randolph, C., Rankin, G., & Hoy, W. (1995). Wechsler adult intelligence scale-revised in the evaluation of anterior temporal lobectomy candidates. *Epilepsia*, 36(5), 480–487.
- Hermann, B. P., Loring, D. W., Bowden, S. C., & Sarkis, R. (2023). Epilepsy and neuropsychology. In G. G. Brown, T. Z. King, K. Y. Haaland, & B. Crosson (Eds.), *APA handbook of neuropsychology*, American Psychological Association (Vol. 1, pp. 311–334).
- Hohmann, L., Berger, J., Kastell, S. U., & Holtkamp, M. (2023). Domain-specific relationships of subjective and objective cognition in epilepsy. *Epilepsia*, 64(7), 1887–1899.
- Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In K. S. McGrew, J. K. Werder, & R. W. Woodcock (Eds.), *Woodcock-Johnson technical manual* (pp. 197–232). Riverside Publishing.
- Jewsbury, P. A., Bowden, S. C., & Duff, K. (2017). The Cattell–Horn–Carroll model of cognition for clinical assessment. *Journal of Psychoeducational Assessment*, 35(6), 547–567. <https://doi.org/10.1177/0734282916651360>
- Kelley, J. (2006). The master race. *Boston Magazine*. <https://www.bostonmagazine.com/uncategorized/2006/05/15/the-master-race/>
- Kernan, C. L., Asarnow, R., Siddarth, P., Gurbani, S., Lanphier, E. K., Sankar, R., & Caplan, R. (2012). Neurocognitive profiles in children with epilepsy. *Epilepsia*, 53(12), 2156–2163.
- Kevles, D. J. (1985). *In the name of eugenics: Genetics and the uses of human heredity*. Harvard University Press.

- Koh, S., Pyo, S. I., Jang, S., Kim, G., Won Seo, S., Huh, K., Kwon, S. S., & Young Choi, J. (2024). Shifting perspectives: From "Epilepsy" to "Cerebroelectric disorder." *Epilepsy & Behavior*, 157, 109842.
- Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, 27(3), 151–177.
- Larry P. v. Riles. (1979). 495 F. Supp. 926 (N.D. Cal).
- Laughlin, H. H. (1922). *Eugenical sterilization in the United States*. Eugenics Record Office.
- Lezak, M. D. (1988). IQ: R.I.P. *Journal of Clinical & Experimental Neuropsychology*, 10(3), 351–361.
- Lombardo, P. A. (2008). *Three generations, no imbeciles: Eugenics, the supreme court, and Buck v. Bell*. Johns Hopkins University Press.
- Lombardo, P. A. (2024). "Ridding the race of his defective blood" – Eugenics in the journal, 1906–1948. *New England Journal of Medicine*, 390(10), 869–873.
- Loring, D. W. (2010). History of neuropsychology through epilepsy eyes. *Archives of Clinical Neuropsychology*, 25(4), 259–273.
- Loring, D. W., & Hermann, B. P. (2017). History of epilepsy neuropsychology. In W. Barr, & L. Bieliesksas (Eds.), *The Oxford handbook of neuropsychology* (pp. 813–840). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199765683.013.18>
- Loring, D. W., Strauss, E., Hermann, B. P., Barr, W. B., Perrine, K., Trenerry, M. R., Chelune, G., Westerveld, M., Lee, G. P., Meador, K. J., & Bowden, S. C. (2008). Differential neuropsychological test sensitivity to left temporal lobe epilepsy. *Journal of the International Neuropsychological Society*, 14(3), 394–400.
- Meador, K. J., Seliger, J., Le, S., Li, Y., Razavi, B., Falco-Walter, J., King, A., Graham, E., Cunningham, E., Leeman-Markowski, B., Boyd, A., Loring, D. W., & Gerard, E. (2025). Cognitive and behavioral effects of perampam 4 mg daily dose. *Epilepsy & Behavior*, 164, 110268.
- Meyer, V., & Yates, A. J. (1955). Intellectual changes following temporal lobectomy for psychomotor epilepsy: Preliminary communication. *Journal of Neurology, Neurosurgery & Psychiatry*, 18, 44–52.
- NFL Players' Concussion Injury Litigation. (2015). No. 2:12-md-02323-AB. 307 F.R.D. 351 (E.D. Pa).
- Norman, M., Wilson, S. J., Baxendale, S., Barr, W., Block, C., Busch, R. M., Fernandez, A., Hessen, E., Loring, D. W., McDonald, C. R., & Hermann, B. P. (2021). Addressing neuropsychological diagnostics in adults with epilepsy: Introducing the international classification of cognitive disorders in epilepsy: The IC CODE initiative. *Epilepsia Open*, 6(2), 266–275.
- Offen, M. L. (2003). Dealing with "defectives: " Foster Kennedy and William Lennox on eugenics. *Neurology*, 61(5), 668–673.
- Osborn, H. F. (1921). The second international congress of eugenics. *Science*, 54(1397), 311–312.
- Paul, D. B., & Moore, J. (2012). The race question and the eugenics movement. In A. Bashford, & P. Levine (Eds.), *The Oxford handbook of the history of eugenics* (pp. 83–104). Oxford University Press.
- Pearson, K. (1907). *On further methods of determining correlation*. Cambridge University Press.
- Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47(7), 702–709.
- Richardson, N. (2011). *Howard Andrew Knox: Pioneer of psychological testing at Ellis Island*. Columbia University Press.
- Schneider, A. L. C., Pike, J. R., Elser, H., Coresh, J., Mosley, T. H., Diaz-Arrastia, R., & Gottesman, R. F. (2024). Traumatic brain injury and cognitive change over 30 years among community-dwelling older adults. *Alzheimer's & Dementia*, 20(9), 6232–6242.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, 20, 11–21.
- Shorvon, S. (2024). Should we use the word "epilepsy?" *Epilepsy & Behavior*, 157, 109865.
- Siegel, L. S. (1989). The development of reading disabilities. *Journal of Abnormal Child Psychology*, 17(3), 325–338.
- Sofair, A. N., & Kaldjian, L. C. (2000). Eugenic sterilization and a qualified Nazi analogy: The United States and Germany, 1930–1945. *Annals of Internal Medicine*, 132(4), 312–319.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. Macmillan.
- Sripada, C., Angstadt, M., Taxali, A., Clark, D. A., Greathouse, T., Rutherford, S., Dickens, J. R., Shedden, K., Gard, A. M., Hyde, L. W., Weigard, A., & Heitzeg, M. (2021). Brain-wide functional connectivity patterns support general cognitive ability and mediate effects of socioeconomic status in youth. *Translational Psychiatry*, 11(1), 571.
- Sternberg R. J. (1999). Successful intelligence: Finding a balance. *Trends in Cognitive Sciences*, 3(11), 436–442. [https://doi.org/10.1016/s1364-6613\(99\)01391-1](https://doi.org/10.1016/s1364-6613(99)01391-1)
- Stringer, A. Y., Fuchs, K., Bordes Edgar, V., Bristow, T., Correia, S., Penna, S., Reyes, A., Whiteside, D. M., Block, C., Bodin, D., Butt, S., Calamia, M., Didehban, N., Dodzik, P., Dotson, V. M., Fernandes, M., Freece, K., Fuller, R., Getz, G., ... Schmitter-Edgecombe, M. (2025). The Minnesota Conference proposed guidelines for education and training in clinical neuropsychology. *The Clinical Neuropsychologist*, 1–42 (Advanced online publication).
- Terman, L. M. (1925). *Genetic studies of genius: Vol. 1. Mental and physical traits of a thousand gifted children*. Stanford University Press.
- Human Varieties. (2023, May 27). IQ scores by ethnic group in a nationally representative sample of 10-year-old American children. <https://humanvarieties.org/2023/05/27/iq-scores-by-ethnic-group-in-a-nationally-representative-sample-of-10-year-old-american-children/>
- Wallin, J. E. W. (1912). Eight months of psycho-clinical research at the New Jersey State Village for epileptics, with some results from the Binet-Simon testing. *Epilepsia*, A3, 366–380.
- West Virginia Legislature. (2013). House Bill No. 2463: An Act to repeal sections 27-16-1 through 27-16-5 relating to sterilization of mentally incompetent persons (2013 Regular Session).
- Yerkes, R. M. (1921). Psychological examining in the United States Army. *Memoirs of the National Academy of Sciences*, 15, 1–890.
- Zenderland, L. (1998). *Measuring minds: Henry Herbert Goddard and the origins of American Intelligence Testing*. Cambridge University Press.