Uses and Misuses of the Test of Premorbid Functioning

- Lawrence G. Weiss, PhD

- Pacific Northwest Neuropsychological Society
- September, 2023

Lawrence G. Weiss, Ph.D.

Education

Doctorate, Industrial and Organizational Psychology
Texas A&M University

Master of Arts Clinical Psychology
Trinity University, San Antonio

Bachelor of Arts Psychology
State University of NY at Binghamton

Employment

1990-2018

Vice President, Global Research & Development (Retired)
Pearson Clinical Assessment

Awards

2012, 2013, 2014, 2015

Nominated by the American Psychological Association, Division 15, for the Senior Scientist Award

Lawrence G. Weiss, Ph.D.

Books

2019 WISC-V Clinical Use and Interpretation. Second Edition: San Diego: Academic Press
2016 WISC-V Assessment and Interpretation: Scientist-Practitioner Perspectives: San Diego: Academic Press
2013 WAIS-IV/WMS-IV/ACS: Advanced clinical interpretation. San Diego: Academic Press
2010 Bayley–III clinical use and Interpretation. San Diego: Academic Press
2010 WAIS–IV clinical use and Interpretation: Scientist–practitioner perspectives. San Diego: Academic Press
2008 WISC–IV clinical assessment & Intervention 2e (second edition). San Diego: Academic Press
2006 WISC-IV Advanced Clinical Interpretation. San Diego: Academic Press
2005 WISC–IV clinical use and interpretation: Scientist–Practitioner Perspective. San Diego: Academic Press
2003 Culture and Children's Intelligence; A Cross-Cultural Analyses of the WISC-III. San Diego: Academic Press

Publications & Presentations

30 peer-reviewed research articles
50 book chapters
40 international presentations
6 invited papers and monographs
30 national presentations
20 state and local presentations
20 technical reports

Proper Use of the Test of Premorbid Functioning

- Step One
  - Enter obtained TOPF Reading Standard Score in Advanced Clinical Solutions software.
  - Compare the Actual and Predicted TOPF Reading Standard Scores
  - Use the higher score as input for the next step.
  - Why?
- Step Two
  - The ACS software uses the TOPF Reading Standard Score found in step one to estimate a baseline score for each broad cognitive ability.
  - Compare the patient's obtained cognitive ability score to their predicted baseline score.

Key Points

- The obtained TOPF score is not the baseline!
  - TOPF is a word reading tests
  - It produces a Word Reading Standard Score
  - The Reading Standard Score is used to predict a baseline score
  - The predicted baseline score is unique to each cognitive ability
    - It corrects for regression to the mean
    - Adjusts for lower correlations between word reading and cognitive abilities less related to reading (e.g., processing speed, episodic memory)
- In ACS, TOPF is never compared directly to the obtained ability score.

Misuses of TOPF

1. Establish a single point “baseline” using the TOPF Reading Score

2. Subtract all other test scores from TOPF Reading Standard Score

3. Shift middle of bell curve to match plaintiff’s TOPF Reading score

4. Declare deficits for all scores 15+ points below TOPF Reading Score

Flaws in Direct Subtraction with Wechsler Scales

- Direct Subtraction of the TOPF reading standard score from Wechsler test scores **will always overestimate the size of any differences** for high ability people because:
  - It assumes the tests are perfectly correlated
  - It does not account for regression to the mean

Misuse: Subtracting TOPF directly from all other scores

- Simple Reaction Time
- Complex Reaction Time
- Stroop – C-W
- Stroop- Interference
- Go/No Go
- Running Memory CPT (simple-V)
- Running Memory CPT (simple-A)
- CPT complex
- Trail Making Test A
- Trail Making Test B
- Switching
- WCST- Categories Completed
- WCST - Trials to Complete
- WCST - Perseverative Errors
- WCST - Non- Perseverative Errors
- Right/Left Orientation
- COWAT
- Proactive Interference
- Category Fluency
- Reys CFT IM
- Reys CFT DR
- Reys CFT Recognition
- Finger Tapping – Dominant
- Finger Tapping – Non-Dominant
- Grip Strength – Dominant
- Grip Strength – Non-Dominant

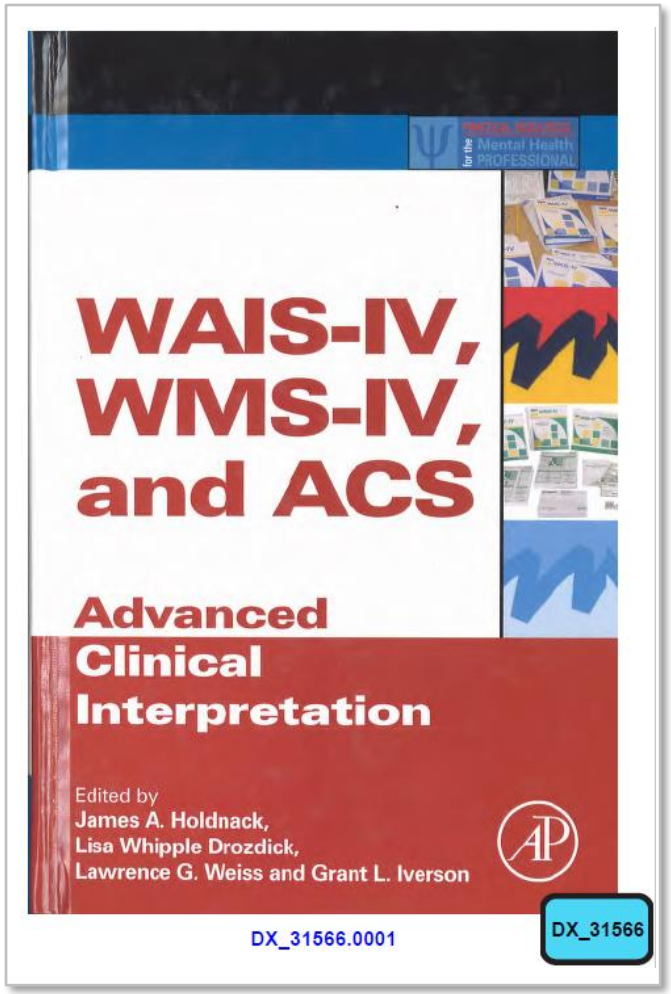
Flaws in Direct Subtraction with Other Test Scores

- DITTO
- Direct Subtraction of the TOPF reading standard score from scores on any cognitive test **will always overestimate the size of any differences** for high ability people because:
  - It assumes the tests are perfectly correlated
  - It does not account for regression to the mean
  - There is no base rate data to inform clinical judgements

TOPF Can't Predict Physiological Skills Accurately

- TOPF is less correlated with tests not associated with reading:
  - Reaction time
  - Finger Tapping
  - Fine Motor skills
  - Grip strength
- Using the TOPF Reading Score as a baseline for physiological abilities is like using your spelling score to predict your ability to ride a bicycle.
- It results in frequent discrepancies from "*baseline*" - falsely suggesting brain injury

Base Rates: What Are They?



exceeded the critical value or not. The base rate column provides the percentages of cases in the standardization sample having the same or larger differences between estimated premorbid and actual abilities.

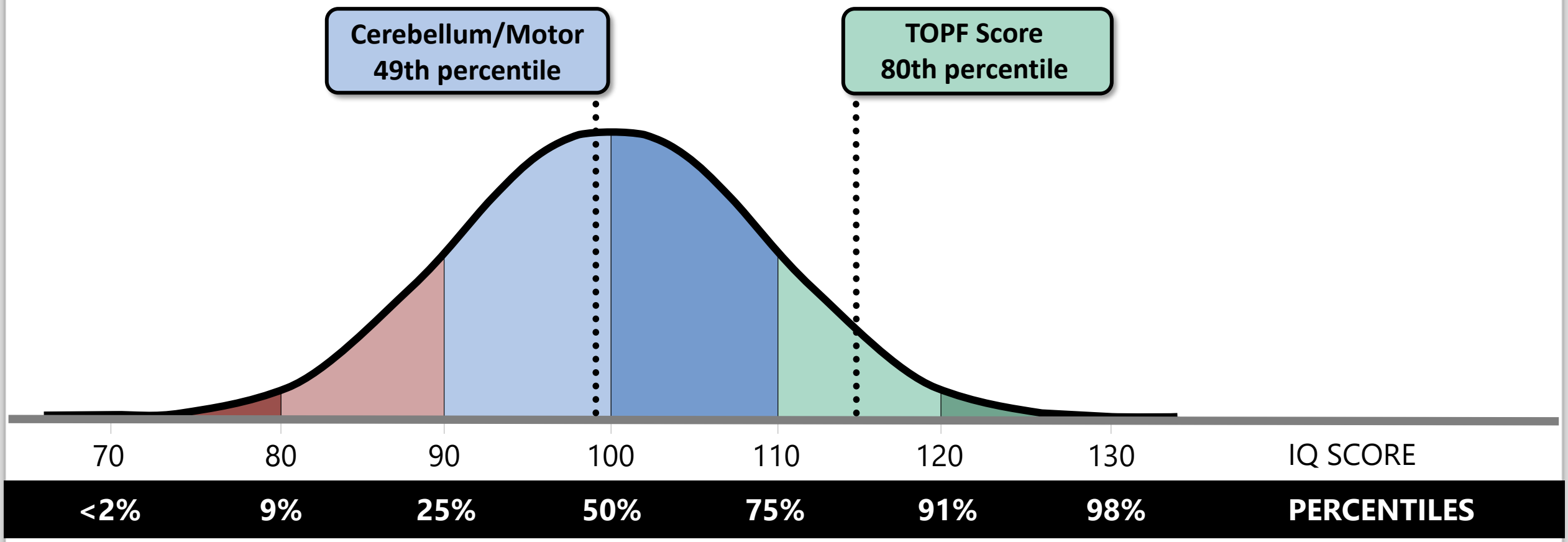
Rule of Thumb:

- Differences which occur in less than 10 or 15 percent of normal people are considered evidence of possible brain injury

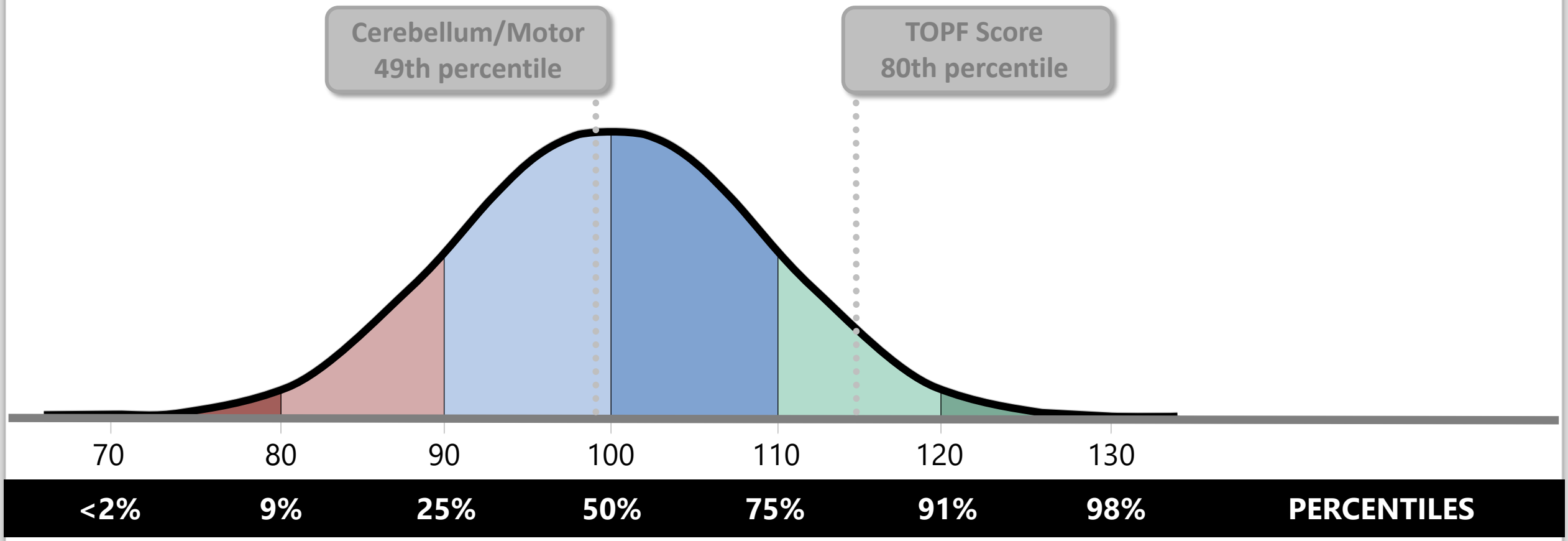
Base Rate Data is Not Always Available

- But, TOPF base rate data only exists for WAIS-IV and WMS-IV
- So, one NP invented his own method for estimating base rates with other tests

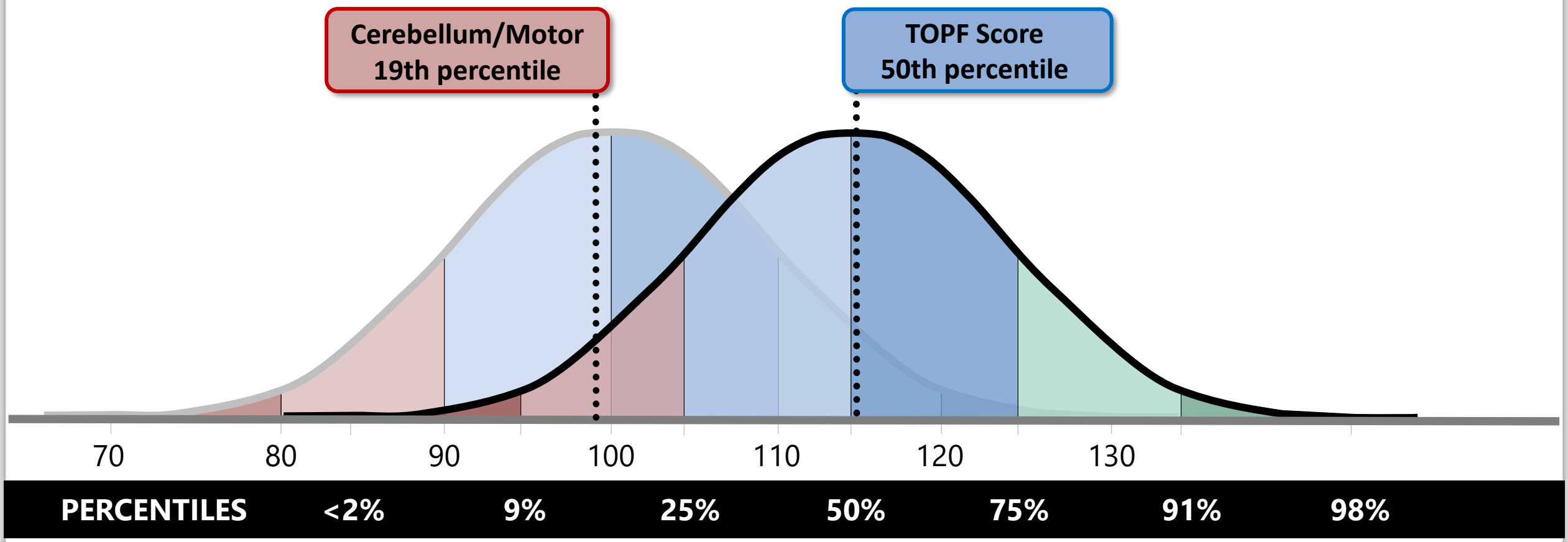
Curve Shifting Flawed Methodology



The Curve Shifts



And the percentiles shift with it



Base Rate Percentages are not Normal Curve Percentiles

- This method misuses normal curve percentile ranks as if they were base rates percentages.
- Percentile Ranks are for comparing a scores on one test to the population.
- They are not for comparing a person's score on test A to their own score on test B.
  - The proper interpretation of a score at the 15th percentile would be that 15 percent of the population obtained that score or lower.
- It does not mean that a particular individual has a 15% chance of obtaining that score given their estimated "baseline"
- That information can only come from base rates of difference scores between two tests.

Issues with Curve Shifting

- Percentile ranks are not an equal interval scale
- So, shifting the curve changes the difference in PRs between the baseline and cognitive ability scores.
- Example:
  - Baseline 130 (98th PR) and ability score of 115 (84th PR)
    - PR rank difference = 14 Percentile points
  - After shifting curve to center on baseline 130:
    - Baseline 130 is at 50th PR, and ability score of 115 is at the16th PR
    - PR rank difference = 34 Percentile points

"It is a serious interpretive error to apply the logic of the bell curve to the interpretation of isolated low scores across a battery of tests."

(Brooks, Iverson, & Holdnack, 2013, page 87).

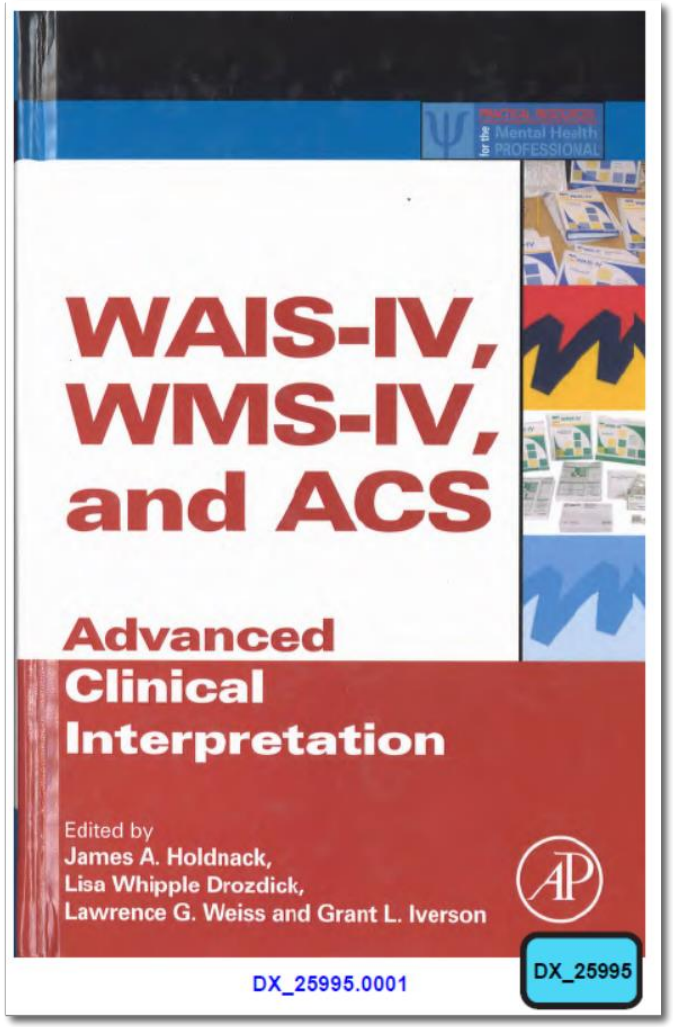
Basic Principle of Neuropsychology: Normal Human Variability

Case Study: An individual who can talk their way out of any situation but has difficulty with higher math like geometry.

Their visual-spatial ability is weak relative to their verbal ability. But, that may be a normal human weakness - not impairment.

- Few people are average on all types of cognitive abilities
- Large patterns of cognitive strengths and weaknesses are normal and natural
- This is why deficits from baseline are not necessarily considered evidence of brain injury

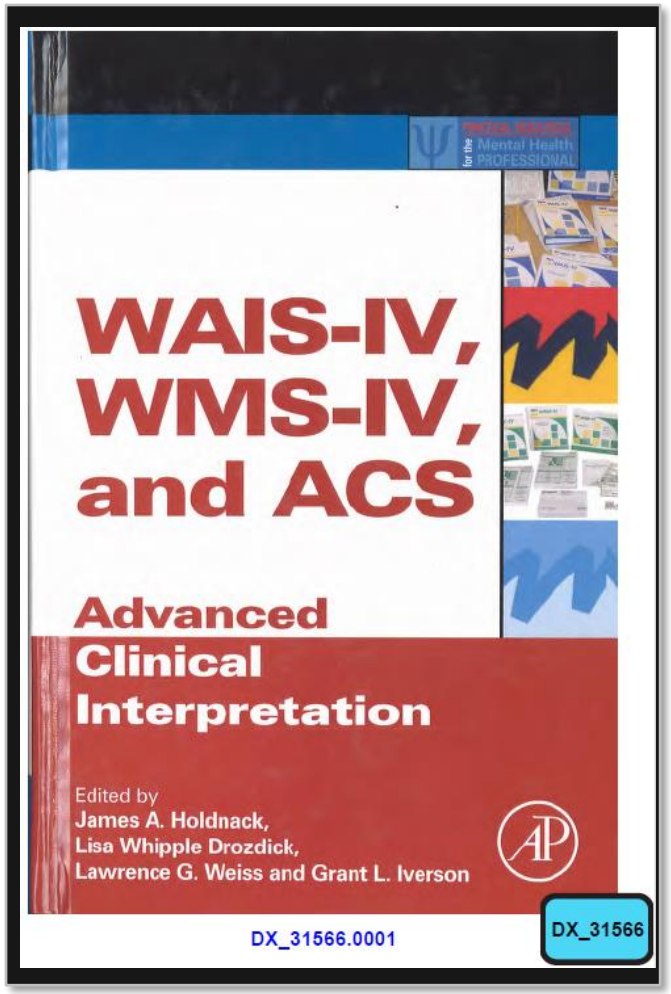
It Is Normal To have Abnormal Scores



variability.

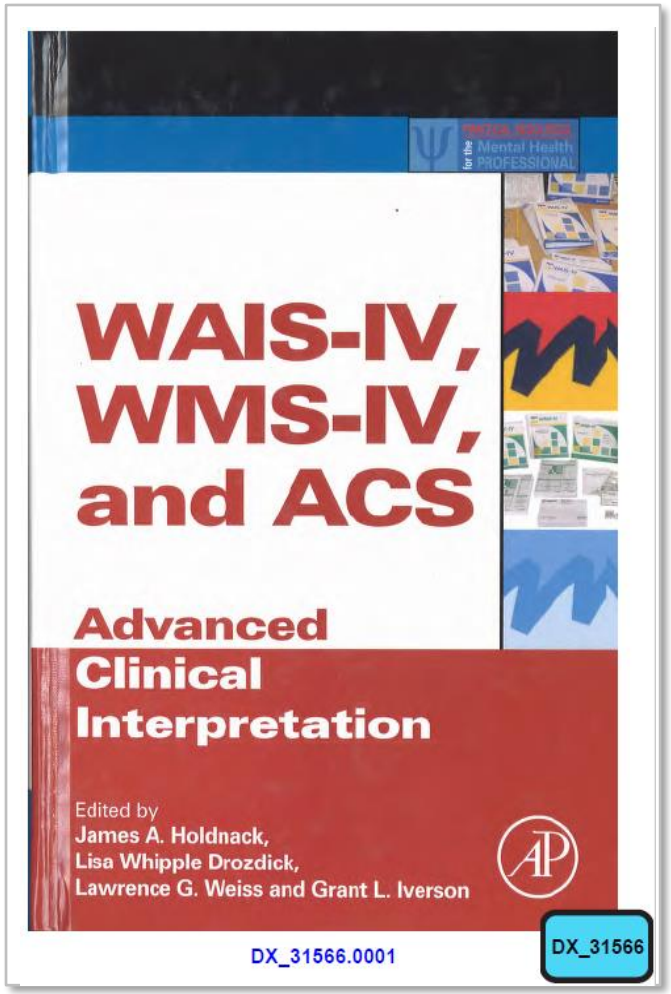
In addition to normal test score scatter, it is also normal for healthy children, adults, and older adults to obtain some low scores, particularly when multiple tests are administered. Two fundamental assumptions relating to test scores can lead to errors in clinical judgment. The first is a belief that if a person is "normal" then *all* of his or her scores should be "normal." The second is a belief that an "abnormal" test score is synonymous with atypicality, particularly brain injury. The clin-

Normal Patterns of Strengths & Weaknesses are Surprisingly Large



- The average difference between highest and lowest WAIS-IV / WMS-IV index scores is 28 IQ points in the standardization sample.
- A difference of 40 points would be required to achieve a base rate less than 10 percent –
  - Considered a rare or unusual finding indicative of brain impairment
- (Chapter 3, Table 3.8).

Multivariate Base Rates



administered and then interpreted in an integrated fashion. Over the past few years, however, practitioners have relied heavily on the immutable characteristics of the bell curve—and often unwittingly applied the logic of a single test score to a battery of tests. Put simply, when a client obtains a test score at the 5th percentile, the psychologist is tempted to believe that only 5% of healthy people would have that particular score (or a lower score). This is true when considering one test score from a normative sample that conforms to the bell curve. However, when considering a battery of tests the interpretive context changes dramatically. Having one or more scores at or below the 7th

common or uncommon it is to get certain patterns of low scores. It is a serious interpretive mistake to apply the logic of the bell curve to the interpretation of isolated low scores across a battery of tests. For

Multivariate Base Rates: How Many Dice Are Thrown?

- The more tests given the more chance of finding some low scores, even among healthy adults
- Technically, this is called “multi-variate base rates”
  - Just a fancy term for how odds change when many dice are thrown at once
- The key diagnostic questions are:
  - “How many scores were low?”
  - “What percent of normal people have that many low scores?”
- Subtracting TOPF from hundreds of test scores is like playing with loaded dice.

Multivariate Base Rates

Evidence-Based Neuropsychological Assessment Following Work-Related Injury

Grant L. Iverson
University of British Columbia

Brian L. Brooks
Alberta Children's Hospital
University of Calgary

James A. Holdnack
Pearson Assessment

Book Title: Neuropsychological Assessment of Workplace Injuries
Editors: Shane S. Bush, Ph.D. & Grant L. Iverson, Ph.D.
Publisher: Guilford Press

March 3, 2011

Introduction

Cognitive impairment following workplace injuries can be time-limited or permanent. Traumatic brain injuries, depression, anxiety disorders, and chronic pain can be associated with subjectively reported and/or objectively documented cognitive problems. The challenge for neuropsychologists who evaluate injured workers lies in accurately identifying problems with cognition, quantifying the deficits, estimating the impact on day-to-day functioning, and apportioning causation. The purpose of this chapter is to promote and encourage evidence-based neuropsychological assessment following work-related injuries.

Conceptualizing Evidence-Based Neuropsychological Practice

In 2005, the American Psychological Association (APA) approved a policy statement relating to evidence-based practice in psychology. Excerpts from this policy statement are reprinted below.

> Evidence-based practice in psychology (EBPP) is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences. This definition of EBPP closely parallels the definition of evidence-based practice adopted by the Institute of Medicine (2001, p. 147) as adapted from Sackett and colleagues (2000): "Evidence-based practice is the integration of best research evidence with clinical expertise and patient values." The purpose of EBPP is to promote effective psychological practice and enhance public health by applying empirically supported principles of

Figure 2. Base rates of low scores across batteries with different numbers of scores being interpreted: Cutoff < 1SD and ≤ 5th percentile.

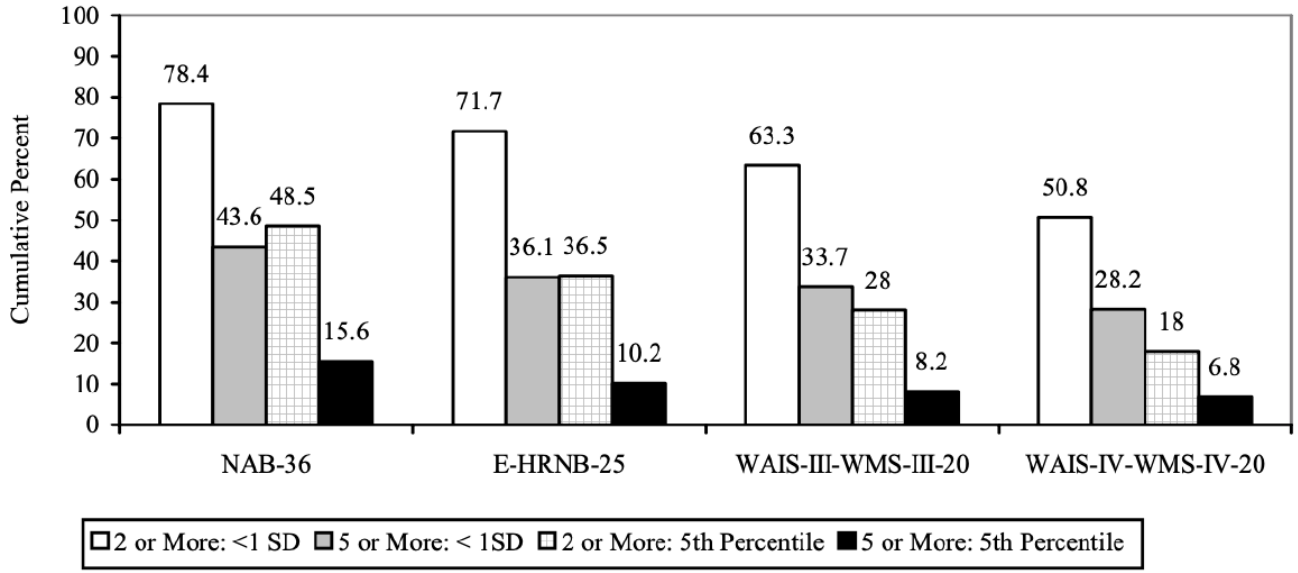
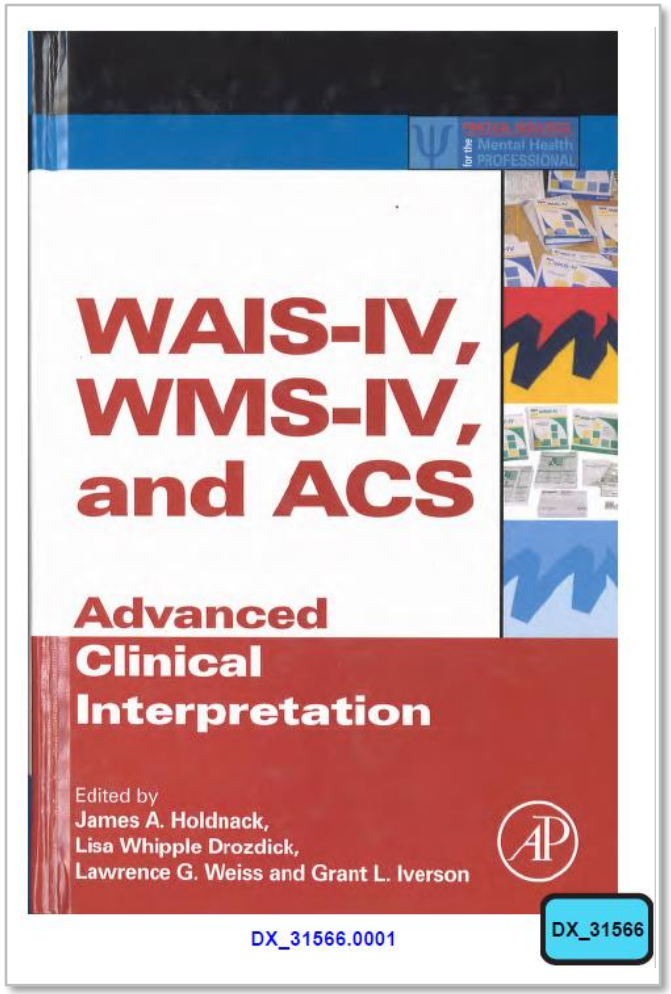


Figure Note: SD = Standard Deviation; NAB-36 = all 36 scores from the full Neuropsychological Assessment Battery; E-HRNB-25 = 25 scores from the Expanded Halstead-Reitan Neuropsychological Battery; WAIS-III-WMS-III-20 = all 20 primary subtest scores from the Wechsler Adult Intelligence Scale – III and Wechsler Memory Scale – III; and WAIS-IV-WMS-IV-20 = all 20 primary subtest scores from the Wechsler Adult Intelligence Scale – IV and Wechsler Memory Scale – IV. Bars represent percent of healthy adults from standardization samples who had (a) 2 or more or (b) 5 or more scores below 1SD (T<40; SS ≤ 7) or at or below 5th percentile (i.e., T=34 or SS=5).

High Ability People have an even Greater Range of Scores

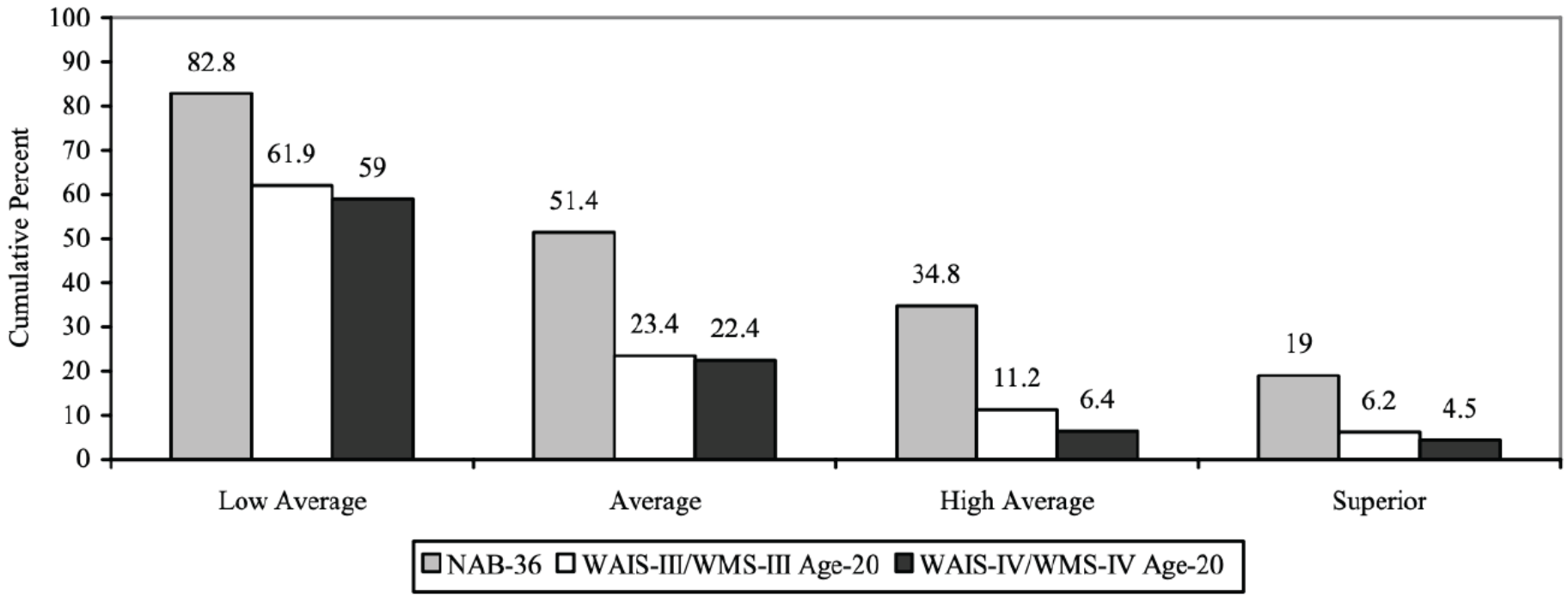


Variability by Highest Score

"Profile variability is significantly correlated with the person's highest obtained subtest and index score. Thus, gifted individuals show more performance variability than other samples."

High Ability People Have A Greater Range of Scores

Figure 8. Prevalence of low scores across different test batteries stratified by level of intelligence: 2 or more scores ≤ 5th percentile.



Note: NAB scores are stratified by RIST scores. WAIS/WMS scores, which are age adjusted, are stratified by either WTAR (WAIS-III/WMS-III) or ToPF (WAIS-IV/WMS-IV) scores. Numbers after the names of the tests in the legend indicate the number of subtest scores considered simultaneously.

Higher Ability S's Have Greater Variability Between Scores

- But, this finding is masked with by direct subtraction and curve shifting
- From Oakes, et. al., 2013, chapter 3, tables 26-28

Baseline (VCI)	PSI	Base Rate	Shifted Percentile
105	84	9%	9%
115	94	16%	9%
124	103	50%	9%

NFL Concussion Settlement

Switch to Adobe and show pdf of NFL Impairment Criteria

Direct Subtraction vs NFL Criteria

- TOPF Reading Standard Score = 125 (Above Average)
  - Four EF scores = 110 each (T=57)

- NFL Criteria for Above Average Reading Level:
  - no EF scores below T-40, thus no impairment

- Direct Subtraction method:
  - 1 SD difference from “baseline” for all EF scores, thus impaired.

Comparing NFL to Direct Subtraction Method

- No single point estimate of baseline
- No direct subtraction of cognitive scores from baseline
- No demographic adjustments of cognitive test scores
- No shifting of bell curve to center on a baseline score
- Relies on general population norms
- Accounts for normal human variability (e.g., 2 of 4 EF scores < T-40).
- Scores of 89 (t-43) or higher are never considered evidence of impairment
  - In the direct subtraction method, average or above average cognitive scores are considered impaired when 15 or more points below TOPF Reading score.

Direct Subtraction Methodology Is Unreliable

- **NOT been subject to peer review**
  - No known evidence for the validity of his procedure
- **NOT sufficiently established in the field**
  - No known textbooks recommend this method
- **NOT generally accepted in the field**
  - No one else does it this way
- **CONTRARY to standards of best practice**
  - Misuses TOPF and bell curve
  - Dosen't account normal human variability
    - or multivariate base rates

Bottom Line

- Direct subtraction method leads to overidentification of cognitive deficits
- Curve shifting method leads to overidentification of cognitive deficits

My Role in Forensic Cases

- Passionate about proper use of the instruments my team helped create
- Defend NP's when unfairly attacked for how they used a test
- Defend NP's when the tests itself is under attack
- Opine in cases where tests are being misused by the opposing NP
- Able to opine with strong credibility due to my role in developing these tests
- But, I do not test plaintiffs myself, nor offer diagnoses.
  - that's your role
  - my role is to have your back

One Last Story

- A very famous NP was attacked for using WASI in a civil court case
- Why? Because the WASI manual states;
  - *"In general, the WASI-II FSIQ should not be used for legal, judicial, or quasi-legal purposes (e.g., statutorily mandated diagnosis or determination of a disability)."*
- What happened next….?

Thank You!

- Contact Information:

- Email: LGweiss48@Yahoo.com
- Mobile: (210) 300 – 5665
- Time Zone:
  - Eastern, November thru May
  - Mountain, June thru October
- Website: https://lgweiss48.wixsite.com/dr-larry-weiss
- Linked In: https://www.linkedin.com/in/larry-weiss-350a569/

Discussion Questions

- How would you determine BI using TOPF and unlinked tests?
- How would you determine BI for a very high ability patient?
- Are there better ways of determining BI in these situations?
  - What new data are needed?