

# Explaining Psychometric Data to IEP Teams

Emily S. Rosen  
Linda S. Larrivee

---

---

---


---

---

---

---

---



## Introduction

- One responsibility of SLPs is to report assessment information in team meetings
- It is one thing to understand scores, another to explain their importance to other professionals
- SLPs often feel that parents cannot understand psychometric data

---

---

---


---

---

---

---

---



## Why Is This Information Important?

- Some team members and some parents have no idea of the meaning of different scores
- Many different scores are available from standardized tests
  - Some scores are more accurate in interpretation of students' abilities
  - Other scores can lead to misinterpretation and misuse

---

---

---

---

---

---

---

---

## Normal Curve

- Definition: An ideal distribution of scores in which the mean, median, and mode are identical
- Measures of central tendency
  - Mean: the statistical average of a set of scores
  - Median: the score for which 50% of the scores are higher and 50% of the scores are lower
  - Mode: the most frequently occurring score
- Variance
  - Describes the dispersion of a set of scores around the mean of the distribution
  - Standard deviation (SD) is the square root of the variance
    - SD forms the basis for understanding standard scores

---

---

---

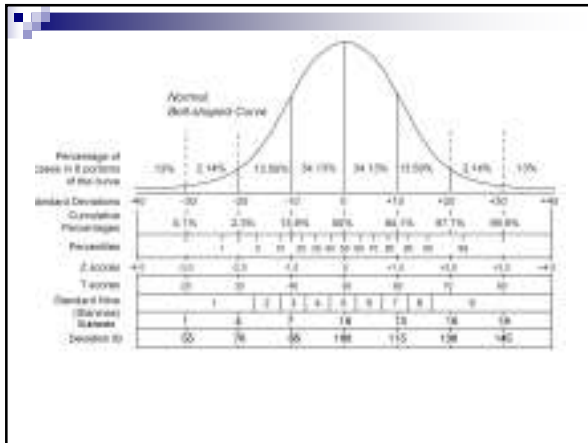
---

---

---

---

---




---

---

---

---

---

---

---

---

## Types of Scores: Raw Score

- Raw Score: represents the total number of stimulus items the student answered correctly.
- Are these scores meaningful?
  - Used to derive other types of scores (standard scores, percentile ranks, etc.)
  - Insignificant because there is no basis for comparison to other test-takers' performances.
  - *Cannot be used for comparison*
  - *Cannot be used to measure progress over time*

---

---

---

---

---

---

---

---

## Types of Scores: Standard Scores

- Standard scores: indicate how many standard deviations away the score is from the mean for a particular age group
  - Derived from raw scores
- Most commonly used scales
  - Deviation (Weschler) IQ (mean=100, SD=15)
  - t-score (mean=50, SD=10)
  - Normal curve equivalent (NCE) (mean=50, SD=20)
  - Subtest scores (mean=10, SD=3)
  - Other: z-score (mean=0, SD=1), stanines (separates the curve into 9 equal section, with 4 being the midpoint)

---

---

---

---

---

---

---

---

## Types of Scores: Standard Scores (continued)

- Are they meaningful?
  - Report a range of normal to which students' scores can be compared
  - Follow an equal interval scale, so they can be used to measure students' progress over time (at 3 year re-evaluation)
  - Can be compared across language modalities, subtests, and tests

---

---

---

---

---

---

---

---

## Types of Scores: Percentiles

- Percentile Rank (PR): a measure of relative standing in terms of percentage of scores occurring above or below that point
- PROBLEM: percentiles are not percentages
  - A student with a PR of 50 means that he scored better than 50% of the people taking the test: 50 is average
  - However, scoring 50% on a non-standardized test (i.e., classroom test) means that he answered 50% of the questions correctly: 50% at this test is failing

---

---

---

---

---

---

---

---

## Types of Scores: Percentiles (continued)

- Are they meaningful?
  - Report a range of normal to which students' scores can be compared
  - Does not follow an equal interval scale, but can measure students' progress over time (at 3 year re-evaluation)
  - Can be compared across language modalities, subtests, and tests
  - But are easily misinterpreted

---

---

---

---

---

---

---

---

## Confidence Intervals

- Range in which the "true" score is expected to fall
- Because no test is 100% reliable, the "true" score is expected to be within a range
- Standard error of measurement (SEM): Amount of error (in standard score units) to be considered in interpreting scores
- Used for classification or placement decisions
  - May be useful in "borderline" cases
  - Can allow room for clinical judgment

---

---

---

---

---

---

---

---

## Confidence Intervals (continued)

- Problems:
  - NO ONE seems to understand these scores
    - When to use them?
    - Why to use them?
  - Choice of intervals:
    - 68% (smallest range)
    - 80%
    - 90% (largest range)
  - Using these scores to qualify students can be a slippery slope

---

---

---

---

---

---

---

---

## Types of Scores: Age and Grade Equivalent Scores

- Age Equivalent (AE) Scores
  - represent the supposed age at which a student is functioning for the construct that was tested.
- Grade Equivalent Scores
  - represent the supposed grade at which a student is functioning for the construct that was tested.

---

---

---

---

---

---

---

---

## Age and Grade Equivalent Scores: Problems

- Do not take into consideration the range of normal performance
- Promote typological thinking: the average X year-old does not exist
- Promote untrue generalization
  - An 18 year-old who scores at the 3 year level is not actually functioning like a 3 year-old

---

---

---

---

---

---

---

---

## Age and Grade Equivalent Scores: Problems

- Derived through interpolation and extrapolation
  - scores represent a mean score of a group of children that were not actually tested.
- AE scores not derived for students receiving extremely high or low scores.
- Do not follow an equal interval scale: cannot be used to demonstrate progress over time.

---

---

---

---

---

---

---

---

## Limitations of AE Scores: Evidence

- Despite well-documented limitations, AE scores continue to be used
- Research study:
  - examined how AE scores described performance on the PPVT-III for a group of 19-year-olds whose scores fall within the range of normal, as determined by standard scores
  - aimed to provide empirical evidence for theoretical limitations associated with AE scores

---

---

---

---

---

---

---

---

## Method

- Participants:
  - n=50
  - 19;0 to 19;11
  - 32% male; 68% female
  - Fluent in English
  - 94% Native English Speakers
  - Normal hearing
  - 98% college students
  - 4% learning disability
- Procedure:
  - One testing session
  - 1 page questionnaire pertaining to demographic info
  - Administration of PPVT-III (Form A)
  - PPVT protocols were compiled and analyzed
  - Participants (n=7) whose scores fell outside range of normal (SS between 85-115) were excluded from analyses

---

---

---

---

---

---

---

---

## Results

- Mean chronological age (CA) for all participants = 19;6
- CAs ranged from 19;0 to 19;11
- 7 participants whose scores fell outside the range of normal were excluded
- Frequency distributions were plotted for raw scores, standard scores, and AE scores for the remaining participants (n=43)

---

---

---

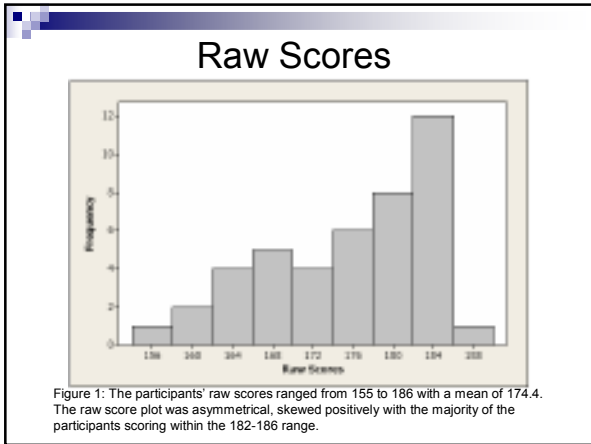
---

---

---

---

---




---

---

---

---

---

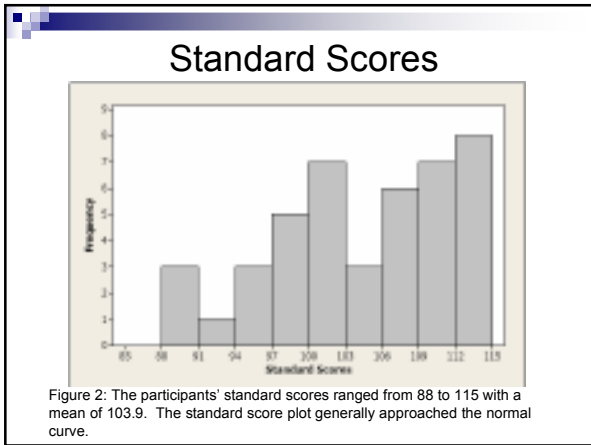
---

---

---

---

---




---

---

---

---

---

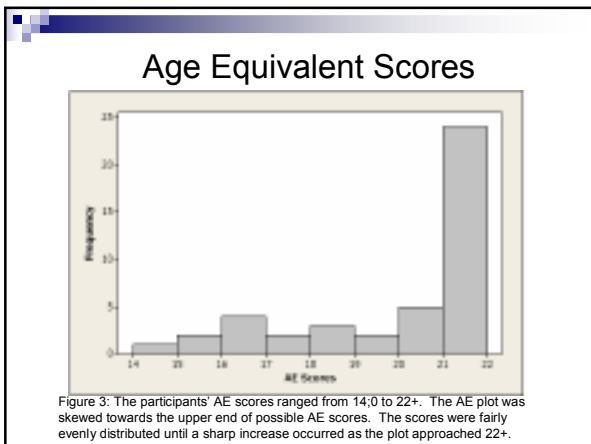
---

---

---

---

---




---

---

---

---

---

---

---

---

---

---

### Difference Between CAs and AE Scores

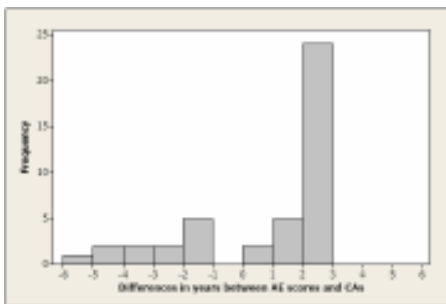


Figure 4: The differences between participants' CAs and AE scores ranged from - 5.58 to 3.00 years. Thirty percent earned AE scores below their CA. Seventy-two percent earned AE scores above their CAs. Seventy-four percent earned AE scores that were 2 years above or below their CAs.

---

---

---

---

---

---

---

---

### Discussion: Limitations of AE Scores

- Results support the concerns associated with AE scores proposed in the literature
- Range of AE scores was wide (14;0 to 22+)
- Example:
  - A 19-year-old who earns a standard score of 88 could earn an AE score of 14;0
- Finding demonstrated the potential for AE scores to be grossly misinterpreted
- AE scores are likely to promote typological thinking
  - Was the 19-year-old who earned an AE score of 14;0 performing like the "average" 14-year-old?

---

---

---

---

---

---

---

---

### Discussion: Limitations of AE Scores (cont.)

- AE scores provided limited information for examinees who earned extremely high scores
  - participants who earned standard scores between 104 and 115 (56%) all earned AE scores of 22+
- Limitations applied to most participants
  - 74% of participants earned AE scores that differed from their CAs by at least 2 years
- Method in which AE scores are derived is problematic –
  - AE scores for higher raw scores extrapolated for ages in much larger increments than for younger children earning lower raw scores
  - only two extrapolated AE scores are within 19-year-old range
    - 19;1 and 19;9.

---

---

---

---

---

---

---

---



## Implications for SLPs

- Evidence for some of the limitations associated with AE scores that have been proposed in the literature
  - AE scores can lead to gross misinterpretation of test performance
  - AE scores are likely to promote typological thinking
- AE scores are highly ineffective in reporting performance on standardized tests
- **SLPs should:**
  - Use more useful and appropriate scores that take into consideration the range of normal performance
  - Avoid using AE scores because they lead to misinterpretation and misuse.

---

---

---

---

---

---

---

---

## Explaining Scores to IEP Teams

- Always use a visual: A picture of the normal curve is helpful
- Avoid technical terms
- When technical terms must be used, use examples

---

---

---

---

---

---

---

---

## Explaining Raw Scores

- Should you report raw scores?
  - No need to report these scores to IEP team members
  - Provide no insight into how the student performed compared to his or her same-aged peers
  - What if your district requires you to report them?
    - If this is the case, there is no harm in reporting them
    - Be sure to downplay raw scores and place more emphasis on other types of scores that provide information about how the test-taker performed in relation to others.

---

---

---

---

---

---

---

---

## Explaining Standard Scores

- Standard scores should be the basis of presentation to the IEP team:
  - Report a range of normal to compare students' scores
  - Measure students' progress over time
  - Compare students' performance
    - across language modalities (strengths and weaknesses)
    - with performance on other IEP team member's testing,
      - school psychologists (cognition)
      - special ed teachers (reading)
      - etc.

---

---

---

---

---

---

---

---

## Explaining Standard Scores (continued)

- Always use a visual when explaining standard scores to the IEP team
- Be sure to present the bell curve and explain:
  - where average falls for a particular scale: this is considered to be within normal range.
  - scores yield mild, moderate, and severe ranges
  - what these terms mean in your district (ex: does mild in your district not qualify a student for services?)

---

---

---

---

---

---

---

---

## Mild? Moderate? Severe?

- Many districts have definitions of mild, moderate, and severe
- In addition, tests often have their own definitions which may or may not agree with the district

---

---

---

---

---

---

---

---

## Example of Guidelines for Describing Severity of Language Disorder

- From CELF-4 Examiner's Manual (p.112)

| Standard Scores | Classification           | Relation to Mean     |
|-----------------|--------------------------|----------------------|
| 115 and above   | above average            | + 1 SD and above     |
| 86 - 114        | average                  | within +/- 1 SD      |
| 78 - 85         | marginal/borderline/mild | within -1 to -1.5 SD |
| 71 - 77         | low range/moderate       | within -1.5 to -2 SD |
| 70 and below    | very low range/severe    | -2 SD and below      |

---

---

---

---

---

---

---

---

## Explaining Standard Scores (continued)

- Example of definition for the team (parents, in particular):
  - After scoring your child's responses/performance on this test, a score was calculated by comparing his performance to other children his age.
  - This score is called a standard score.
  - We use standard scores to see where your child is performing compared to other children his age and to measure his progress over time.

---

---

---

---

---

---

---

---

## Explaining Percentile Ranks

- Should be used to support standard scores
  - Report a range of normal to which students' scores can be compared
  - Does not follow an equal interval scale, but can measure students' progress over time (at 3 year re-evaluation)
  - Can be compared across language modalities, subtests, and tests
  - But are easily misinterpreted
    - be sure to explain the difference between percentile ranks and percentages

---

---

---

---

---

---

---

---

## Explaining Percentile Ranks (continued)

- Always use a visual when explaining percentiles to the IEP team and demonstrate where in the bell curve the PR falls
  - Explain where average falls
  - Percentiles of 16 - 85 are within normal limits (which is another problem of percentile ranks!)

---

---

---

---

---

---

---

---

## Explaining Percentile Ranks (continued)

- Example of definition for the team (parents, in particular)
  - The percentile rank indicates how many students in your child's age range performed above him or below him on the test.
- Be sure to give an example
  - Your child received a percentile rank of 30.
  - This means that he scored better than 30% of the other children his age on this test.

---

---

---

---

---

---

---

---

## Conclusions

- IEP teams consist of many members with a variety of backgrounds in the use and interpretation of psychometric data.
- Members of the team may consist of:
  - school psychologists with extensive knowledge,
  - teachers with limited knowledge
  - parents with no knowledge in this area
- Focus your presentation on standard scores and percentile ranks, regardless of what scores are available in the manuals (*and* stay away from AEs)

---

---

---

---

---

---

---

---

## Conclusions (continued)

- SLPs often have both the knowledge and the ability to interpret these scores across various types of tests.
- Therefore, it is often necessary that SLPs explain psychometric data to team members with varying levels of knowledge
- Use of clear terminology and visual examples will greatly increase understanding

---

---

---

---

---

---

---

---

## Contact Information

- [Emsros@aol.com](mailto:Emsros@aol.com)
- [Linda.Larrivee@worcester.edu](mailto:Linda.Larrivee@worcester.edu)

---

---

---

---

---

---

---

---