


Statistics for Assessors with a deep dive into WRIT tables

Confidence in Scoring
Sarah Beard
Dyslexia Specialist Assessor, Tutor and Trainer
dittas



1 © dittas 2023

Learning outcomes

By the end of the session Assessors will be able to:

Understand:

- Statistical terminology
- Prevalence and significance

Know:

- How to use the WRIT Tables to work out statistical significance, prevalence and subtest scatter
- How to calculate confidence bands for standard and scaled scores

Communicate:


- This information to others

2 © dittas 2023

Overview

- **Part 1**
- Review of statistical terminology
- **Part 2**
- Calculating Confidence bands
- **Part 3**
- Deep dive into WRIT tables

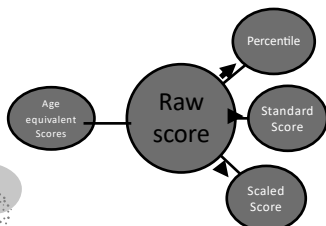
3 © dittas 2023



PART 1

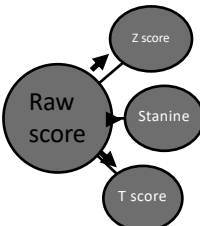
4 © dittas 2023

Converting Raw Scores



5 © dittas 2023

Converting Raw Scores




6 © dittas 2023

Activity

- Cut and paste the link below to complete an activity matching statistical terms

<https://wordwall.net/play/51636/579/268>

Or scan the QR code



wordwall.net/resource/51636579

Standard Scores

- Standard scores show the individual's position compared to the mean for their age group and are the most appropriate type of score to use when evaluating an overall picture of results.
- Most tests use a scale with a mean of 100 and a standard deviation of 15.

Scaled Scores

- Scaled scores, also referred to as standard scores, are standard scores with a range of 1-20, a mean of 10 and a standard deviation of 3.
- The average range is 7-13.
- Can be converted by using a conversion chart or
- $(5 \times \text{scaled score} + 50)$

8	90
9	95
10	100
11	105
12	110

Percentiles

- Percentiles (or centiles) show the percentage of individuals who have a score the same or lower than the testee.
- i.e. a percentile score of 58 means that 58% of chronological age peers would score the same or lower than the testee.
- Percentiles of 15 and below are below average. Percentiles of 85 and above are above average.
- Note: the gap between percentiles is not regular so percentile scores can be misleading.

Composite Scores

- Composite scores combine scores from several subtests, often these subtests scores are in scaled score form, but when combined a composite score is obtained and this is usually given as a standard score.
- For example, on the CTOPP Rapid Naming test, a scaled score of 6 on Digit Naming and 7 on Letter Naming would yield a composite score of 79.
- Composite scores are sometimes called index scores.

Stanines

- Stanine is short for 'standard nine'. The distribution curve is therefore separated into nine sections.
- A stanine score ranges from a low of 1 to a high of 9.
- Stanines cover a range of standard scores and can only be converted to a standard scores range i.e. a stanine of 3 is equivalent to a standard score of between 82-88.

Stanine Conversion

Stanine	Standard score equivalent	Range descriptor/s
1	73 and below	Very low
2	74-81	Low
3	82-88	Below average - Low average
4	89-96	Average
5	97-103	Average
6	104-111	Average - High average
7	112-118	High average - Above average
8	119-126	High
9	127 and above	Very high

Age Equivalent Scores

Not recommended for secondary age and beyond, because:

- Rate of reading progress slows down as an individual gets older.
- A reading age of 12-13 years is sufficient to decode most material.
- A reading age of 12 for a 15 year old may suggest a difficulty that is not present.
- Not good for self esteem.

Why we don't use age equivalent scores

Child age 10 (WRAT 5 Blue):

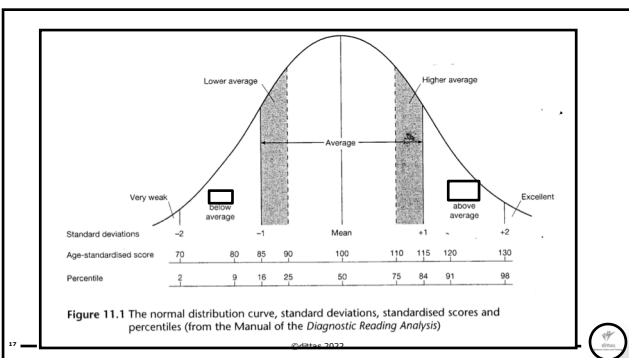
Raw score	Standard score	Centile	Age equivalent score (years)
36	86	18	8.4 - 8.7
49	114	82	12.0 - 12.05

A 19 year old with a raw score of 53 has a standard score of 89 but a reading age of 13.6-13.11

Can you see the problem?

Normal Distribution Curve

- A smooth symmetrical curve (bell shaped) with mathematical properties.
- A normal distribution curve is a distribution that occurs naturally in many situations.
- The mean is the central point – the arithmetical average score for the reference group on whom the test was standardised.
- To measure the variation in the scatter of scores the standard deviation is calculated.



Standard Deviation

- The Standard Deviation is the average amount by which scores differ from the mean. One SD is 15 points.
- The distribution curve is sliced into vertical parts, which are each one standard deviation wide.
- 68% of the population fall 1 SD either side of the mean (-1 and +1) This is the 'average range.'

Range	Standard Scores	Centile	Standard deviation from mean	% of population
well above average	131 and above	98	+2	2%
above average	116 - 130	84 - 97	+1 to +2	14%
average	85 - 115	16 - 83	-1 to +1	68%
below average	70 - 84	2 - 15	-1 to -2	14%
well below average	69 and below	below 2	-2	2%

Reliability

Reliability is the degree to which a test provides consistent results

Several different types of reliability:

- Test/retest reliability
- Administrator reliability
- Parallel form reliability

What reliability coefficient should we look for in a test?

Internal Reliability

- Internal reliability is represented by the reliability coefficient.
- $r = .87$ means 87% of variance is true and 13% likely to be due to sampling error. **When choosing a test you should look for tests with reliability coefficient of over 0.8**

"A reliable test produces the same results on repeated trials and remains stable and consistent over time"

Validity

Does the test really measure what it says it does?

Many types of validity including:

- Content Validity
- Construct Validity
- Criterion Related Validity
- Predictive Validity

Deficit or Discrepancy

These terms are often confused:

- Deficit is a very weak standard score of at least 1.5 standard deviations (SD) below the norm
e.g. 77 is considered a mild deficit and 69 is considered a significant deficit.

Deficit or Discrepancy


These terms are often confused:

- A discrepancy between scores means a difference between two scores, **usually of at least one SD**

Standard Error of Measurement (SEM)

- Allows us to quantify the extent to which a test provides an accurate score.
- SEMs are based on a calculation using the SD and reliability.
- SEMs are used to calculate the confidence band.

21 © ditta2023



PART 2 Confidence Bands and SEM

22 © ditta2023

Standard Error of Measurement (SEM)

Where to find SEM

SEM can usually be found in the test manual or the examiner booklet.

The SEM for the following tests are on the examiner booklets:

- CTOPP2
- GORT 2
- TOWRE
- DASH 9-16 and 17+

21 © ditta2023

Confidence Intervals

Confidence bands or intervals provide a ring of confidence around a standard score

- e.g. if someone achieved a standard score of 90 on the TOWRE2 Sight Word Efficiency test we can be 95% confident that their score will always fall between a standard score of 80 -100.

22 © ditta2023

Calculating Confidence Bands

- Some test manuals provide tables with all the confidence bands for each standard score i.e. WRIT, WRAT4
- Other manuals provide the confidence interval that has to be applied to the obtained standard score i.e. WRAT5, WIAT- 3T
- Other manuals provide only the SEM and require the assessor to calculate the confidence band.

23 © ditta2023

Calculating Confidence Bands

1. Find the SEM on the examiner booklet or in the manual	⇒	1. e.g. the TOWRE2 Sight Word Reading, SEM is 5
2. Multiply the SEM by two	⇒	2. $5 \times 2 = 10$
3. Apply to the standard score by subtracting and adding to obtained score	⇒	3. For a standard score of 90, subtract 10 from one side and add 10 to the other. We can say we are 95% confident that on any day this individual's score would fall between a standard score of 80-100.

24 © ditta2023

Activity

Let us work another confidence band out.

The SEM for CTOPP2 Phonological Awareness is 4, so let us work out the confidence band for a standard score of 82.

I will give you a moment to work this out and then I will go through it on the next slide.

Confidence Bands

1. SEM of CTOPP PA is 4
2. So $2 \times 4 = 8$
3. Standard score was 82.
So take 8 from 82 to get 74.
Then add 8 to 82 to get 90.

So the 95% confidence band for a standard score of 82 is 74 – 90.

Activity

Calculate the confidence bands for the following scores:

TEST	STANDARD SCORE	SEM
TOWRE 2 Word Reading Efficiency	80	3
CTOPP2 Rapid Naming	85	4
DASH 9-16	70	5


Answers

Calculate the confidence bands for the following scores:

TEST	STANDARD SCORE	SEM	95% confidence
TOWRE 2 Word Reading Efficiency	80	3	74 - 86
CTOPP2 Rapid Naming	85	4	77 - 93
DASH 9-16	70	5	60 - 80

Activity

- Cut and paste the link below or scan the barcode to complete an activity calculating SEM and confidence bands.
Note watch out for the WIAT 3T !
- <https://wordwall.net/resource/51637168>



Calculating confidence bands for scaled scores

1. Find the SEM on the examiner booklet or in the manual → 1. e.g. CTOPP2 PM gives a SEM of 1
2. Multiply the SEM by two → 2. $1 \times 2 = 2$
3. Apply to the standard score by subtracting and adding to → 3. Standard score of 6, subtract 2 from one side and add 2 to the other so we get a confidence band of 4 - 8.
4. Convert the 3 scaled scores to standard scores → 4. So we can say we are 95% confident that on any day the confidence band for a standard score of 80 is 70 - 90.

Calculating Confidence Bands (index scores)

If the SEM is a decimal you should multiply by 1.96 not 2

- Find the SEM on the examiner booklet or in the manual → 1. e.g. TOMAL SEM for ACI of a 10 year old is 2.01 $2.01 \times 1.96 = 3.94$ (round up to 4)
- Multiply the SEM by 1.96 →
- Apply to the standard score by subtracting and adding to obtained score. → 2. Index score of 90, subtract 4 from one side and add 4 to the other. We can say we are 95% confident that the score would fall between a standard score of 86 - 94 on any day.

Calculating Confidence Bands (scaled scores)

If the SEM is a decimal you should multiply by 1.96 not 2

- Find the SEM on the examiner booklet or in the manual → 1. e.g. TOMAL SEM for Letters Backwards, age 10, is 0.72
- Multiply the SEM by 1.96 → 2. $0.72 \times 1.96 = 1.41$ (round down to 1)
- Apply to the standard score by subtracting and adding to obtained score → 3. Scaled score of 7, subtract 1 from one side and add 1 to the other so confidence band is 6 - 8.

Calculating Confidence Bands

If the SEM is a decimal you should multiply by 1.96 not 2

- We can say we are 95% confident that the score would fall between a scaled score of: → 4. 6-8
- We now need to convert that to standard score → 5. So 80-90

Calculate the confidence bands for the following scores on TOMAL for a 17yr old:

TEST	SCALED SCORE	SEM	Confidence band scaled	Confidence band Standard score
Digits Forward	6	.56		
Digits Backwards	5	.52		

Calculate the confidence bands for the following scores on TOMAL for a 17yr old:

TEST	SCALED SCORE	SEM	Confidence band scaled	Confidence band Standard score
Digits Forward	6	.56	5-7	75-85
Digits Backwards	5	.52	4-6	70-80



PART 3 WRIT TABLES

Wide Range Intelligence Test (WRIT)

In this section of the presentation we will look at:

- Verbal/visual discrepancies
- Statistical significance/critical values
- Prevalence/base rates
- Sub Test scatter



WRIT Tables

- | | |
|---|-----------|
| • 6.1 Range descriptors | • Page 65 |
| • 6.2 Minimum and maximum IQs by age | • Page 67 |
| • 6.3 Statistical significance of verbal and visual IQ discrepancies | • Page 68 |
| • 6.4 Cumulative Percentages obtaining verbal/visual IQ discrepancies | • Page 70 |



WRIT Tables

- | | |
|---|---------------------|
| • Average of 4 subtests contributing to general IQ | • Table 6.5 page 75 |
| • Differences between individual subtests required for statistical significance | • Table 6.6 page 76 |
| • Cumulative Percentage obtaining various amounts of subtest scatter | • Table 6.7 page 77 |



Statistical Significance

Statistical significance refers to the probability that's the difference between two scores has not occurred by chance.

When a finding is significant, it simply means you can feel confident that's it not by chance.



Verbal/Visual Score Discrepancies

- Table 6.3, on page 68 of the WRIT manual provides information on the statistical significance of verbal/visual score discrepancies.
- It is important to decide if to combine visual/verbal scales to calculate a general ability standard score based on the statistical analysis.



Significance

- The 0.01 significance level means that there is only one chance in a hundred that this could have happened by chance.
- The 0.05 level tells you that it could have happened by chance five times in a hundred, that it has a 95% chance of being a real difference.



Section of table 6.3, page 68; verbal/visual discrepancies

Age	Level of statistical significance	Verbal/Visual
13 - 18	0.05	13
13 - 18	0.01	17
Adult (19+)	0.05	9
Adult (19+)	0.01	12

Activity

Using table 6.3, page 68 or the the table on the next slide, find out if the differences are significant and if so, what is the significance.

1. A verbal score of 112 with a visual score of 88 for a 14 year old.
2. A verbal score of 91 with a visual score of 78 for a 15 year old.
3. A verbal score of 85 with a visual score of 75 for a 22 year old.

Answers

Age	Level of statistical significance	Verbal/Visual
13 – 18	0.05	13
13 – 18	0.01	17
Adult (19+)	0.05	9
Adult (19+)	0.01	12

1. A verbal score of 112 with a visual score of 88 for a 14 year old
➤ 24 points 0.01

2. A verbal score of 91 with a visual score of 78 for a 15 year old.
➤ 13
➤ 13 points 0.05

3. A verbal score of 85 with a visual score of 75 for a 22 year old
➤ 10 points

Prevalence

- Just because a score is statistically significant it does not mean it is important or rare. Table 7.4 page 70 show what score differences occurred in the standardisation sample.
- To find out if the size of a score discrepancy is very unusual we look at the table on page 70 (Table 6.4) or page 77.

Table 6.4 (p 6.4) Cumulative Percentages Obtaining Various Verbal IQ - Visual IQ Discrepancies.

- This table shows us what % of the standardised sample obtained a specific discrepancy.
- Note this table is "two tailed" meaning is has to be divided by two (see page 69 of the manual)

Using table 6.4 page 70 in your manual

- Work out the prevalence for:
- A discrepancy of:
 - Age 14 of 24 points
 - Age 15 of 13 points
 - Age 22 of 10 points

Answers

- Work out the prevalence for:
- A discrepancy of:
 - Age 14 of 24 points (9/2) 4.5%
 - Age 15 of 13 points (31/2) 15.5%
 - Age 22 of 10 points (46/2) 23%

Table 6.5 page 75 Average of 4 subtests contributing to general IQ

This table allows us to see if a subtest score is significantly different to the general ability. Using table 6.5 we can compare this to the general ability score to see if it significant at the 0.01 or 0.05 level.

Helps us consider if to combine

Table 6.6 page 76 Differences Between Individual Subtest

- Table 6.6 page 76 shows the statistical significance between individual subtests.

Activity

- Work out the statistical significance for the following scores:
- Vocabulary 112 and Analogies 97 for an 8 year old
- Matrices 89 and Diamonds 101 for a 20 year old

Answers

- Work out the statistical significance for the following scores:
- Vocabulary 112 and Analogies 97 for an 8 year old - 0.05
- Matrices 89 and Diamonds 101 for a 20 year old - 0.05

Table 6.7 page 77 Cumulative Percentage obtaining various amounts of subtest scatter for individual subtests

- Take the lowest subtest score away from the highest
- Matrices 118
- Diamonds 112
- Analogies 110
- Vocabulary 89

Subtest scatter is 29 points and if you look at table 6.7 you will see this occurred in 25% of the standardization sample.

Work out the subtest scatter - table 6.7 p77

Age 13	Age 10
<ul style="list-style-type: none"> • Matrices 100 • Diamonds 79 • Verbal Analogies 120 • Vocabulary 110 • Difference • Subtest scatter 	<ul style="list-style-type: none"> • Matrices 104 • Diamonds 106 • Verbal Analogies 74 • Vocabulary 91 • Difference • Subtest scatter

Work out the subtest scatter - table 6.7 p77

Age 13	Age 10
<ul style="list-style-type: none"> • Matrices 100 • Diamonds 79 • Verbal Analogies 120 • Vocabulary 110 • Difference 41 • Subtest scatter 5% 	<ul style="list-style-type: none"> • Matrices 104 • Diamonds 106 • Verbal Analogies 74 • Vocabulary 91 • Difference 30 • Subtest scatter 25%

Activity

- In your course materials find the handout

'Activity WRIT calculations'

Pause the presentation to complete this.

Discussion point

Consider the quote below.


Assessors must remember however that, *'tests don't diagnose, people do'* and base their diagnoses exclusively on test results, a hazardous enterprise at best. "Test Results are merely observations.... But do not tell the examiner why the person performed as he or she did."

(Wagner, Torgesen and Rashotte, 1999, in Backhouse and Morris 2005, dyslexia? assessing and reporting: The Patoss Guide, Hodder Murray) page 11

Activity

Cut and paste the link below to complete an activity matching statistical terms

<https://wordwall.net/resource/51952626>



Next steps

1. Complete the Interactive activities if you have not already done so.
2. Fill in the google form called 'SASC Accredited Training Certificate Request ' Once this is received your certificate will be sent within 10 days

- This is the link:
- <https://forms.gle/76W5BT81vgitdyVA8>
- Have questions ?
- Email us or leave a comment on the google form.

Thank you

- Visit our website for details of our services including webinars, mentoring, tuition and assessments
- Leave a comment on our Facebook Page
- Dittas Dyslexia Services (<https://www.facebook.com/dittaslimited>)



© ditta 2023

