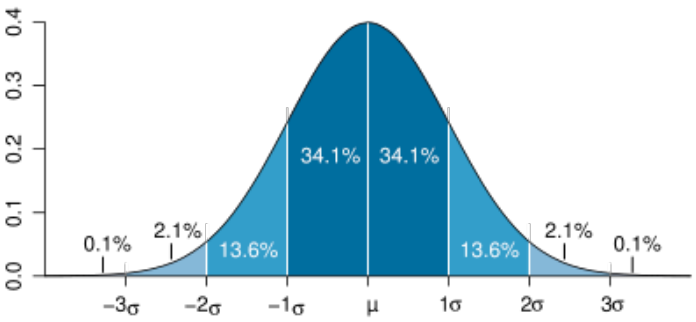
Week 8

- Tuesday:
  - Interpretation of NP Test Results (begin)
- Thursday:
  - Interpretation of NP Test Results (finish)
  - Neuropsychological Tests - Practice Lab

Interpretation of NP Test Results

- Standard Scores
- Normative Scores
- Demographic Corrections
  - Age, Education, Gender and Race/Ethnicity
  - Demographic Norms Example
    - Example paper (Diehr et al 1998)
- Premorbid Functioning
  - Qualitative Estimates
  - Quantitative Estimates
- Psychometric Issues

Normal Distribution

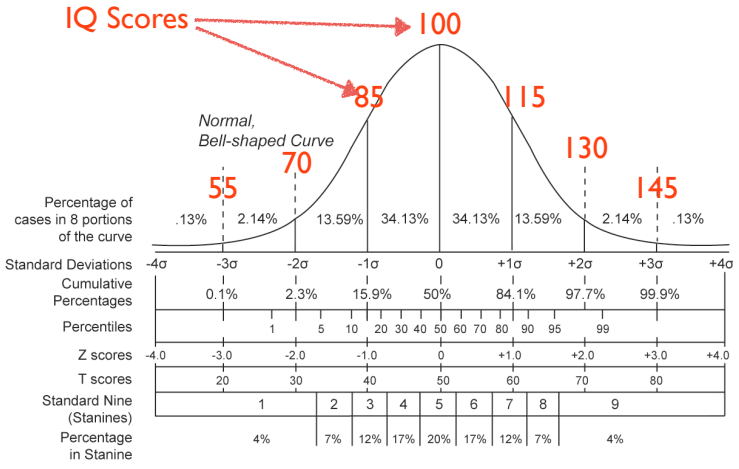


Many measures show a normal distribution

Raw vs. Standard Scores

- Raw score - the # or percent correct
  - range : depends on test
  - mean : depends on test
  - standard deviation : depends on test
- Standardizing scores
  - put all tests on same metric
  - allows comparisons across tests
  - method
    - Describe scores as distance from Mean, with distance measured against the SD
- Problem:
  - too many "standard" standard scores!

Standard Scores



Standard Scores

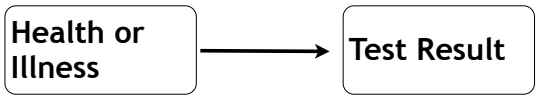
	Z scores	IQ scores	T scores	Scaled Scores
Mean	0	100	50	10
SD	1	15	10	3
1 SD below the mean	-1.0	85	40	7
2 SD below the mean	-2.0	70	30	4

Normative Data

- Normative Group
  - For each test, give test to a group of "normal" people and measure scores.
    - group should match society (age, education, SES, gender, ethnicity…)
  - Determine Mean (SD), call this "normal"
- For an individual under assessment…
  - compare observed score to normative data
    - data tables
    - formulas
    - computer programs

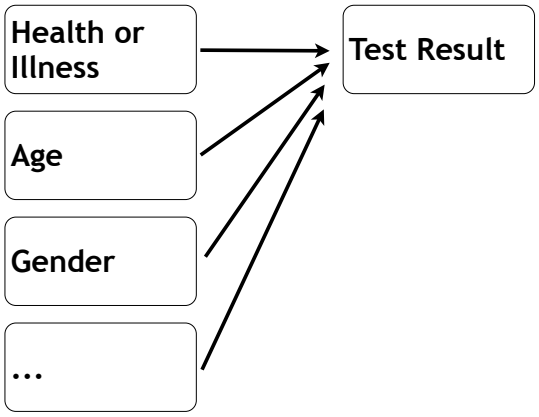
Demographic Corrections

- Normative Data would be simple if there was one definition of "normal"
- Simple model



Demographic Model

- More realistic model



Common Demographic Covariates

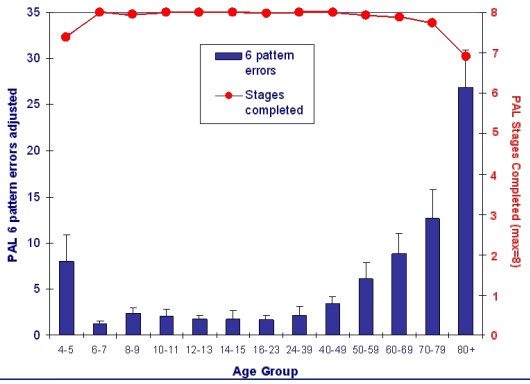
- Age
- Sex
- Race/Ethnicity
- Education
- Language

Age

- Once you hit 35 years old, it's all downhill!
- Gets significantly worse
  - fluid reasoning
  - speed of processing
  - working memory
  - long term memory
- Gets a little worse or stays same
  - Crystallized information
  - semantic memory
- May get better
  - reasoning
  - vocabulary
  - emotion
  - metacognition

CANTAB Paired Associate Learning

Paired Associate Learning (errors)– Computerized Task (n=1444)



Gender

- Historical conceptions
- Modern data

Gender and IQ

- It was commonly accepted in the 1800s that men were intellectually superior to women
- Darwin, Descent of Man (1871) "The chief distinction in the intellectual powers of the two sexes is shewn by man's attaining to a higher eminence, in whatever he takes up, than can woman - whether requiring deep thought, reason, or imagination, or merely the use of the senses and hands"
- Book was edited by Darwin's daughter Henrietta and wife Emma.
- Darwin was in other ways socially liberal

Gender and IQ

- Modern scientific consensus is that men and women, on average, have equal IQ scores.
- Differences are small and generally insignificant (1-3 IQ points when differences are found)
- Men's IQ scores tend to be slightly more variable (higher variance) so more men tend to fall at either end of the spectrum
- Some evidence that males are better at stereotypical "male" tasks (visuospatial skills) whereas women are better at "female" tasks (language).
  - Genes vs. Environment?
  - Evolutionary reasons?
  - Hormones such as Testosterone?

Race vs. Ethnicity

- Race - genetic heritage
- Ethnic group -- population whose members identify with each other
- "National, religious, geographic, linguistic and cultural groups do not necessarily coincide with racial groups: and the cultural traits of such groups have no demonstrated genetic connection with racial traits. Because serious errors of this kind are habitually committed when the term "race" is used in popular parlance, it would be better when speaking of human races to drop the term "race" altogether and speak of 'ethnic groups'.

Effects of Race & Ethnicity on NP Scores

- Many NP tests show effects of Race
  - widespread agreement
- Reasons for these differences?
  - great disagreement
  - Genetics?
  - Proxy for other variables (nutrition, education, social & environmental opportunities & rewards, money…)

The Flynn Effect - 100 years

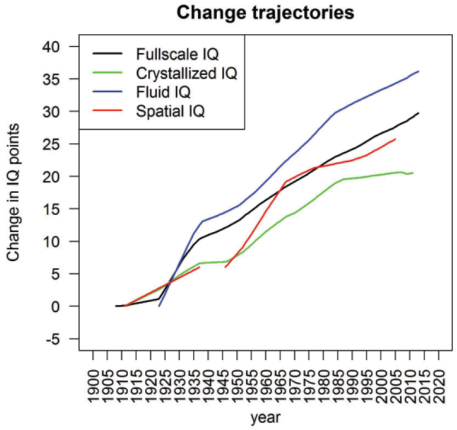
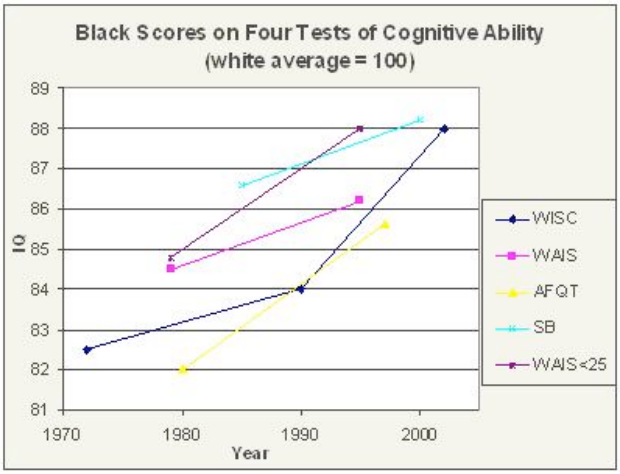


Fig. 1. Domain-specific IQ gain trajectories for 1909–2013. Changes are based on weighted average annual IQ changes in all available data.

Flynn Effect: shrinking racial differences



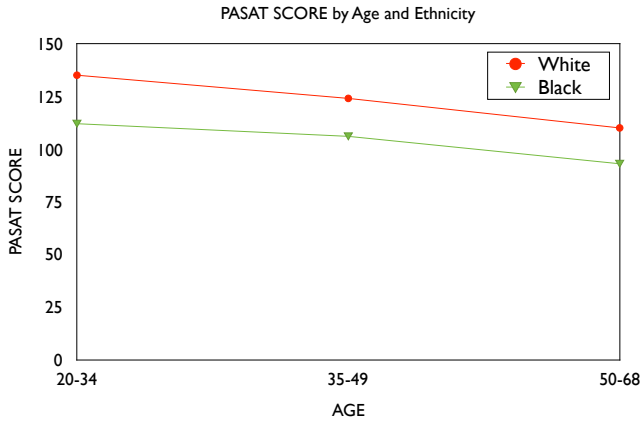
Diehr et al. (2003)

- Paced Auditory Serial Addition Test
- Demographic variables:
  - ethnicity : self-identified as either White / Black
  - age
  - gender
  - education

Diehr et al. (2003) Results

- Statistically and Clinically-significant differences in PASAT test result found for
  - Age
  - Education
  - Ethnicity
- Results not statistically significant:
  - Gender

Diehr et al. (2003) Results



Education

- Is education a demographic variable?
- Can educational level be a result, not a cause, of brain development?

Demographic Variables: Conclusions

- Age difference : Large
- Gender differences on IQ tests are small (less than .2 SD) but still controversial
- Ethnic differences are fairly large (1.0 SD)
  - Explanations:
    - Test bias? some found, but doesn't explain most of difference
    - Genetic differences? yes, but recent results suggest this % has been vastly over-estimated.
    - Environmental differences? yes, explains a great deal, and explains recent reductions (e.g. Flynn effect)
- Education differences : Large
  - probably a proxy for other factors, but works well

Estimating Premorbid Functioning

- Interpretation of NP test results is most accurate when comparing pre- and post- test results
- Rarely have premorbid tests
- What to do?

- Estimates from demographic variables
  - age, education, gender, ethnicity
- Estimates from vocation/skills
  - nature of work done

- Ideal: a test which is highly correlated with IQ but not affected by most forms of brain damage

Premorbid Functioning : the NAART

- North American Adult Reading Test
- Pronunciation of 61 word reading list
- Correlates highly with IQ
- Correlates very weakly with many forms of brain damage

- Automatic Process
- Not timed

NAART Estimating Premorbid Verbal IQ

- Given NAART score, estimate premorbid Verbal IQ (via on their Vocabulary test Scaled Score)

- 4th equation includes Education correction

WAIS–R Vocabulary
$= 31.30 + 0.622 \times \text{NAART (SEE} = 5.14)$
WAIS–R Vocabulary
$= 25.71 + 0.566 \times \text{NAART} + 0.508$
$\times \text{Education (SEE} = 5.02)$
WAIS–R Vocabulary Scaled
$= 5.383 + 0.179 \times \text{NAART (SEE} = 1.71)$
WAIS–R Vocabulary Scaled
$= 4.112 + 0.167 \times \text{NAART} + 0.115$
$\times \text{Education (SEE} = 1.69)$

Psychometric Issues

- Standardization
- Reliability
- Validity
- Type I and II errors - Prevalence & Hit Rates

Standardization

- Tests must be standardized in methodology

- Issues
  - feedback, encouragement
  - motivation

Reliability

- Reliability
  - whether a test gives a consistent result in the same situation
  - how much statistical noise is present
  - Example:
    - We believe IQ is relatively fixed
    - IQ test should give same results when administered
      - over time
      - by different examiners

Validity

- Validity
  - whether a test result "means what it says"
  - is the conclusion (or prediction) made using the test accurate
  - Example:
    - We believe IQ is related to employment success
    - Someone with low IQ
      - Prediction: "They will not win Jeopardy"
      - Result : ???

Types Of Errors

		The patient actually is...	
		Sick	Healthy
You decide patient is...	Sick	True Positive 1-β **"Hit"**	False Positive Type I Error α **"False Alarm"**
	Healthy	False Negative Type II Error β **"Miss"**	True Negative 1-α **"Hit"**

Hit Rates

- Hit Rate
  - % of people correctly classified
- Sensitivity
  - probability of detecting illness
- Specificity
  - probability of detecting health
- Positive Predictive Value
  - probability of being ill when you test ill
- Negative Predictive Value
  - probability of being healthy when you test healthy

Base Rates (Prevalence) and Hit Rates

- If the disease is very rare (has a low Base Rate or Prevalence)
  - most of the diagnoses are false positives
- SPIN / SNOUT:
- SPIN : When **S**pecificity is high, **P**ositive result rules **IN** the condition
- SNOUT : When **S**ensitivity is high, **N**egative result rules **OUT** the condition.

Low Prevalence Example

- Imagine a disease that's only found in 1/1000 patients.
- You have a test with 95% specificity, 95% sensitivity (in other words, a good test).
- In a group of 1000 patients,
  - you will diagnose 50 of them as having the disease
  - but only 1 actually does

Exercise 2: NP Test Practice

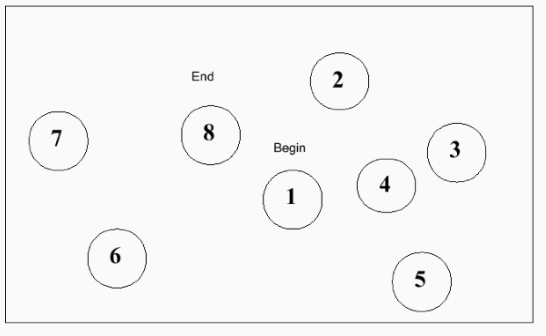
- In which we practice giving two common neuropsychological tests
- And we practice scoring the tests, converting to standardized scores, and interpreting the results

Trail-Making Test A and B ("Trails A and B")

- History
  - Originally part of Army Individual Test of General Ability (1944)
  - Incorporated into Halstead-Reitan Battery
- Procedure
  - Part A: subject connects the dots in order
    - 1,2,3,4,5...
  - Part B: subject connects the dots in order, alternating numbers with letters
    - 1,A,2,B,3,C...

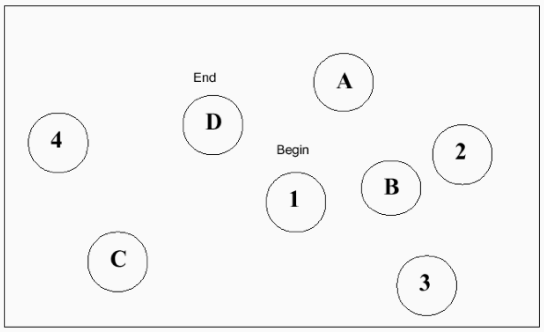
Trails A

Trail Making Test Part A – *SAMPLE*



Trails B

Trail Making Test Part B – *SAMPLE*



Trails - Practice

- Break into teams of 2 or 3
- Administer Trails A, and Trails B to each other
- Use your non-dominant hand?
  - more interesting results

Trails A and B - Abilities

- Trails A
  - visual perception
  - motor control
  - number sequencing
  - load on Working Memory: Low
- Trails B
  - same as trails A, plus…
  - working memory - letter/number sequencing
  - executive control : *inhibition* of over-learned responses
    - (temptation to go from 1 to 2, rather than 1 to A, or from A to B, rather than A to 2)

Trails A&B - Demographic Effects

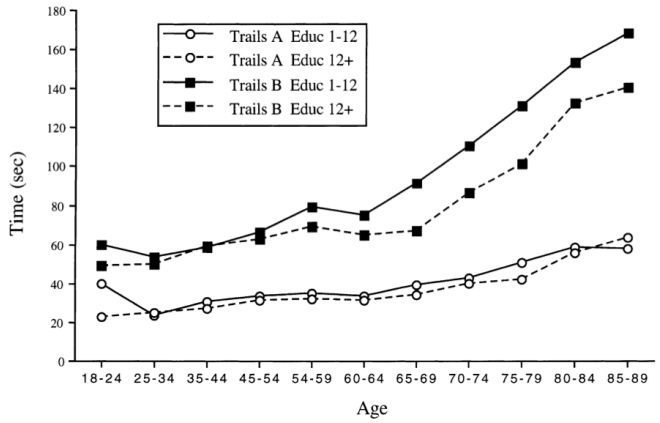
- Age: r=0.58 to 0.62
- Education : r=0.17 to 0.25
- Gender : 0.05 (NS)
- Trails A vs. B: 0.74

Table 1
Correlations of age, education, gender with time (s) to complete Trails A and B

	Age	Education	Gender	Trail A
Age				
Education	−.17**			
Gender	−.08*	−.03		
Trail A	.58**	−.17**	−.05	
Trail B	.62**	−.25**	−.05	.74**

* $p < .05$.
** $p < .01$.

Trails A&B - Demographic Effects



Trails A - Scoring Example

- First, determine your Raw score
  - X = _____
- Next, determine demographic variables (age, education, gender…)
- For given demographics, determine normative score
  - mean (M) = _____
  - standard deviation (SD) = _____
- Next, convert raw score to a standard score using formula:
  - Z score = (X-M) / SD
  - the Z score is simply the distance from the mean, expressed in units of 1.0 SD
  - if your score is 1 SD below the mean, then Z = -1.0, etc.
  - convert standard score to description "High average" etc.

Estimating Years of Education

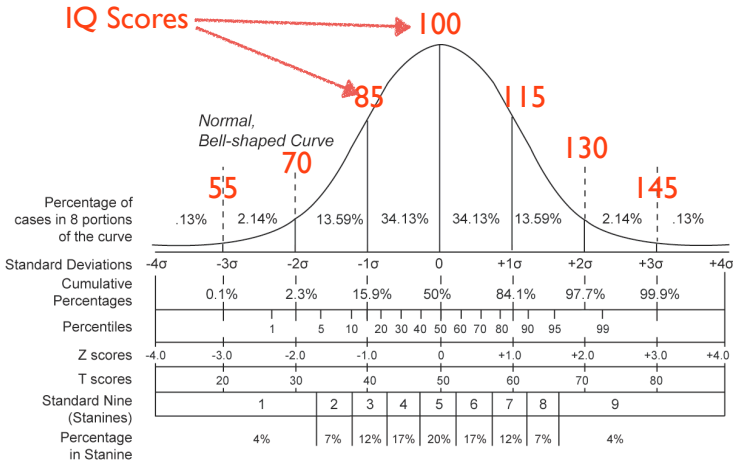
- Many normative data sets include demographic corrections for education
- Rules about how to calculate education: idiosyncratic
- When in doubt, read the original article
- In general, use highest level reached - don't double-count multiple degrees
  - 8 = Middle School
  - 12 = High School
  - 14 = Associates Degree (2-year college)
  - 16 = Bachelor's Degree (4-year college)
  - 18 = Master's Degree
  - 20 or 21 = Ph.D. or M.D.
- Example: person with a AA, B.S., M.D. and M.S.
  - 21 years (not 23 or 25 or 27)

Trails A,B Norms

Table 2
Statistical properties for age, education, gender, Trails A and B (s) for each normati

	Statistics		
Age groups	Mean (S.D.)	Median	Minimum–maximum
Age group 18–24 ($n = 155$)			
Age	20.17 (1.48)	20.00	18–24
Education	12.92 (1.01)	13.00	10–15
Gender	1.59 (0.49)		
Trail A (s)	22.93 (6.87)	21.70	12–57
Trail B (s)	48.97 (12.69)	47.00	29–95
Age group 25–34 ($n = 33$)			
Age	29.42 (2.87)	30.00	25–34
Education	14.18 (1.61)	14.00	11–18
Gender	1.58 (0.50)		
Trail A (s)	24.40 (8.71)	23.00	10–45
Trail B (s)	50.68 (12.36)	50.00	29–78
Age group 35–44 ($n = 39$)			
Age	39.74 (2.94)	41.00	35–44
Education	13.59 (2.06)	14.00	10–20
Gender	1.59 (0.50)		
Trail A (s)	28.54 (10.09)	26.00	12–50
Trail B (s)	58.46 (16.41)	58.00	29–95

Standard Scores



Standard Scores

	Z scores	IQ scores	T scores	Scaled Scores
Mean	0	100	50	10
SD	1	15	10	3
1 SD below the mean	-1.0	85	40	7
2 SD below the mean	-2.0	70	30	4

Describing Performance (WAIS-4 Terminology)

Classification	IQ Score	Z Score	T-Score	Percentile (%ile)
Very superior	above 130	above 2.0	70 and above	98% and above
Superior	120-129	1.3 to 2.0	63-39	91% - 97%
High average	110-119	0.6 to 1.3	56-62	75% - 90%
Average	90-109	-0.6 to +0.6	44-55	25% - 74%
Low average	80-89	-0.6 to -1.3	43-37	9% - 24%
Borderline	70-79	-1.3 to -2.0	36-30	2% - 8%
Extremely low	69 and below	below -2.0	29 and below	below 2%

COWAT - Verbal Fluency

- Controlled Oral Word Association Test
  - aka Verbal Fluency
  - aka "FAS, Animals"
- History
  - Incorporated into Halstead-Reitan Battery
- Procedure
  - FAS - phonemic
    - name as many words as possible starting with F, A, S
  - Animals - semantic
    - name as many animals as possible

COWAT - Abilities

- FAS
  - verbal fluency / processing speed
  - short term verbal memory
  - semantic memory
  - word associations by phonemic lookup
  - mediated by Frontal Lobe?
- Animals
  - word association by semantic lookup
  - mediated by Temporal lobe?

COWAT - Demographic Effects

- Age: r=.02 to .07
- Education : r=.08
- Race : r=.10
- Gender : NS

COWAT - Administration - Practice

Appendix A

Instructions for Letter (FAS) and Category (Animals) Fluency Tasks

Verbatim Instructions for the Controlled Oral Word Association Test (FAS)

"I am going to say a letter of the alphabet to you, and I want you to tell me as many words as you can think of that begin with that letter. But none of the words can be proper names of people or places. For instance, if I gave you the letter "B," you could say "brook, bottle, black," and so forth, but you could not say "Barbara" since that is a person's name, nor could you say "Boston," since that is the proper name of a place. Also, do not give me the same word with different endings, such as sit, sits, and sitting."

"The first letter we will use is "F." Go ahead and tell me as many words as you can think of that begin with "F."

(Begin timing. Record all responses verbatim. Do not interrupt the respondent or ask him or her to slow down. It is permissible to repeat instructions if the respondent loses set or forgets what he or she is supposed to be doing. Stop the respondent after 60 seconds. "A" and "S" trial are introduced in the same manner as above.)

Verbatim Instructions for Category Fluency

"Now we are going to do something a little different. This time I want you to tell me all of the animal names that you can think of. It doesn't matter what letter they start with. Just tell me all of the animal names that you can think of."

(Record the animal names in the same manner as above.)

COWAT - Scoring Example

- First, determine your Raw score
  - X = _____ (for FAS, total across all letters)
- Next, convert Raw score to Scaled score (see table)
  - Scaled = _____
- Next, determine demographic variables (age, education, gender, ethnicity)
- Use formula to convert Scaled score to T score
  - T-score = _____
- Now, double-check by comparing computed T score with T-score from lookup table

COWAT - Raw to Scaled Conversion

Table 4
Scaled Score Equivalents to Raw Scores for Letter (FAS) and Category (Animal) Fluency

	Raw score	
Scaled score	FAS	Animal
19	78+	37+
18	73 - 77	33 - 36
17	67 - 72	31 - 32
16	63 - 66	30
15	58 - 62	29
14	54 - 57	27 - 28
13	50 - 53	25 - 26
12	46 - 49	23 - 24
11	42 - 45	21 - 22
10	37 - 41	19 - 20
9	33 - 36	17 - 18
8	29 - 32	15 - 16
7	26 - 28	14
6	21 - 25	13
5	18 - 20	12
4	15 - 17	11
3	13 - 14	10
2	0 - 12	8 - 9
1		0 - 7

COWAT - Formula

Appendix B

Letter (FAS) and Category (Animal) Norms Formulas

Demographically corrected *T* scores for fluency can be calculated as follows:

Letter (FAS) *T* score = 14.796 + (3.584 x FAS Scaled Score) – (0.914 x Education) + (0.177 x Age) + (5.470 x Race)

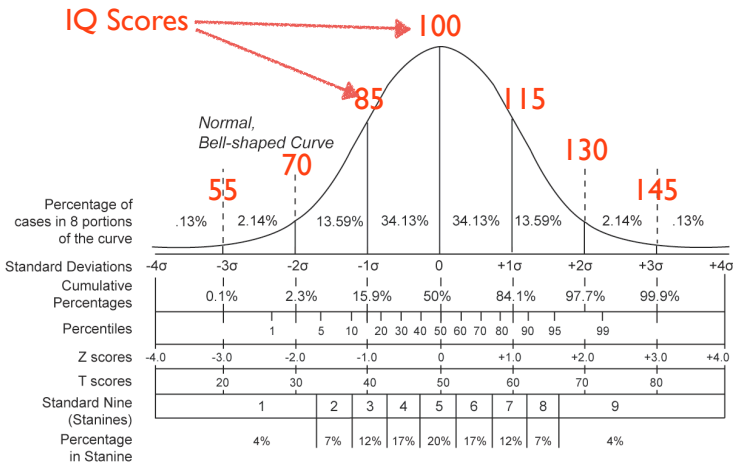
Category (Animal) *T* score = 10.450 + (3.558 x Animal Scaled Score) – (1.048 x Education) + (0.301 x Age) + (8.476 x Race)

Education = years of education successfully completed.

Age = actual age (if age is 20-34 years, age is coded as 34 years).

Race: Caucasian = 0, African American = 1.

Standard Scores



Describing Performance (WAIS-4 Terminology)

Classification	IQ Score	Z Score	T-Score	Percentile (%ile)
Very superior	above 130	above 2.0	70 and above	98% and above
Superior	120-129	1.3 to 2.0	63-39	91% - 97%
High average	110-119	0.6 to 1.3	56-62	75% - 90%
Average	90-109	-0.6 to +0.6	44-55	25% - 74%
Low average	80-89	-0.6 to -1.3	43-37	9% - 24%
Borderline	70-79	-1.3 to -2.0	36-30	2% - 8%
Extremely low	69 and below	below -2.0	29 and below	below 2%