

PSYC2101 – Assessment, Personality and Psychopathology

LECTURE 1 – INTRODUCTION

Food for thought

- Can we quantify psychological constructs (e.g., intelligence, personality traits, emotional experience)?
- How do we assess these constructs?

What is Abnormal?

- When does a personality trait become a mental disorder?
 - Shyness vs. social phobia
 - Statistical deviance
 - Are all mental disorders simply extremes of normal behaviour?
 - If so, what is the “normal” equivalent of this belief?
 - *“The thoughts of Eamonn Andrews come into my mind. There are no other thoughts there, only his. . .He treats my mind like a screen and flashes his thoughts onto it like you flash a picture”*. -First-person account from a patient with schizophrenia (from Mellor, 1970)
 - Does developmental stage matter?
 - Is what is abnormal socially determined?
 - Witches have an extra internal organ that flies away at night and poisons people’s blood (Boyer 2002)
 - This is a common belief amongst the Fang people of Cameroon; if you don’t believe it, you’re weird
 - Is psychopathology entirely social determined?
 - What role does culture play in abnormal behaviour?
 - Is all suffering pathological?
 - E.g., relationship breakup, bereavement, posttraumatic stress
 - reflect suffering but are normative and are not usually considered disorders
 - Do all mental disorders involve suffering?
 - E.g., Mania, aggressiveness & conduct disorders
 - Don’t necessarily involve suffering on the part of the individual, but are usually considered disorders
-

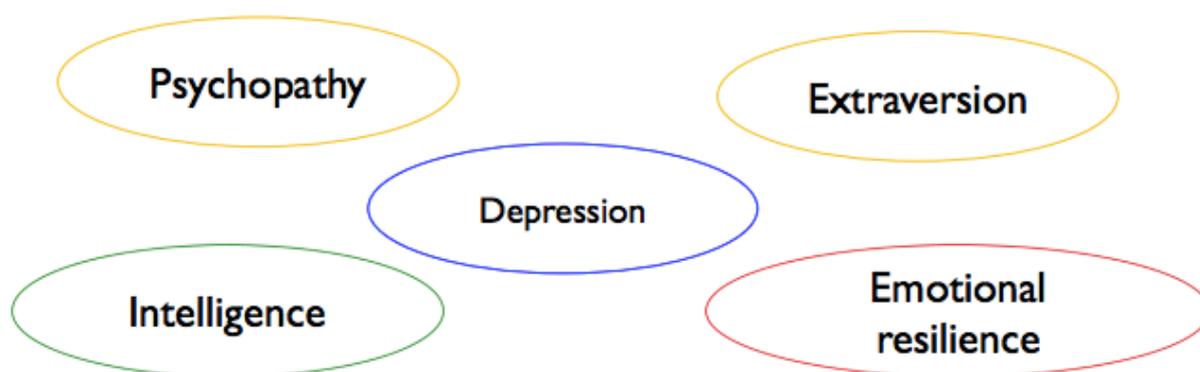
LECTURE 2 – RELIABILITY AND VALIDITY

Psychometrics

- The branch of psychology dealing with the design, administration, and interpretation of quantitative tests for the measurement of psychological variables
- Psyche = mind, metric = measure
- Psychometrics involves the measurement of mental constructs (e.g., intelligence, aptitude, personality, etc.) as opposed to physical constructs (e.g., height, weight, antibody levels, etc.)

What is a mental construct?

- Theoretical, intangible quality or trait
 - Think: conceptual variable
- A psychological test aims to *measure* mental constructs by inferring an individual's true score on the construct from their observed test score.



To bear in mind...

Psychometrics makes 2 fundamental assumptions about the mental constructs it attempts to measure:

1. They exist
2. They can be measured and quantified

What do you see?



- Depression Anxiety Stress Scale (DASS) (Lovibond and Lovibond, 1995)

These are two scales of depressiveness – Which one is better? DASS is objective and is quantifiable. There is no way to say which one is better – as it is an empirical question. That

is it a question that can only be measured by data. The better scale is the one that has better reliability and validity.

What makes a good psychometric test?

- **Reliability:** Consistency of measurement
 - Does the test give consistent results when it's given at different times, by different raters, etc.
- **Validity:** Accuracy of measurement. A scale that is valid will give objectively correct measurements.
 - Are the measurements objectively correct?
 - Does the test measure what it purports to measure?

A good scale has to be reliable! But reliability is not sufficient for it to be a good test. A good test also needs to be valid.

Why is it important?

Psychometric testing infuses all aspects of life.

- E.g. career assessment tests done in school – a bunch of tests which tell you which career you would be suited for
- Carrie Buck – American woman in the in her early 20s. she sat an intelligence test and on the basis of this test was feeble minded. She got a low score. On the basis of this score she was forcibly sterilised and got her daughter removed from her.
- Psychological profiling – CAARMS identifies individuals at risk of psychosis. It determines whether they go into treatment programs take medicine and stuff before they become psychotic.

Measuring Mental Constructs

- A psychological test aims to measure mental constructs by inferring an individual's true score on the construct from their observed test score.
- In ideal measurement, a person's OBSERVED score would equal their TRUE score
- However, there is always some measurement error

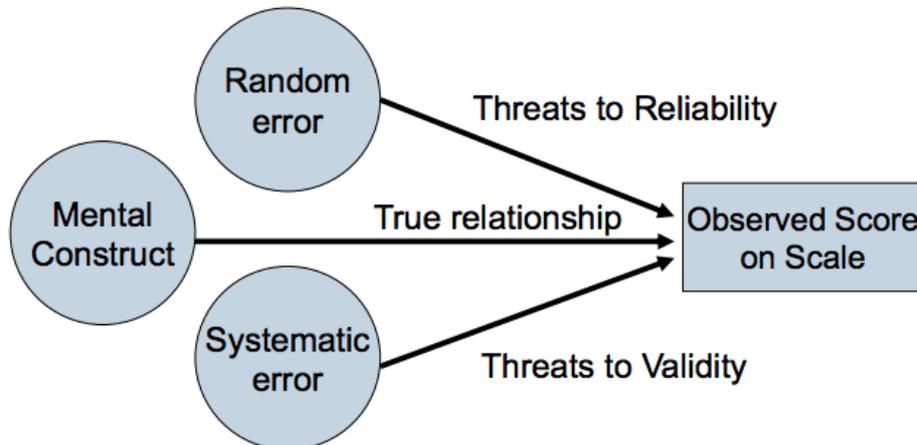


Errors in Measurement

- **Random Error:** unpredictable influences that vary from measurement to measurement.
 - E.g. participant didn't have their morning coffee or has the flu, examiner didn't give clear instructions etc.
 - Influences go in both directions — sometimes increasing a score and sometimes decreasing it
- **Systematic error:** biases that influence scores in a similar way across multiple measurements
 - E.g. using vocabulary as a measure of IQ in ESL participants

- Influences generally go in one direction — consistently increasing or decreasing a score

Random error vs Systematic Error



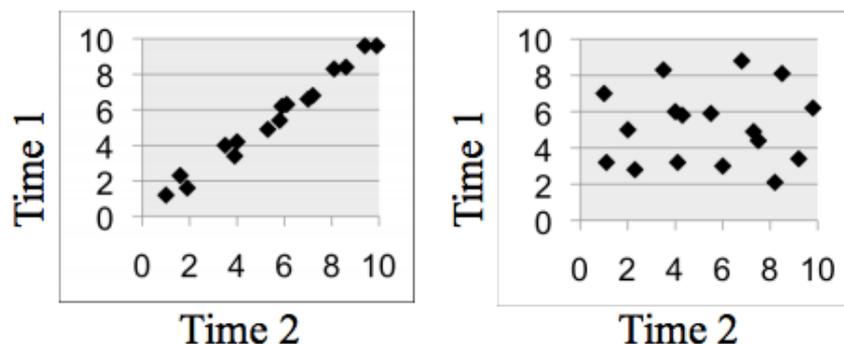
TYPES OF RELIABILITY

Reliability: consistency of measurement. What do we mean by 'consistent'?

- **Temporal Stability:** are the results of the test stable over time?
 - Test-retest reliability
- **Inter-Rater Stability:** are the results of the test stable across raters?
 - Inter-rater reliability
- **Internal Consistency:** Are there correlations between test items that are supposedly measuring the same construct?
 - Split-half reliability
 - Cronbach's alpha

Test-Retest Reliability: Temporal Stability

- Will taking a test again yield the same results?
- Administer a test once. Wait. Administer again to the same group. Then, correlate each set of scores.
- Drawbacks to assessing temporal stability:
 - Practice effects (e.g., memory), Fatigue, Cost
- Temporal stable assessment is good.



Inter-Rater Reliability

- Do two different raters yield the same results?

- Often used with subjectively-scored measures
- Test is scored by two or more raters. Sets of scores are correlated with one another.

Internal Consistency

- Extent to which test items that propose to measure the same construct actually show similar scores
- High correlation between items in a scale

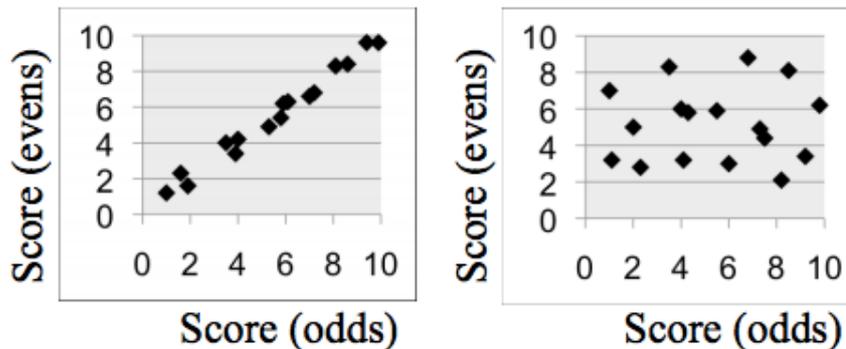
Tom's Bicycle Attitude Inventory (T-BAI) (5 point scale: 1=SD, 2=D, 3=N, 4=A, 5=SA)	
1.	"I like to ride bicycles"
2.	"I've enjoyed riding bicycles in the past"
3.	"I hate bicycles" (Reverse Scored)
4.	"I love watching the Tour de France"
5.	"Lycra is my favourite fabric"
6.	"I think Stevie Wonder is a musical genius"

Item	1	2	3	4	5	6
1	1					
2	0.96	1				
3	0.76	0.88	1			
4	0.7	0.74	0.89	1		
5	0.8	0.58	0.69	0.87	1	
6	0.15	-0.1	0.05	0.21	-0.15	1

Lower internal consistency

Split-half reliability

- Divide the scale in half (e.g., odd vs even items) then correlate the two halves

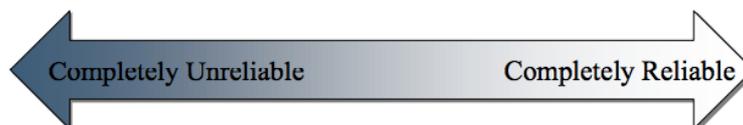


- Cronbach's Alpha (α)
 - Equivalent to calculating the average of all possible split-half correlations

RELIABILITY

- Consistency of measurement
- Extent of freedom from random error

View reliability as a continuum, not a category.



Improving Reliability

- Control test administration
 - Standardized settings and instructions

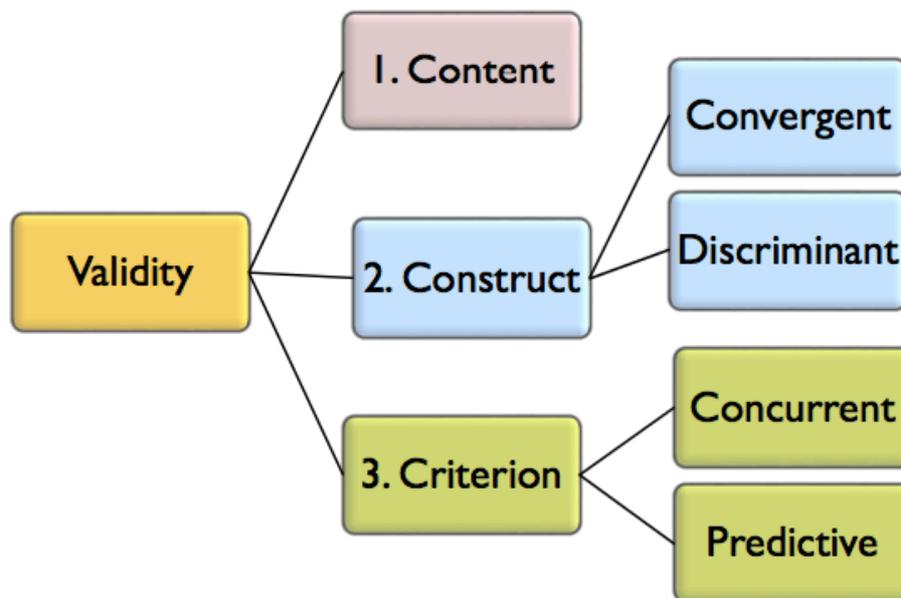
- Increase the number of items
- Use a discriminability analysis
 - identify and remove items that negatively impact your scale's reliability

Let's assume we have a measure that is reliable, i.e., consistent. How do we know that it is actually measuring what it purports to measure?

VALIDITY

- Accuracy of measurement
 - Is the measuring what it claims to be measuring?
 - Are the conclusions drawn from measurement well-founded?
 - How free from systematic error is a particular measure?

TYPES OF VALIDITY



Content Validity

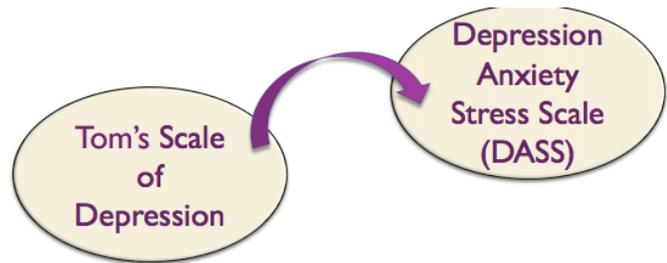
- The degree to which the items adequately sample the mental construct in question
- E.g., Imagine 2 measures of general intelligence (IQ)
 - Scale A only includes a vocabulary measure \diamond poor content validity
 - Scale B Includes measures of vocabulary, arithmetic, working memory, matrix reasoning, mental rotation, etc \diamond better content validity

Construct Validity

- A scale is said to have high construct validity if:
 1. It correlates with other tests purporting to measure the same construct (and hence with which it should theoretically correlate) \rightarrow Convergent Validity
 2. It does not correlate with other tests that do not purport to measure the same construct \rightarrow Discriminant Validity

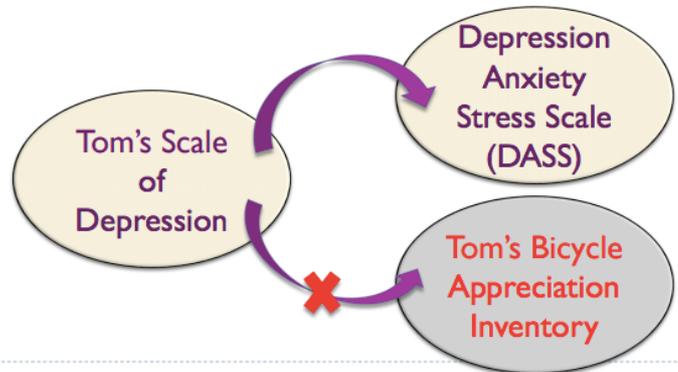
Convergent Validity

Demonstrated when a measure correlates highly with other measures with which it theoretically should correlate.



Discriminant Validity

Demonstrated when a measure does not correlate highly with other measures with which it theoretically should not correlate

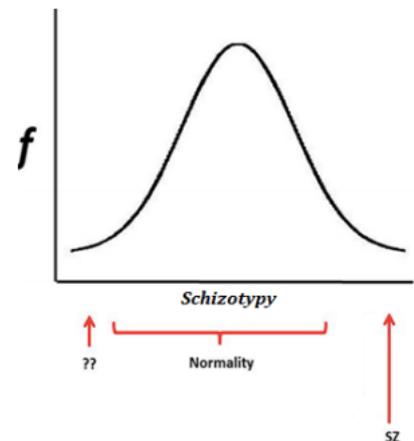


Criterion Validity

- Degree to which the test is able to predict present or future performance on some real-world outcome (i.e., the criterion)
- **Concurrent validity:** Does test score accurately predict present scores on the criterion?
 - I.e., the test and the criterion are assessed at the same time
 - E.g. Predicting who has clinically-diagnosed major depression based on their score on the DASS
- **Predictive validity:** Does test score accurately predict future scores on the criterion?
 - Predicting future transition to psychosis on the basis of an individual's CAARMS score

An example: measuring 'schizotypy'

- Schizotypy is a hypothesized mental construct describing a continuum of characteristics and experiences related to psychosis
- High SZPY is associated with pseudohallucinations, magical thinking, suspicion



Schizotypal Personality Questionnaire (SPQ) (Raine, 1991)

- ▶ Reliability measures
 - ▶ Test-retest: $r = 0.82$
 - ▶ Cronbach's alpha = 0.91
- ▶ Validity measures
 - ▶ Construct Validity (Convergent Validity): $r = 0.81$ with Schizotypy Traits Questionnaire (STA)
 - ▶ Criterion Validity:
 - ▶ Top 10% of scorers on SPQ → 50% met criteria for Schizotypal Personality Disorder
 - ▶ Bottom 10% on SPQ → 0% met criteria for Schizotypal Personality Disorder

Rule of thumb:
r-values above 0.7 - 0.8 are considered 'good'

Rorschach Schizophrenia Index (RSCZI)

▶ Reliability measures

▶ Inter-rater reliability:

$r \sim 0.4$

▶ Validity measures

▶ Construct Validity

(Convergent): $r \sim 0.1$ b/w
RSCZI and Schizophrenia
subscale of the Minnesota
Multiphasic Personality
Inventory (MMPI)

The Bull's-eye

