

## Item Response Theory (alternate model to True Score Theory or Classical Test Theory)

Can explain how the IRT differs in its approach to test construction compared to classical test theory

Classical Test Theory (CTT)	Item Response Theory (IRT)
Classical Test Theory (CTT) focuses on the total score of a scale or subscale. In this way, CTT is often thought of as the theory of total scores. The key equation for CTT is: test score = true score + error.	Item Response Theory (IRT) focuses on the relationship between observed responses to items and overall construct or latent dimension underlying the test/being measured i.e. assumes that there is a relationship between responses to items and the underlying or latent dimension being assessed by the scale (e.g. relationship between a person's ability and the likelihood that they will answer an item correctly – in this way, provides a way to model the probability that a person with ability will be able to perform at the level of Y).
In CTT, estimates of test and item parameters are dependent on the sample from which they were calculated i.e. estimates from traditional item analysis rely on the representativeness of the sample to generalise from one sample to the next.	Estimates from IRT are <i>sample invariant</i> in that they do not depend on the sample from which they were drawn.
Scoring in CTT is usually simpler (e.g. addition of items).	Scoring in IRT is more complex, requiring computer time.
	IRT has stronger assumptions and when these are met, provides stronger findings: <ul style="list-style-type: none"> <li>• Unidimensionality: items assess a single construct</li> <li>• Local dependence: items are not too similar</li> </ul>

## Assessment of Intelligence

*Can define intelligence*

In general, IQ is a measure of cognitive ability.

Sternberg (1981):

Verbal intelligence	Problem solving ability	Practical intelligence
<ul style="list-style-type: none"> <li>- General learning and comprehension</li> <li>- Good vocabulary</li> <li>- Reads with high comprehension</li> <li>- Intellectual curiosity</li> </ul>	<ul style="list-style-type: none"> <li>- Abstract thinking or reasoning</li> <li>- Can apply knowledge to tasks at hand</li> <li>- Plans ahead</li> <li>- Solves problem well</li> </ul>	<ul style="list-style-type: none"> <li>- Real-world adaptive behaviours</li> <li>- Sizes-up situations well</li> <li>- Determines how to achieve goals</li> <li>- Displays awareness of the world</li> </ul>

*Can name the theories of intelligence*

\*'g' = general theory of intelligence

Lumpers – intelligence is only 'g'	<b>Hierarchical models</b>	
<p>Spearman (1927) – intelligence is governed by a single mental ability i.e. if people are good at one thing, they are good at other things</p> <p>Implication: give them one test of intelligence</p>	<p>Vernon (1950) – there are two abilities that feed into 'g': verbal educational (v:ed) and kinetic mechanical (k:m) → beneath those are specific individual tests which feed into these abilities (e.g. reading tests, spelling tests and math tests feed into v:ed)</p> <p>Thurstone (1938) – there are seven primary abilities that feed into 'g': verbal comprehension, word fluency, number</p>	<p>Splitters – intelligence is separate factors</p> <p>Guildford (1967) – intelligence is governed by 120 separate and independent abilities</p> <p>Implication: need to measure each and every single one of those abilities i.e. there are tasks to target separate forms of intelligence</p>

## Assessment in Clinical Practice

---

(Meyer et al., 2001):

- Testing: a particular scale is administered to obtain a specific score and a descriptive meaning can be applied to the score on the basis of normative, nomothetic findings.
- Assessment: concerned with the clinician who takes a variety of test scores, generally obtained from multiple test methods, and consider data in the context of history, referral information, and observed behaviour to understand the person being evaluated, to answer the referral questions, and then to communicate findings to the patient, his or her significant others and referral sources.
- Why assess?
  - Describe current functioning
  - Confirm, refute or modify impressions formed by clinicians
  - Identify therapeutic needs, highlight issues likely to arise in treatment, recommend forms of interventions and offer guidance about likely outcomes
  - Aid in different diagnosis
  - Monitor treatment over time to evaluate the success of interventions
  - Manage risk (untoward treatment reactions, potential legal liabilities)
  - Provide skilled, empathetic assessment feedback as a therapeutic intervention in itself
- Why use standardised tests?
  - (Dahlstrom, 1993): the samples of behaviour that psychologists collect in the brief time that an hourglass takes to empty have been shown to reveal basic aspects of ability, personality and temperament that are operative over long spans of an individual's life
  - Clinicians are unreliable judges
    - Why?
      - Errors in gathering data:
        - Tendency to see patterns where none exist
        - Tendency to seek confirmatory evidence
        - Use of preconceived biases
      - Errors in synthesizing data:
        - (Tversky & Kahneman, 1974): Heuristics in clinical judgment
          - × Representativeness
          - × Availability
          - × Anchoring
          - × Affect (Garb, 2005)
          - × Prototypes
    - (Binet & Simon, 1907): we have made a methodological comparison between the admission certificates filled out for the same children within only a few days... we think we may say without exaggeration that they looked as if they had been drawn by a chance out of a sack