

# Psychometrics to Support

Assessment Design

* Construct Definition
  + A simple form: More or less, high to low
* Item Development
  + Realizations of the construct
* Outcome Space
  + Aspect of response we value – how to score
* Measurement Model
  + How we related scores to constructs

Construct

*Causality*

Item Responses

Measurement Model

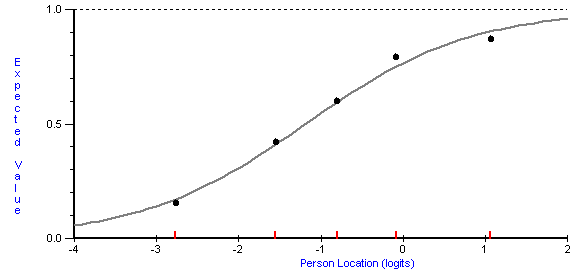
Outcome

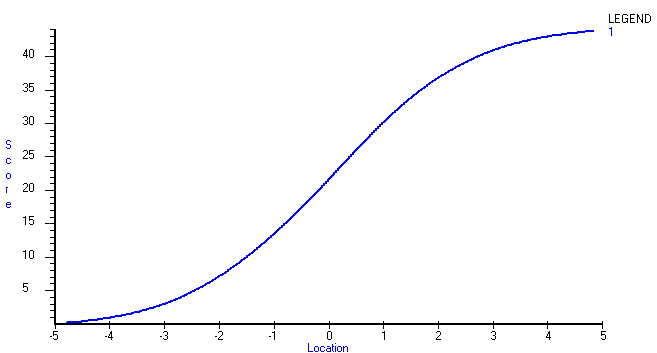
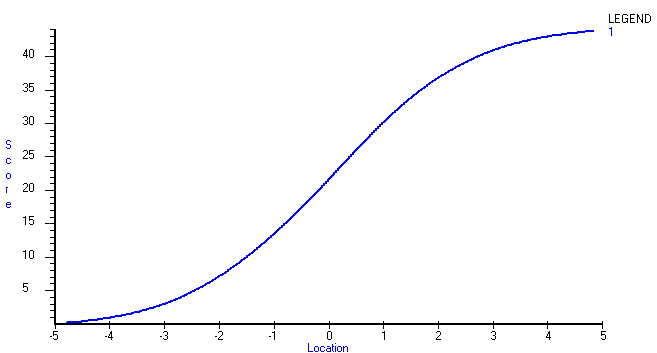
Space

*Inferences*

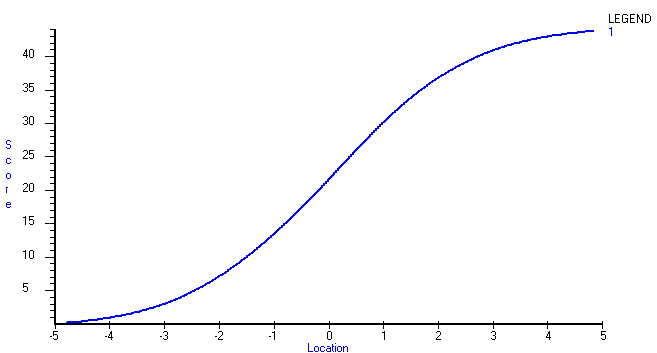
Source: Mark Wilson, 2005

* Most IRT models are based on a paradigm that identifies a model which explains variation in the data – to find a model that best characterizes the data
* Rasch is an approach that is based on the paradigm of constructing a measure which can characterize a construct on a linear scale – such that the total score fully characterizes a person on a given construct
* Rasch models provide a basis and justification for obtaining person locations on a continuum from total scores on assessments.
* Although it is not uncommon to treat total scores directly as measurements, they are actually counts of discrete observations rather than measurements.
* Each observation represents the observable outcome of a **comparison** between a person and item.
* Such outcomes are directly analogous to the observation of the rotation of a balance scale in one direction or another.
* This observation would indicate that one or other object has a greater mass, but counts of such observations cannot be treated directly as measurements.



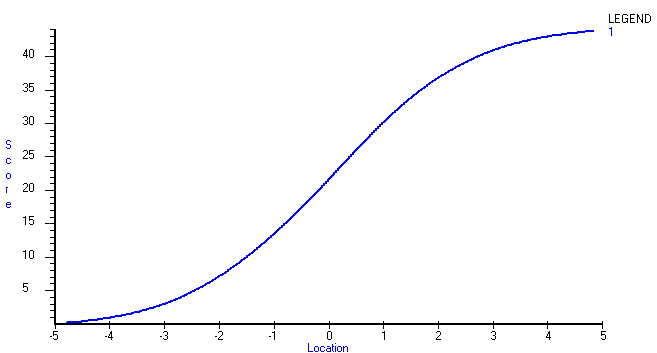


[4 Points on the Raw Score Scale](http://upload.wikimedia.org/wikipedia/en/6/6b/TCC.PNG)



[4 Points on the Raw Score Scale](http://upload.wikimedia.org/wikipedia/en/6/6b/TCC.PNG)

[0.5 on the Rasch Scale](http://upload.wikimedia.org/wikipedia/en/6/6b/TCC.PNG)



[4 Points on the Raw Score Scale](http://upload.wikimedia.org/wikipedia/en/6/6b/TCC.PNG)

[1.0 Point on the Rasch Scale](http://upload.wikimedia.org/wikipedia/en/6/6b/TCC.PNG)

* Numbers themselves do not mean much.
  + Is 10 meters a short distance? Long distance?

We need context to bring meaning to the measure: 10 meters.

However, 10 meters should always be 10 meters, no

matter who takes the measure or how it is taken.

* Is an item with a p-value of .90 easy or difficult?

… 90% passed the item

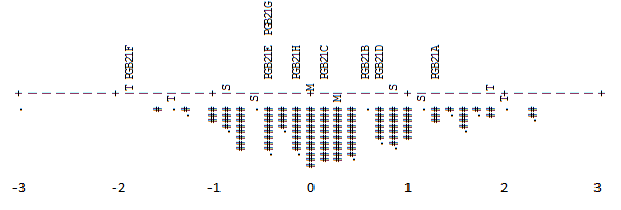
* Is a person with a score of 5 out of 50 items

low in ability?

… correctly answered 10% of the items

* Person-free item difficult
  + Locates the items on the ability continuum
* Item-free person ability
  + Locates the person on the ability continuum
* Places items and persons on the same scale –

the ITEM MAP



1. Explains the construct; interpretation guide
2. Enables design of items that will lead children to give responses that inform important levels of the construct map; identify relevant item features
3. Provides criterion to analyze responses regarding

degree of consistency with construct map

1. Item selection or retention should be based on

informed professional judgment