Multitrait Scaling and IRT: Part I

Ron D. Hays, Ph.D.
(drhays@ucla.edu)

http://www.gim.med.ucla.edu/FacultyPages/Hays/

Questionnaire Design and Testing Workshop

2nd Floor Conference Room

October 16, 2009 (3-5pm)

Multitrait Scaling Analysis

- Internal consistency reliability
  - Item convergence

Hypothetical Item-Scale Correlations

	Trait #1	Trait #2	Trait #3
Item #1	0.80*		
Item #2	0.80*		
Item #3	0.80*		
Item #4		0.80*	
Item #5		0.80*	
Item #6		0.80*	
Item #7			0.80*
Item #8			0.80*
Item #9			0.80*

*Item-scale correlation, corrected for overlap.

Measurement Error

observed = true score + systematic error + random error

(bias)

Cronbach's Alpha

01	55
02	45
03	42
04	35
05	22

Source	df	SS	MS
Respondents (BMS)	4	11.6	2.9
Items (JMS)	1	0.1	0.1
Resp. x Items (EMS)	4	4.4	1.1
Total	9	16.1	

$$\text{Alpha} = \frac{2.9 - 1.1}{2.9} = \frac{1.8}{2.9} = \boxed{0.62}$$

Alpha for Different Numbers of Items and Homogeneity

Number of Items (k)	Average Inter-item Correlation ($\bar{r}$)					
	.0	.2	.4	.6	.8	1.0
2	.000	.333	.572	.750	.889	1.000
4	.000	.500	.727	.857	.941	1.000
6	.000	.600	.800	.900	.960	1.000
8	.000	.666	.842	.924	.970	1.000

$$\text{Alpha}_{st} = \frac{k * \bar{r}}{1 + (k - 1) * \bar{r}}$$

Spearman-Brown Prophecy Formula

$$\text{alpha}_{y} = \left(\frac{N \cdot \text{alpha}_{x}}{1 + (N - 1) * \text{alpha}_{x}} \right)$$

N = how much longer scale y is than scale x

Example Spearman-Brown Calculations

MHI-18

$$\frac{18/32\ (0.98)}{(1+(18/32\ -1)*0.98}$$

= 0.55125/0.57125 = 0.96

Number of Items and Reliability for Three Versions of the Mental Health Inventory (MHI)

Measure	Number of Items	Completion time (min.)	Reliability
MHI-32	32	5-8	.98
MHI-18	18	3-5	.96
MHI-5	5	1 or less	.90

From McHorney et al. 1992

Reliability Minimum Standards

0.70 or above (for group comparisons)

0.90 or higher (for individual assessment)

- SEM = SD $(1 - \text{reliability})^{1/2}$

Intraclass Correlation and Reliability

Model	Reliability	Intraclass Correlation
One-Way	$\frac{MS_{BMS} - MS_{WMS}}{MS_{BMS}}$	$\frac{MS_{BMS} - MS_{WMS}}{MS_{BMS} + (K-1)MS_{WMS}}$
Two-Way Fixed	$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS}}$	$\frac{MS_{BMS} - MS_{EMS}}{MS_{EMS} + (K-1)MS_{EMS}}$
Two-Way Random	$\frac{N(MS_{BMS} - MS_{EMS})}{NMS_{BMS} + MS_{JMS} - MS_{EMS}}$	$\frac{MS_{BMS} - MS_{EMS}}{MS_{BMS} + (K-1)MS_{EMS} + K(MS_{JMS} - MS_{EMS})/N}$

Multitrait Scaling Analysis

- Internal consistency reliability
  - Item convergence
- Item discrimination

Hypothetical Multitrait/Multi-Item Correlation Matrix

	Trait #1	Trait #2	Trait #3
Item #1	0.80*	0.20	0.20
Item #2	0.80*	0.20	0.20
Item #3	0.80*	0.20	0.20
Item #4	0.20	0.80*	0.20
Item #5	0.20	0.80*	0.20
Item #6	0.20	0.80*	0.20
Item #7	0.20	0.20	0.80*
Item #8	0.20	0.20	0.80*
Item #9	0.20	0.20	0.80*

*Item-scale correlation, corrected for overlap.

Multitrait/Multi-Item Correlation Matrix for Patient Satisfaction Ratings

	Technical	Interpersonal	Communication	Financial
Technical				
1	0.66*	0.63†	0.67†	0.28
2	0.55*	0.54†	0.50†	0.25
3	0.48*	0.41	0.44†	0.26
4	0.59*	0.53	0.56†	0.26
5	0.55*	0.60†	0.56†	0.16
6	0.59*	0.58†	0.57†	0.23
Interpersonal				
1	0.58	0.68*	0.63†	0.24
2	0.59†	0.58*	0.61†	0.18
3	0.62†	0.65*	0.67†	0.19
4	0.53†	0.57*	0.60†	0.32
5	0.54	0.62*	0.58†	0.18
6	0.48†	0.48*	0.46†	0.24

Note – Standard error of correlation is 0.03. Technical = satisfaction with technical quality. Interpersonal = satisfaction with the interpersonal aspects. Communication = satisfaction with communication. Financial = satisfaction with financial arrangements. *Item-scale correlations for hypothesized scales (corrected for item overlap). †Correlation within two standard errors of the correlation of the item with its hypothesized scale.

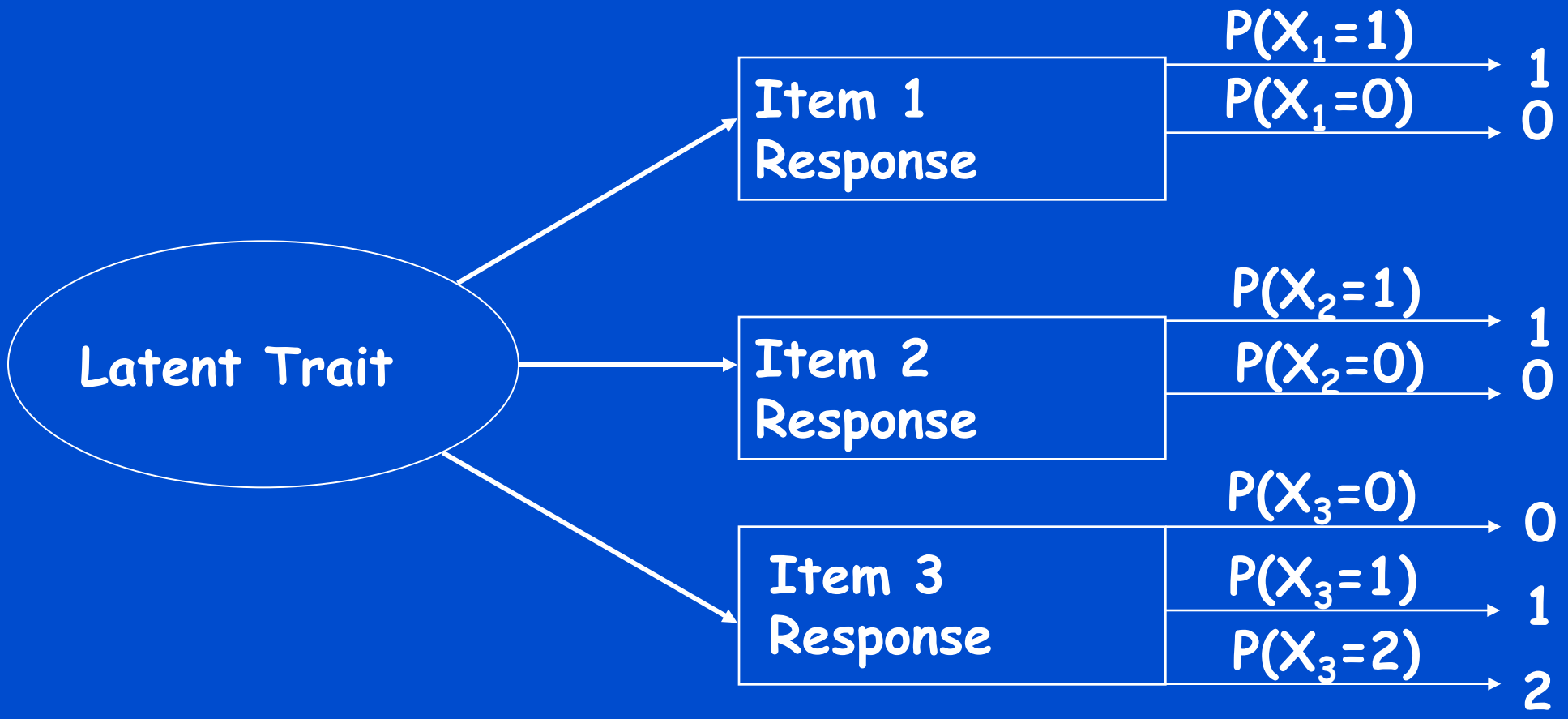
Confirmatory Factor Analysis

- Compares observed covariances with covariances generated by hypothesized model
- Statistical and practical tests of fit
- Factor loadings
- Correlations between factors
- Regression coefficients

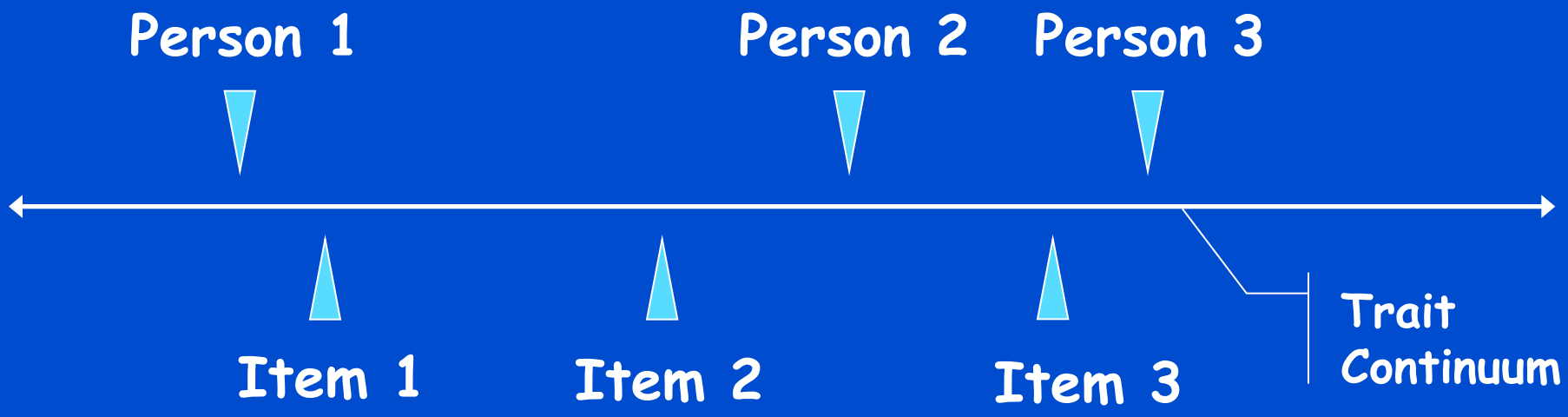
Fit Indices

- Normed fit index: $\dfrac{\chi^2_{null} - \chi^2_{model}}{\chi^2_{null}}$
- Non-normed fit index: $\dfrac{\frac{\chi^2_{null}}{df_{null}} - \frac{\chi^2_{model}}{df_{model}}}{\left[\frac{\chi^2_{null}}{df_{null}} - 1\right]}$
- Comparative fit index: $1 - \left[\dfrac{\chi^2_{model} - df_{model}}{\chi^2_{null} - df_{null}}\right]$

Latent Trait and Item Responses



Item Responses and Trait Levels



Item Response Theory (IRT)

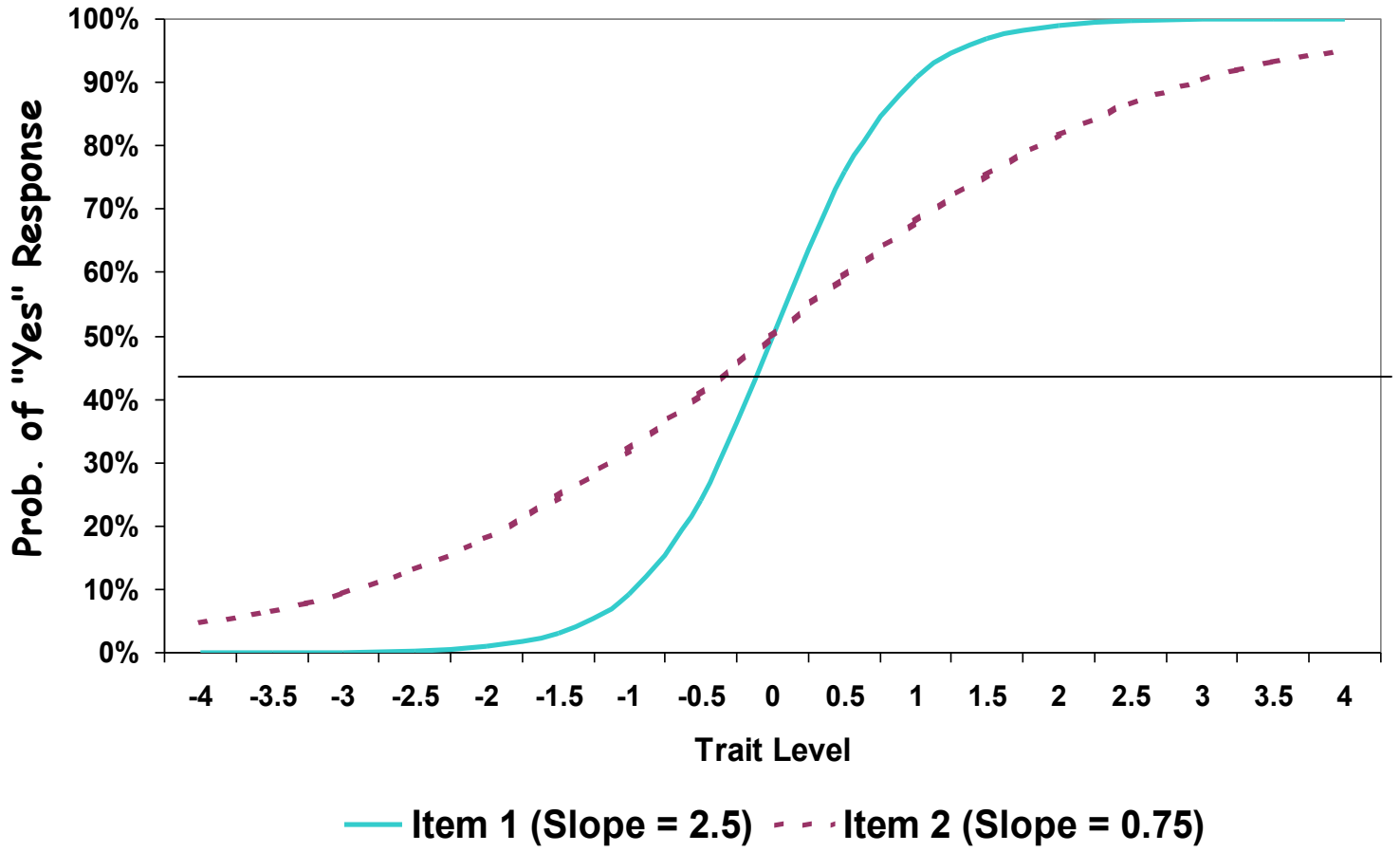
IRT models the relationship between a person's response Y_i to the question (i) and his or her level of the latent construct θ being measured by positing

$$\Pr(Y_i \geq k) = \frac{1}{1 + \exp(-a_i\theta + b_{ik})}$$

b_{ik} estimates how difficult it is for the item (i) to have a score of k or more and the discrimination parameter a_i estimates the discriminatory power of the item.

If for one group versus another at the same level θ we observe systematically different probabilities of scoring k or above then we will say that item i displays DIF

Item Characteristic Curves (2-Parameter Model)



PROMIS Assessment Center

http://www.nihpromis.org/

http://www.assessmentcenter.net/ac1/