IRT Model for Item Clusters

Jon Cohen

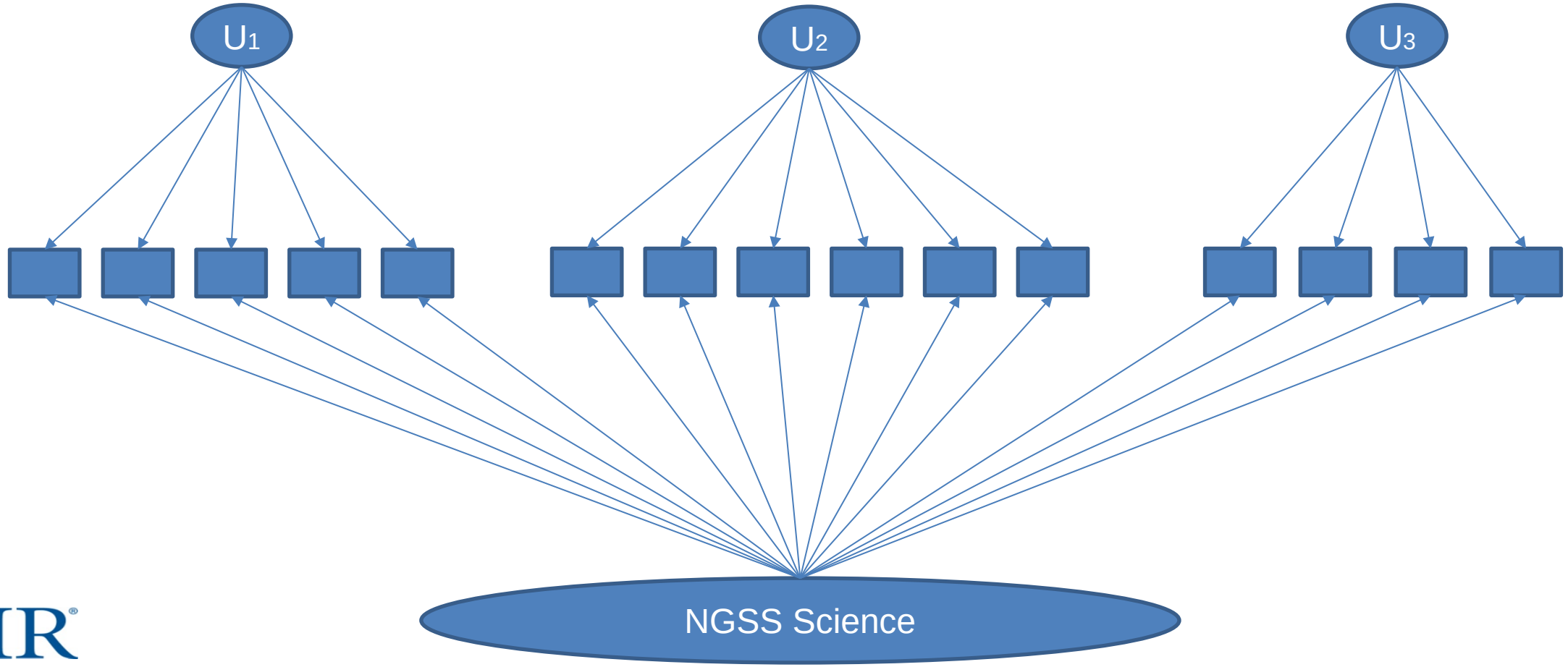
AIR Assessment

A division of the American Institutes for Research

Analyzing item clusters: When does unidimensionality fail?

- Under a unidimensional model,
  - In a nice symmetric world…
    - each cluster has the same number of "points" or *scoring assertions*
    - the intra-cluster correlation is about the same across clusters
    - …point estimates are a bit biased and not consistent (bias does not decrease with the number of items)
    - …standard errors are understated
  - In an asymmetric (real) world…
    - …point estimates are more biased and inconsistent
    - …standard errors are understated
- AIR is recommending that our clients use a "bifactor model," which captures and removes the impact of cluster specific variance

Bifactor Model

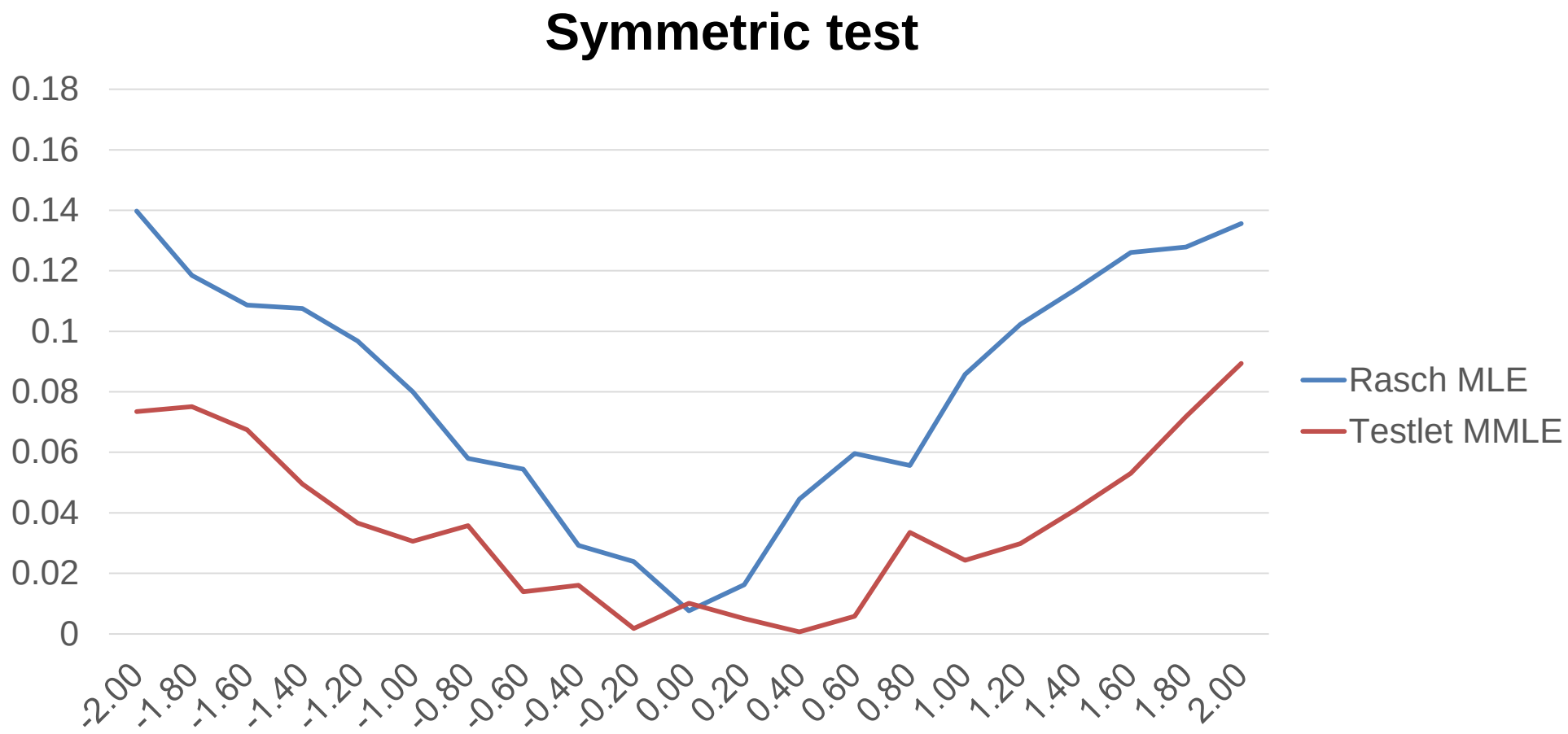


Early empirical results

Percent of systematic variance that is cluster specific (analysis of 13 clusters from Connecticut)

Statistic	Value
Mean	0.39
Standard deviation	0.26
Minimum	0.08
Maximum	0.85

Absolute Bias: ρ=0.5



Statistical inference becomes increasingly problematic with asymmetry

Coverage rates under symmetric assumptions

	Proportion of systematic variance due to cluster-specific factors			
	0.20	0.33	0.44	0.50
1 PL MLE	0.055	0.07	0.085	0.10
Testlet MMLE	0.047	0.045	0.043	0.045

Coverage rates under asymmetric assumptions

	Proportion of systematic variance due to cluster-specific factors			
				0.50
1 PL MLE				0.137
Testlet MMLE				0.047

Implication

- Using the bifactor model, varying the number of points/assertions in each cluster does not bias estimates of the student performance on the main trait.
  - Can examine student performance on each cluster for every meaningful choice they make, rather than forgoing information in the interest of symmetry.
- Points or assertions contribute less to the overall score as the intra-cluster correlation increases
  - Prevents the asymmetry from “tilting” the trait towards performance expectations measured with more points/assertions
  - When the intra-cluster correlation is zero, the cluster is only systematically measuring the main trait, and therefore contributes more information
- Supports a strategy that maximizes *information per minute.*