

AN INVESTIGATION OF THE ALIGNMENT METHOD

FOR DETECTING MEASUREMENT NON- INVARIANCE ACROSS MANY GROUPS WITH DICHOTOMOUS INDICATORS

**Jessica Flake, Erin Strauts, Betsy McCoach, Jane Rogers, Megan Welsh
University of Connecticut**

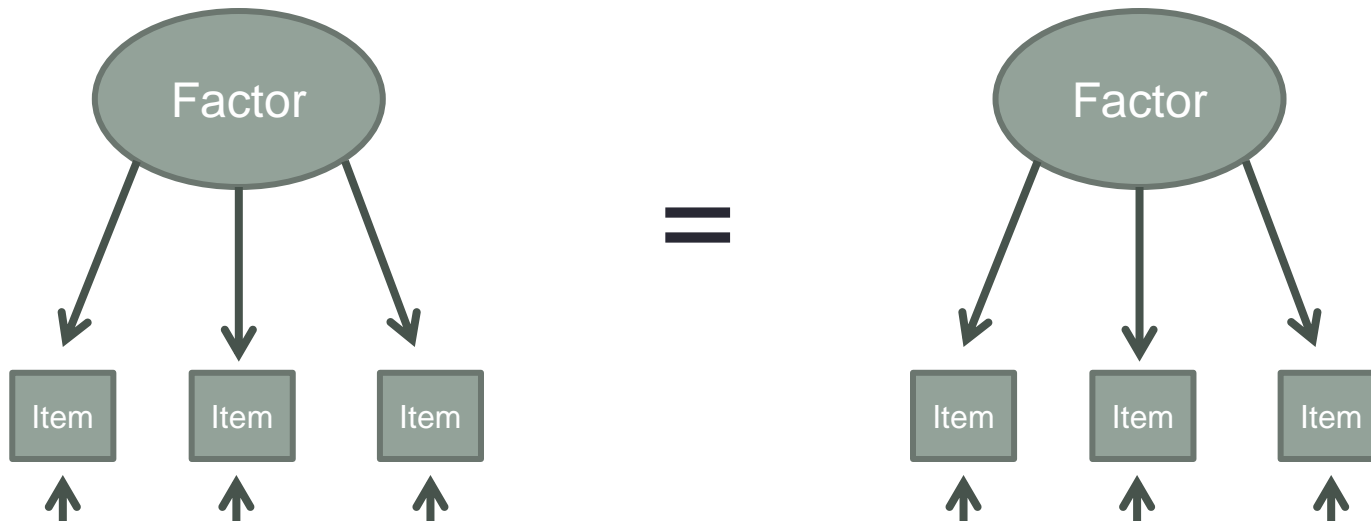
**Modern Modeling Methods Conference
May 2014**

OUTLINE

- **Measurement Invariance**
- **What is Alignment?**
- **Current Study**
- **Implications**

MEASUREMENT INVARIANCE

- measurement properties of a scale do not change, or are **invariant**, across numerous situations
- assumed whenever measurement models are used to answer questions about differences across time or between groups

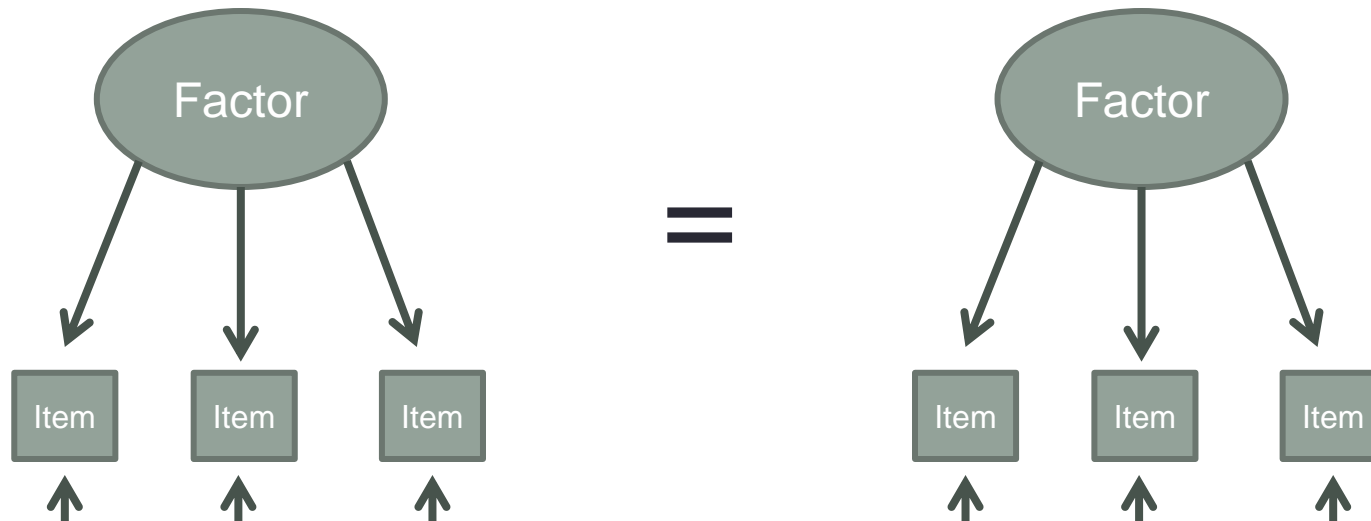


MEASUREMENT INVARIANCE

- **Many ways to assess**
 - Structural Equation Modeling¹
 - Set equality constraints, compare model fit
 - Item Response Theory²
 - Test for group differences with Wald statistics
 - Observed Scores³
 - Regression framework

MEASUREMENT INVARIANCE

- Traditional invariance testing in SEM
 - Configural Invariance
 - Pattern or Metric Invariance
 - Scalar or Strong Invariance
 - Strict Invariance



MEASUREMENT INVARIANCE

- **Easy right?**
- **...item by item**
- **...pairwise group comparison by pairwise group comparison**
- **Invariance becomes harder to ascertain and achieve as items and groups increase**

ALIGNMENT

- **A new, alternative methodology¹**
- **Two parts:**
 - Estimate group specific measurement models
 - Ad-hoc non-invariance testing procedure

ALIGNMENT

- **Part 1: Estimate Group Specific Measurement Models**
 - A configural model is estimated for every group
 - Step 1: estimates all group means at 0, and variance at 1, all loadings and intercepts are free and unequal
 - Perfect fit
 - Group specific means and variances are chosen such that the amount of measurement non-invariance is minimized
 - Component Loss Function

ALIGNMENT

- **Part 2: Ad-hoc invariance testing procedure**
 - Step 1: Determine starting set of invariant groups
 - Series of pairwise tests, if $p > .01$ then groups join the starting set
 - Step 2: Calculate average parameter using starting set
 - Test each group to average
 - If $p > .01$, add that group to average
 - Repeat until stabilization

ALIGNMENT

- **Options**

- Estimation

- Continuous Indicators
 - ML,MLR, MLF, Bayes
 - Dichotomous
 - Bayes, MLR



Brand New in Version
7.2!

- Reference groups

- FREE option: estimates a different factor variance and mean for every group, no reference group
 - FIXED option: first group is the referent

ALIGNMENT

- **Previous Research**
 - Favorable estimates with continuous data¹
 - Initial research with 5 item test
 - N=100- 0 to .20 bias in estimates
 - Greater bias seen with more groups and more non-invariance
 - N=1000- 0 to .02 bias in estimates
 - Across all conditions
 - Recent work with binary data²
 - Application to International Association for the Evaluation of Educational Achievement Civic Test
 - Simulation study in progress
 - No previous research about ad-hoc testing procedure

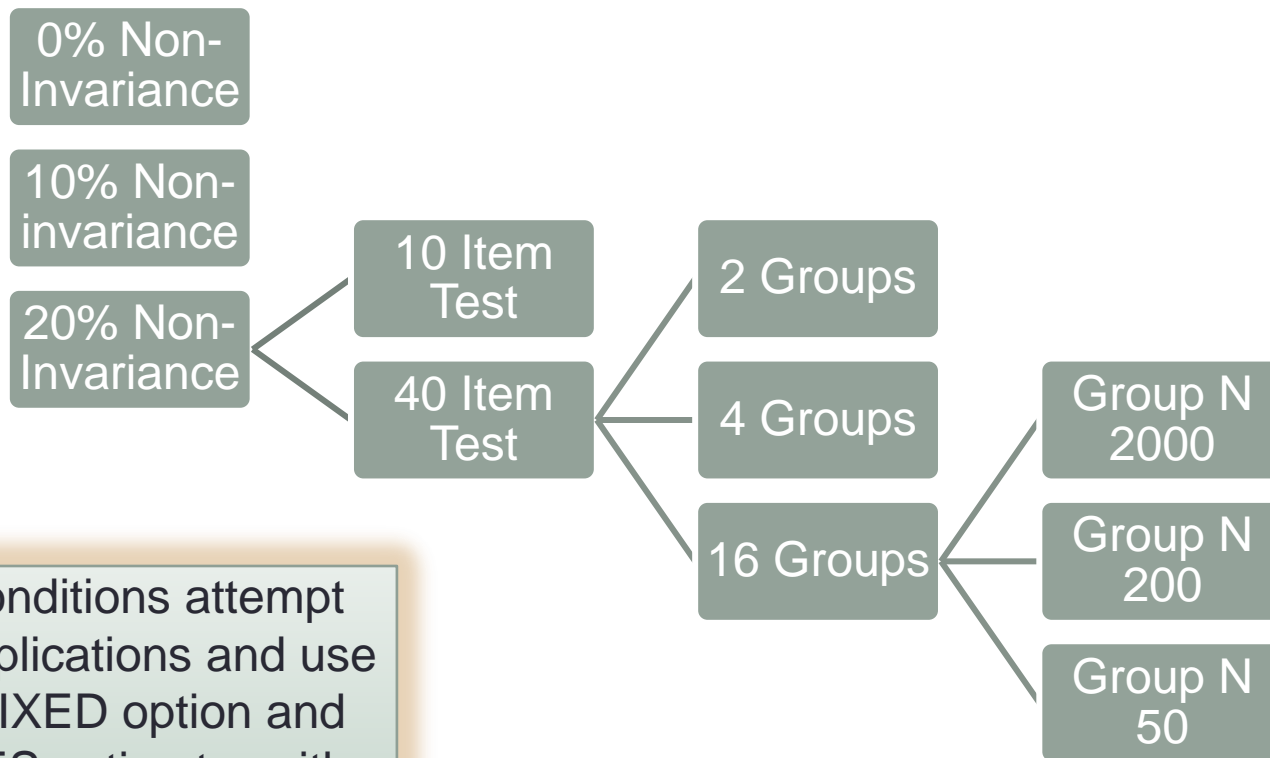


CURRENT STUDY PURPOSE

- **Study 2:**
 - Alignment in the educational context
 - Compare bias and coverage in measurement parameters
 - Investigate accuracy of ad-hoc procedure

CURRENT STUDY CONDITIONS

- Overview



All conditions attempt 100 replications and use the FIXED option and BAYES estimator with non-informative priors

CURRENT STUDY

ITEM AND TEST GENERATION

| 10 Item, Moderate Test | | |
|------------------------|---------------------|-----------|
| | Complete Invariance | |
| | Loading | Threshold |
| 1 | 1.86 | -0.55 |
| 2 | 1.45 | 1.99 |
| 3 | 1.41 | -1.18 |
| 4 | 1.75 | 0.97 |
| 5 | 1.77 | 1.83 |
| 6 | 1.5 | -0.57 |
| 7 | 1.6 | 0.02 |
| 8 | 1.41 | -0.92 |
| 9 | 1.68 | 1.05 |
| 10 | 1.43 | -0.02 |

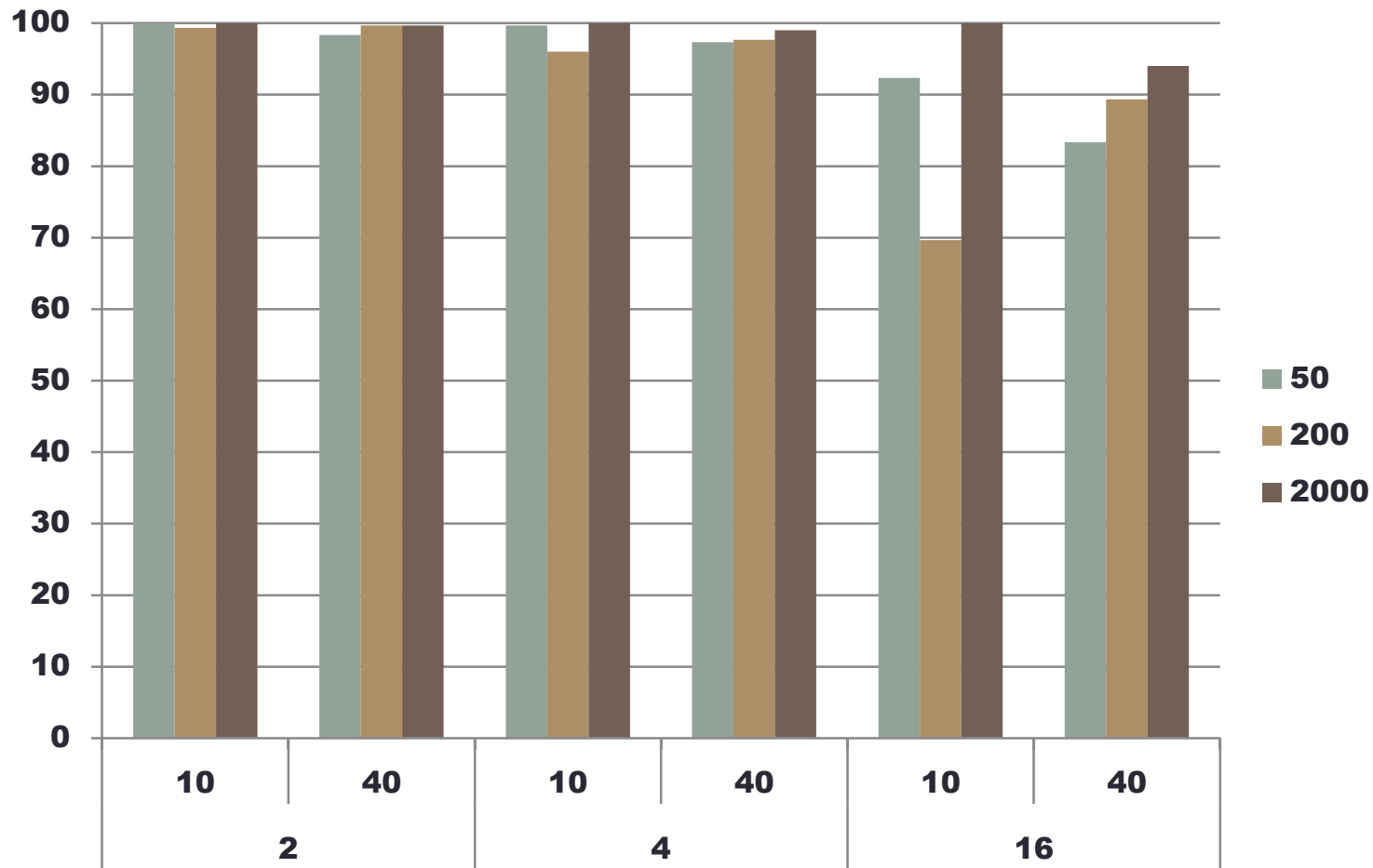
CURRENT STUDY CONDITIONS

- Group specification
 - Many groups, but only 2 group types
 - Odd group always deviates

| 10 Item, Moderate Test | | | | | | |
|------------------------|---------------------|-----------|--------------------|-----------|--------------------|-----------|
| | Complete Invariance | | 10% Non-invariance | | 20% Non-invariance | |
| | Loading | Threshold | Loading | Threshold | Loading | Threshold |
| 1 | 1.86 | -0.55 | 1.86 | -0.55 | 1.86 | -0.55 |
| 2 | 1.45 | 1.99 | 1.45 | 1.99 | 0.95 | 1.99 |
| 3 | 1.41 | -1.18 | 1.41 | -1.18 | 1.41 | -1.18 |
| 4 | 1.75 | 0.97 | 1.75 | 0.97 | 1.75 | 0.97 |
| 5 | 1.77 | 1.83 | 1.77 | 1.83 | 1.77 | 1.33 |
| 6 | 1.5 | -0.57 | 2 | -0.57 | 2 | -0.57 |
| 7 | 1.6 | 0.02 | 1.6 | 0.02 | 1.6 | 0.02 |
| 8 | 1.41 | -0.92 | 1.41 | -0.92 | 1.41 | -0.92 |
| 9 | 1.68 | 1.05 | 1.68 | 1.05 | 1.68 | 1.05 |
| 10 | 1.43 | -0.02 | 1.43 | 0.48 | 1.43 | 0.48 |

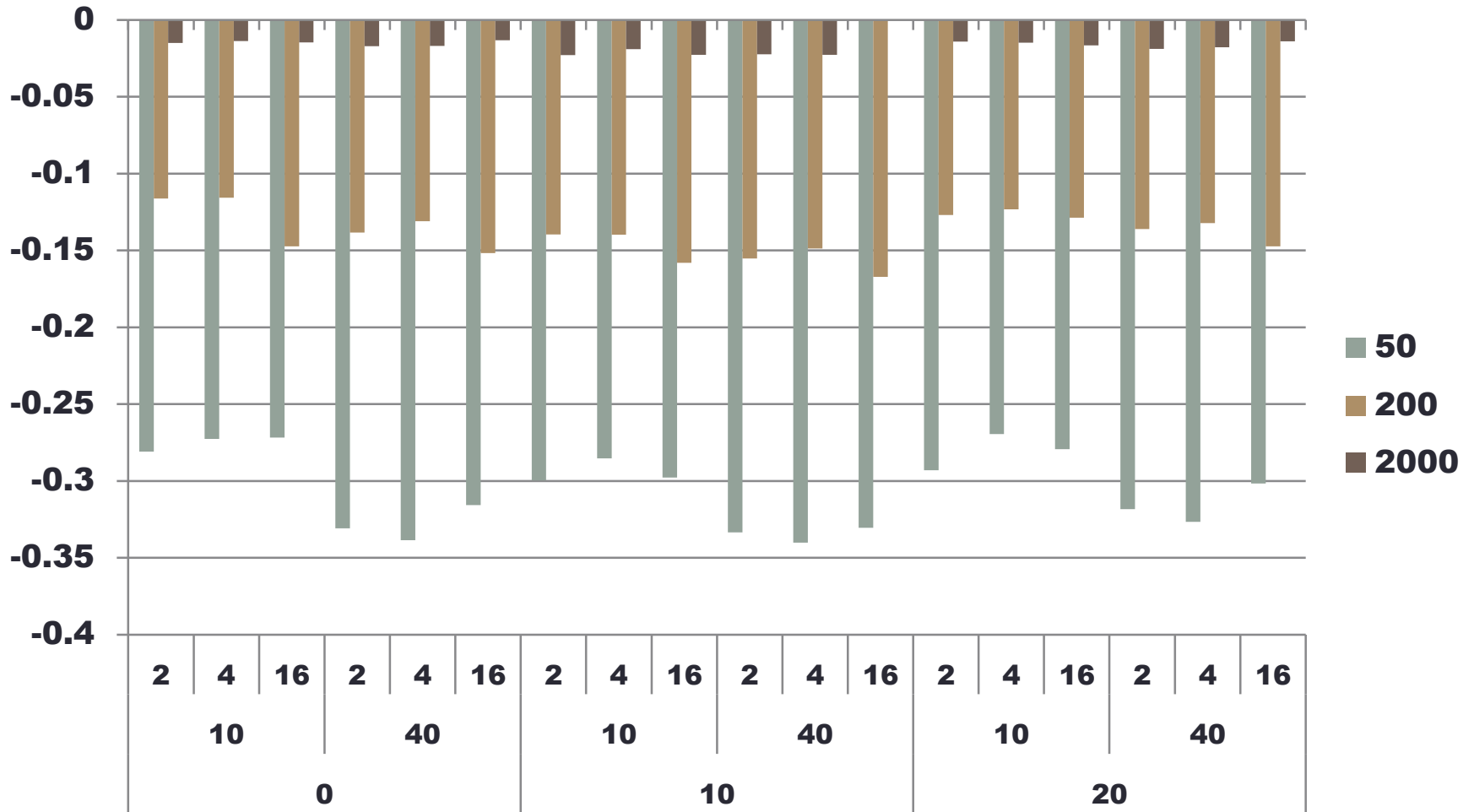
CURRENT STUDY RESULTS

COMPLETION RATES



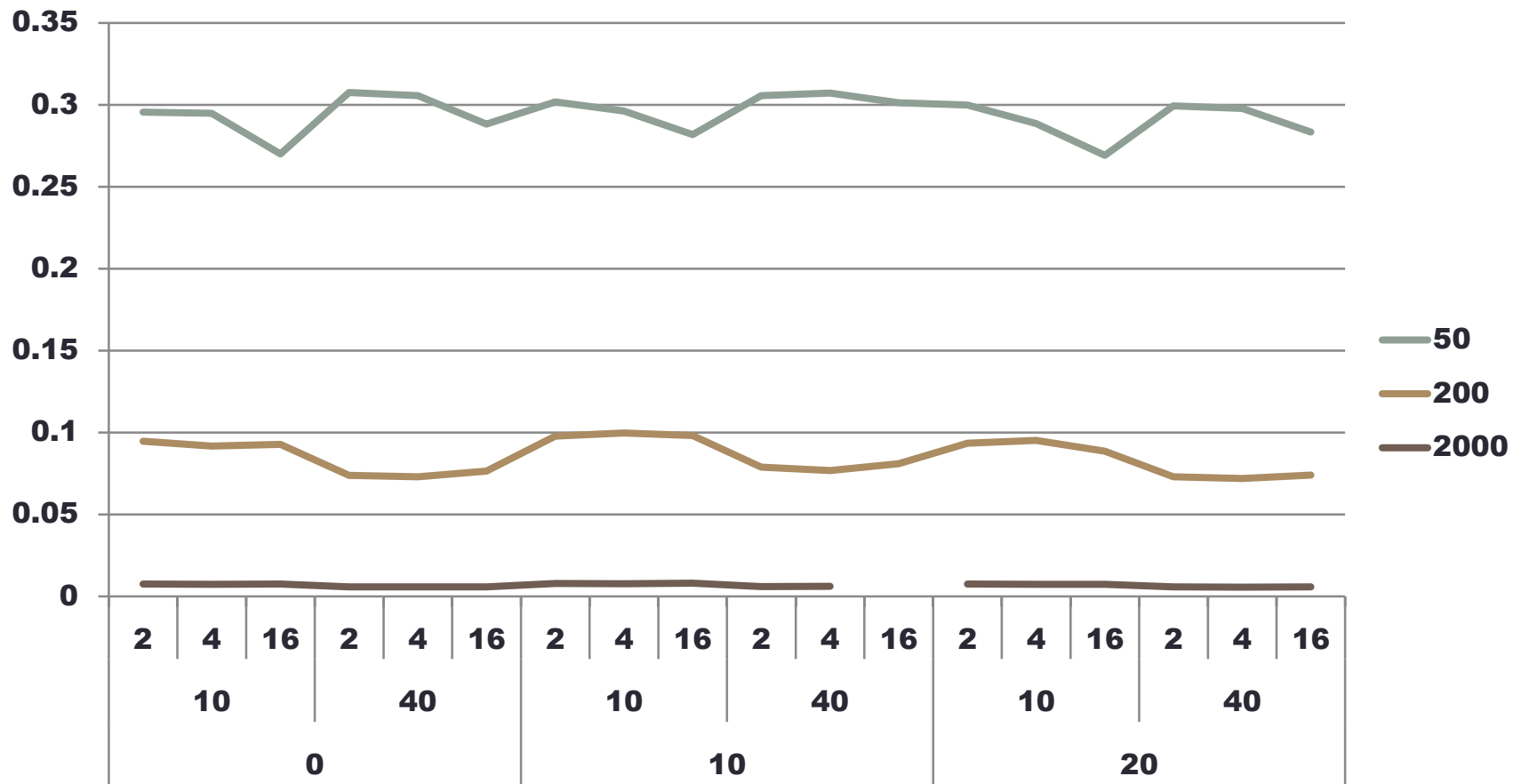
CURRENT STUDY RESULTS

BIAS IN LOADINGS



CURRENT STUDY RESULTS

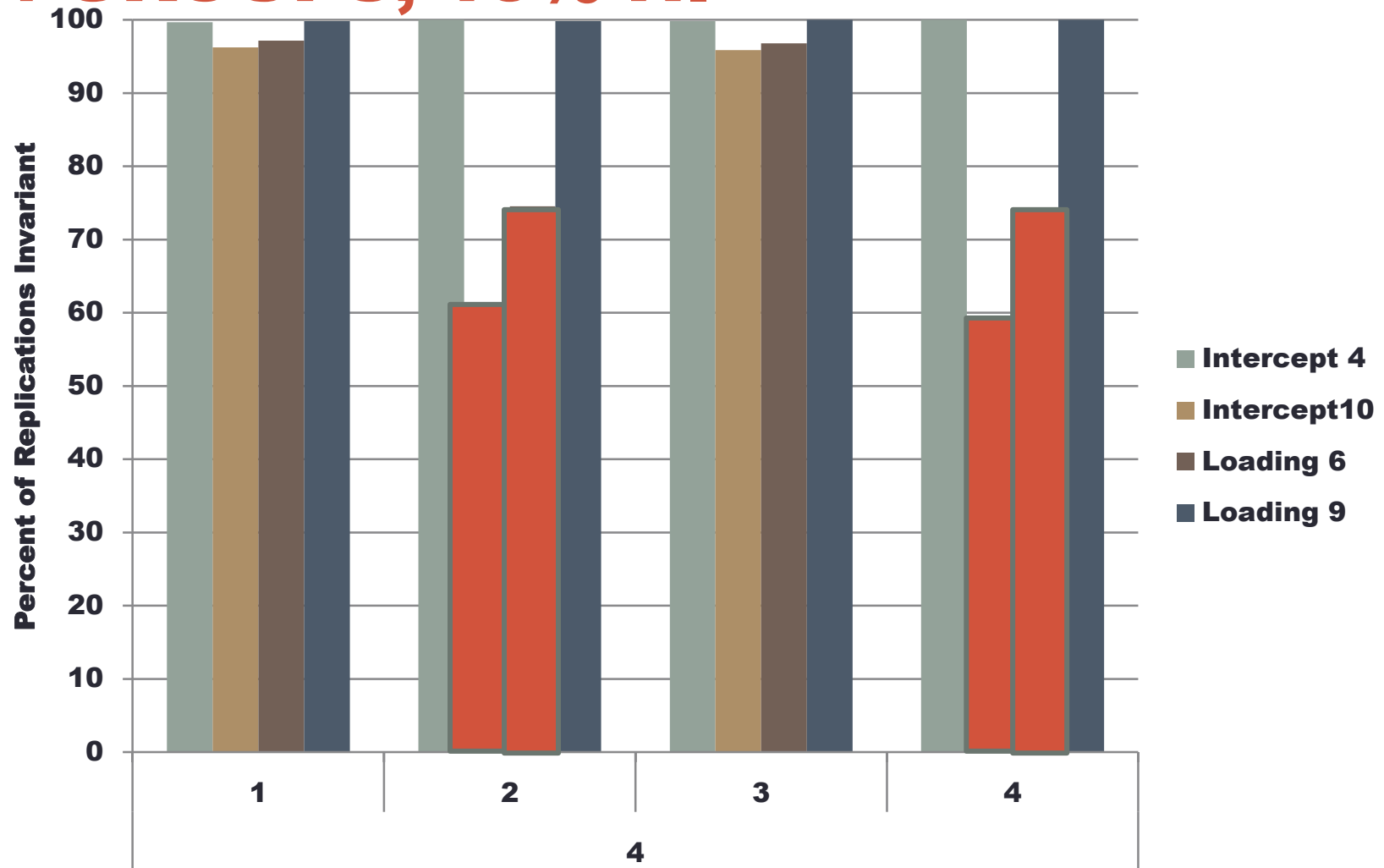
MSE FOR ALL PARAMETERS



CURRENT STUDY RESULTS

AD-HOC TESTING PROCEDURE

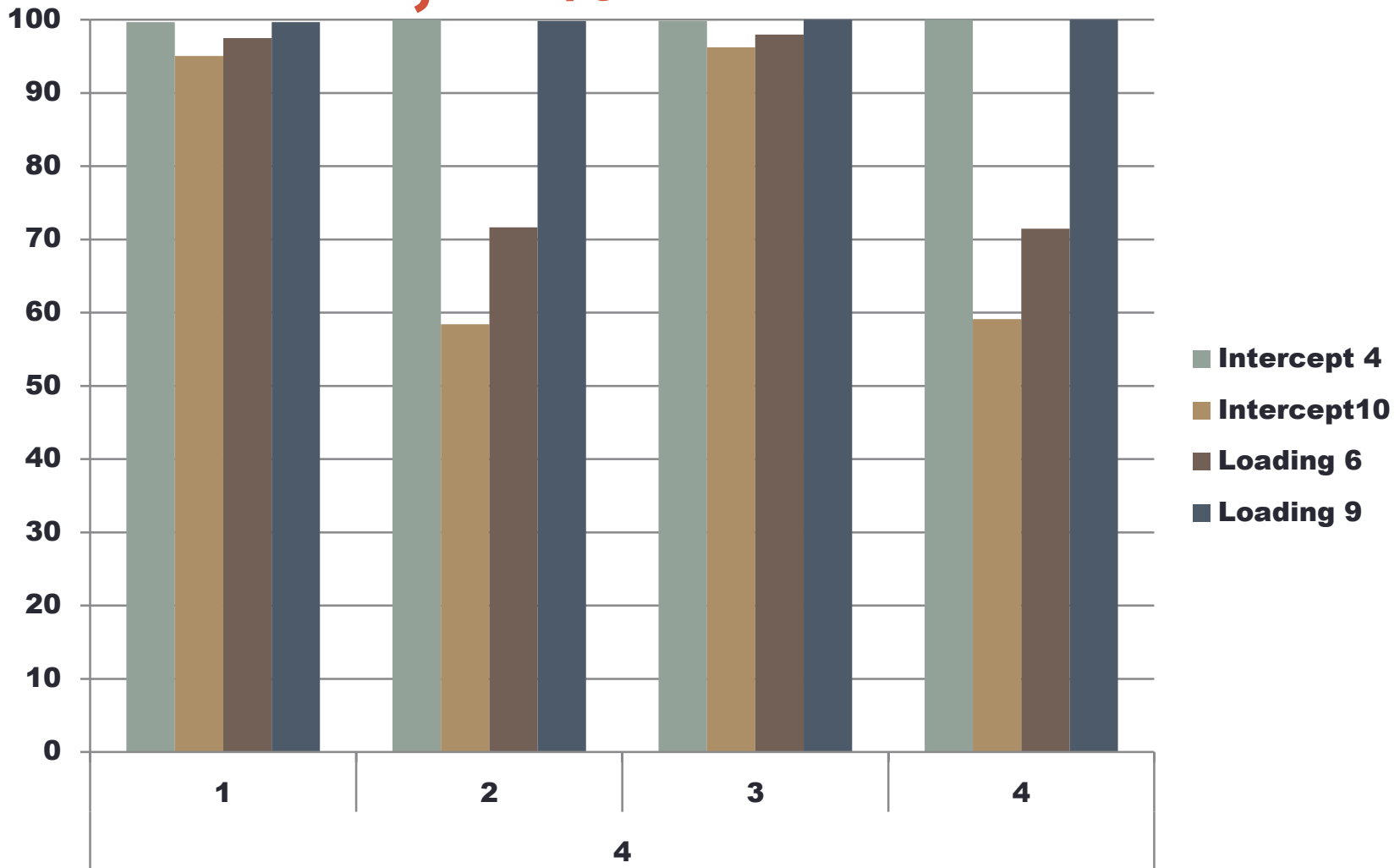
4 GROUPS, 10% NI



CURRENT STUDY RESULTS

AD-HOC TESTING PROCEDURE

4 GROUPS, 20% NI



IMPLICATIONS

- **Less completion with more groups**
 - Convergence may be hard to ascertain, thousands of parameters in the model
- **Bias and MSE in estimates low across all large sample (group n=2000) conditions**
- **Bias for thresholds lower than loading across all conditions**
- **Coverage was good as well (average across all parameters was .94)**
- **Ad-hoc procedure is too conservative**
 - NI thresholds not flagged ~40% of reps
 - NI loadings not flagged ~30% of reps

LIMITATIONS AND FUTURE RESEARCH

- **What is the effect of specifying an informative prior on the bias in smaller samples?**
- **How does MLR work in comparison?**
- **How does the invariance testing procedure perform with a less conservative criterion?**
- **How does the invariance testing procedure perform under various magnitudes of non-invariance?**
 - Our parameters were $\sim .5$ different