

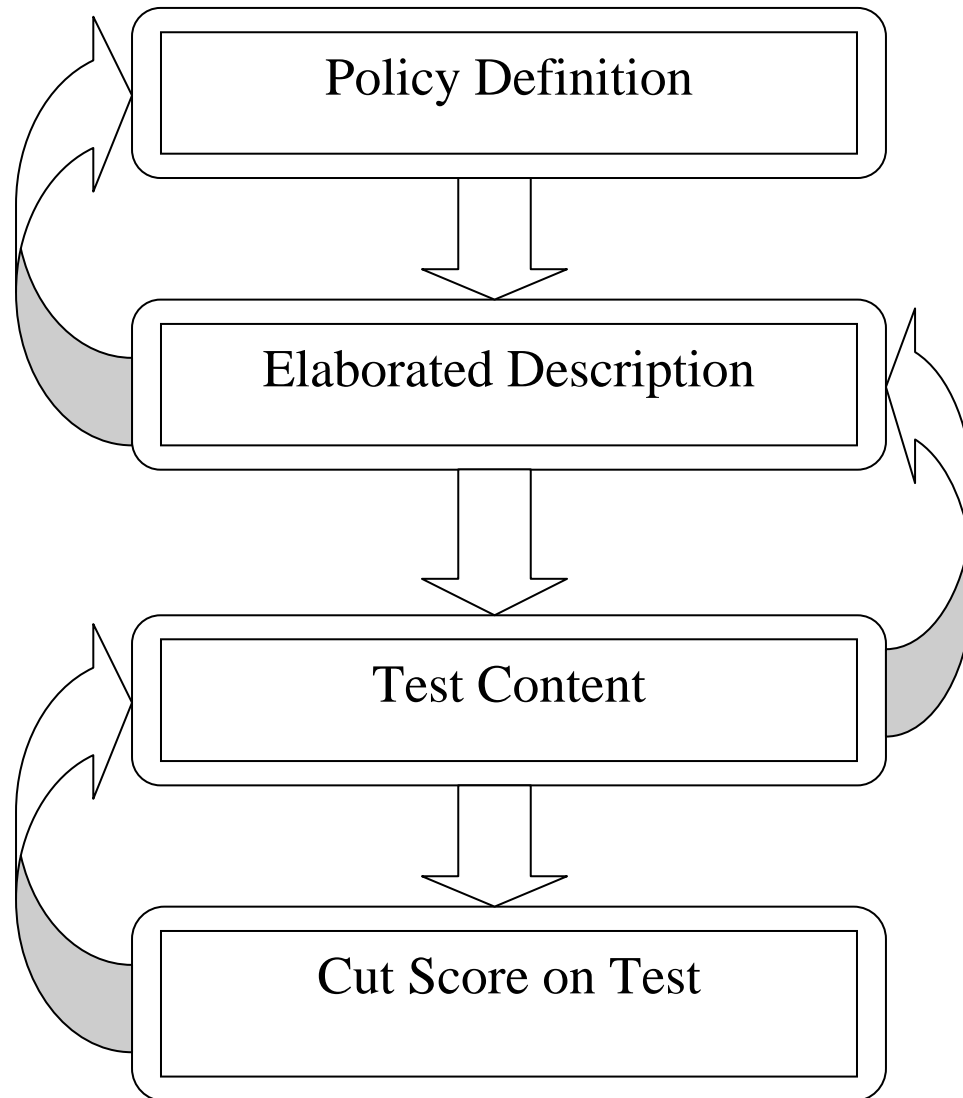
# Standard Setting Theory and Practice: Issues and Difficulties

Mark D. Reckase  
Michigan State University  
reckase@msu.edu

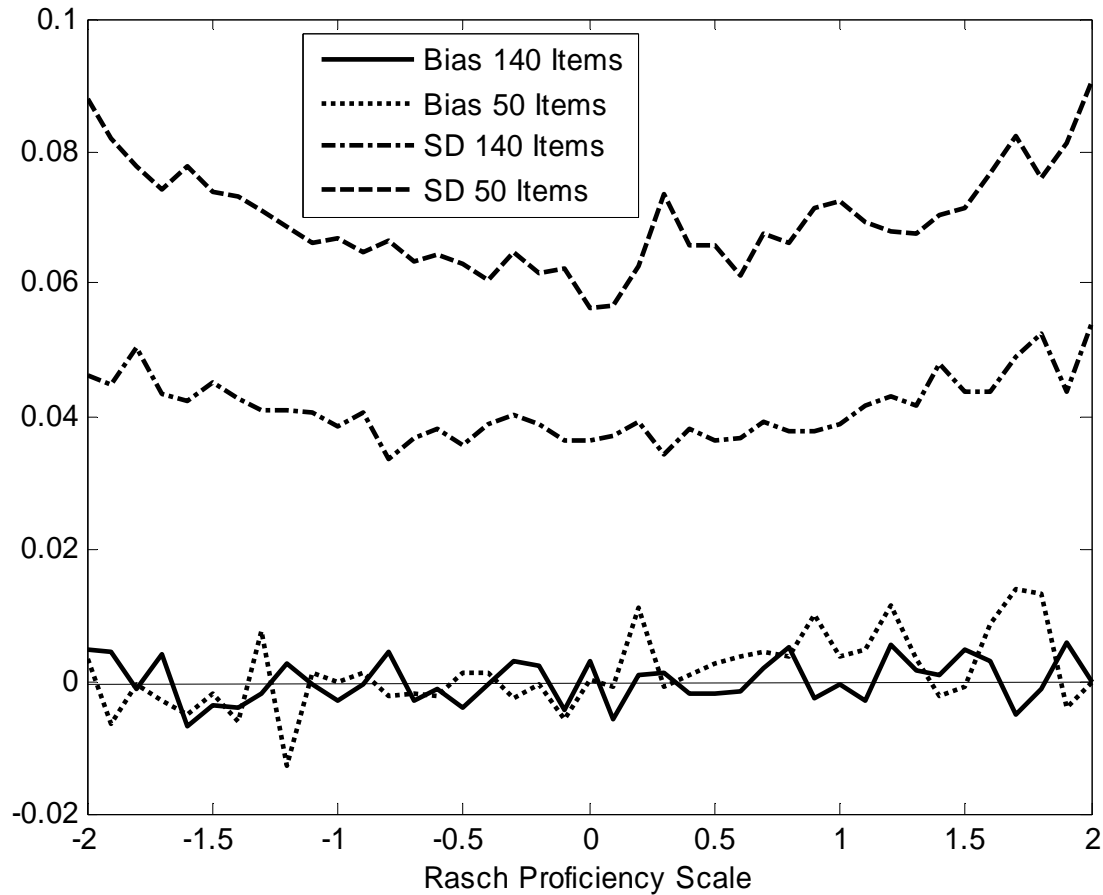


# Definition of a Standard

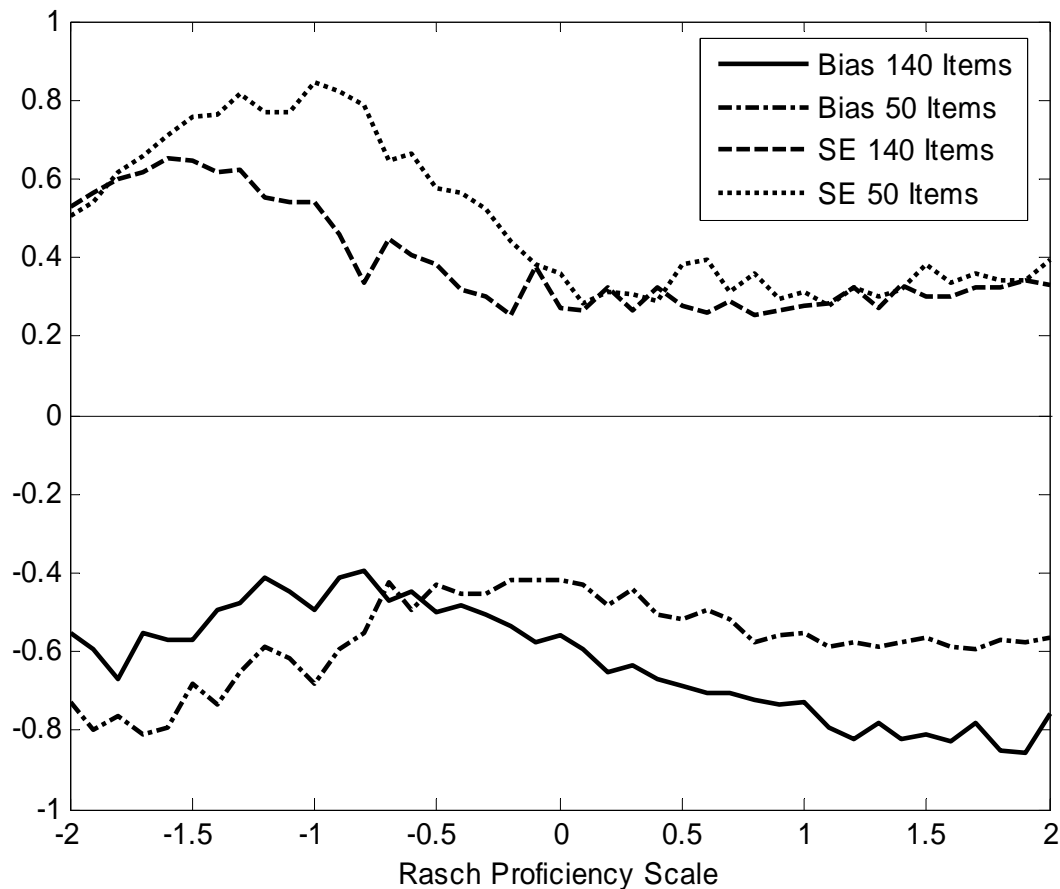
- Something established by authority, custom, or general consent as a model or example: criterion
- Something set up and established by authority as a rule for the measure of quantity, weight, extent, value or quality.



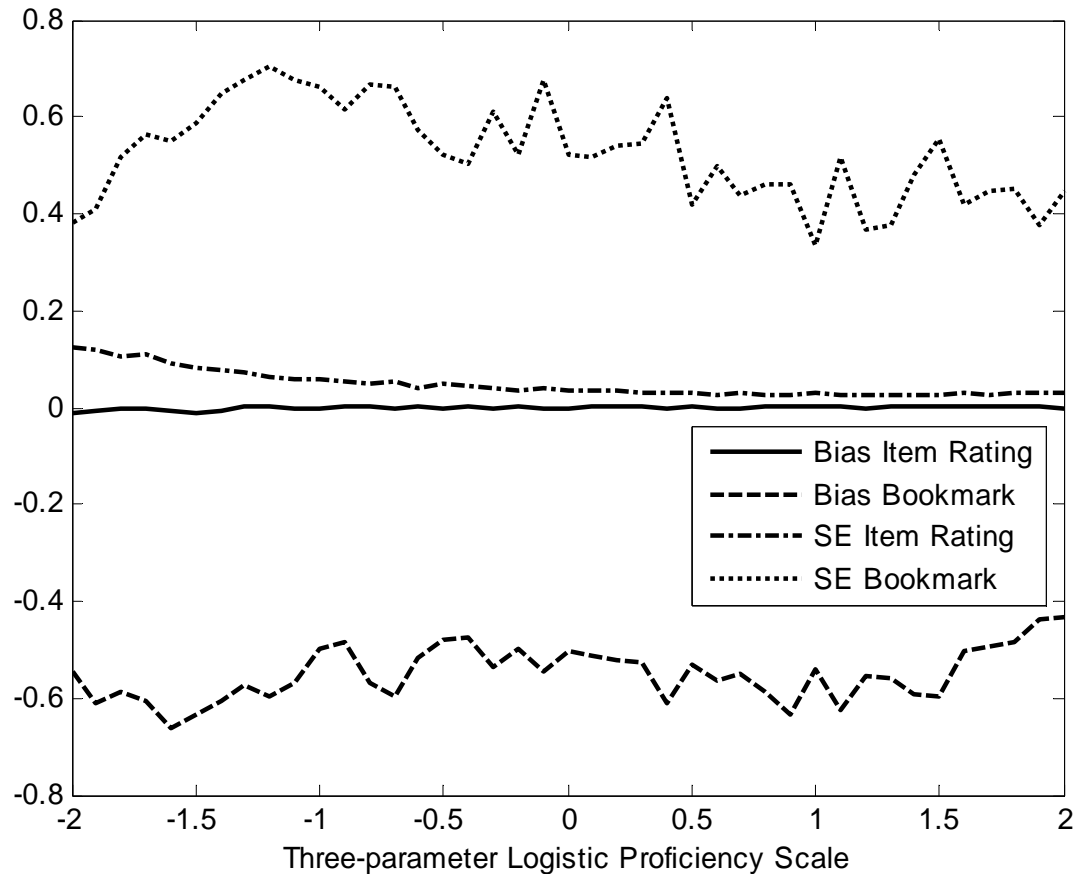
# Results with Judgment Error Angoff Method



# Results with Judgment Error Bookmark Method



# Results with Judgment Error Real Item Pool





# Conclusions

- Bookmark consistently has negative bias – no bias for Angoff item ratings
- Bookmark consistently has larger standard error than item rating
- There is an unusual interaction between test length and bias for the bookmark procedure – longer tests don't necessarily make better standards.
  - Panelists have more opportunities to misplace a bookmark with a long test – error accumulates.

# References

- Reckase, M. D. (2006). A conceptual framework for a psychometric theory for standard setting with examples of its use for evaluating the functioning of two standard setting methods. *Educational Measurement: Issues and practice*, 25(2), 4-18.
- Reckase, M. D. (2006). Rejoinder: evaluating standard setting methods using error models proposed by Schulz. *Educational Measurement: Issues and Practice*, 25(3), 14-17.



# References

- Haertel, E. H. & Lorie, W. A. (2004). Validating standards-based test score interpretations. *Measurement: Interdisciplinary Research & Perspective*, 2(2), 61-103.