

Rasch Psychometric Analysis of a Course Satisfaction Survey

Presented by Andrew D. Carson,
Senior Research Analyst,
Office of Institutional Effectiveness,
Kaplan University

The Theory | The Rasch Model

- Using log-odds transformations, Georg Rasch developed his measurement model to solve a testing problem for the Danish Department of Defense
- Initial model assumed dichotomous items arranged in difficulty along a single dimension; goal to create a scale in which each unit along it means exactly the same thing (i.e., it is objective and a true interval scale)
 - There are a number of ways to determine degree of deviation from the theoretic ideal
 - Variations from the basic model support use of multiple choice items, partial credit scoring, multiple dimensions
- The Rasch approach differs from other theories in some important respects
 - Classical test theory (CTT); now increasingly marginalized to smaller sample or clinical work
 - Item response theory (IRT); conceptually different in some ways from Rasch, but as usually implemented (3-parameter logistic, or 3PL) requires quite large Ns (beware: true-believers!)

Software | Winsteps

- Winsteps, single-user \$149, the basic program
- Facets, single user \$149, the more powerful but more complex version
 - Use Winsteps unless Winsteps won't work on a problem
- Licenses available for use by groups
- Other deals seem to be available as well, more on an ASP model
- Free time-limited license, and a free Ministeps version (restricted features)
- See handout for details

Example | Teacher Subscale on EOT Survey

- 15 Items, 4 of which also comprise Seminar subscale, and 5 more of which comprise the Discussion subscale

19	The professor stated clearly what is expected of students.	Teaching
20	The professor was responsive to my questions and concerns.	Teaching
21	I received helpful feedback from the professor on my course work	Teaching
22.1	The professor engaged students in seminar by: sharing relevant knowledge.	Teaching, Seminar
22.2	The professor engaged students in seminar by: keeping the discussion focused.	Teaching, Seminar
22.3	The professor engaged students in seminar by: building upon the contributions of students.	Teaching, Seminar
22.4	The professor engaged students in seminar by: encouraging all students to participate.	Teaching, Seminar
23.1	The professor added value to the Discussion Thread by: sharing relevant knowledge.	Teaching, Discussion
23.2	The professor added value to the Discussion Thread by: building upon the contribution of students.	Teaching, Discussion
23.3	The professor added value to the Discussion Thread by: providing useful insights.	Teaching, Discussion
23.4	The professor added value to the Discussion Thread by: connecting student responses to course material.	Teaching, Discussion
23.5	The professor added value to the Discussion Thread by: highlighting the practical relevance of course materials.	Teaching, Discussion
24	The professor encouraged me to do my best in this course.	Teaching
25	The professor's grading standards were appropriate.	Teaching
26	Overall, I rate the professor as: ³	Teaching

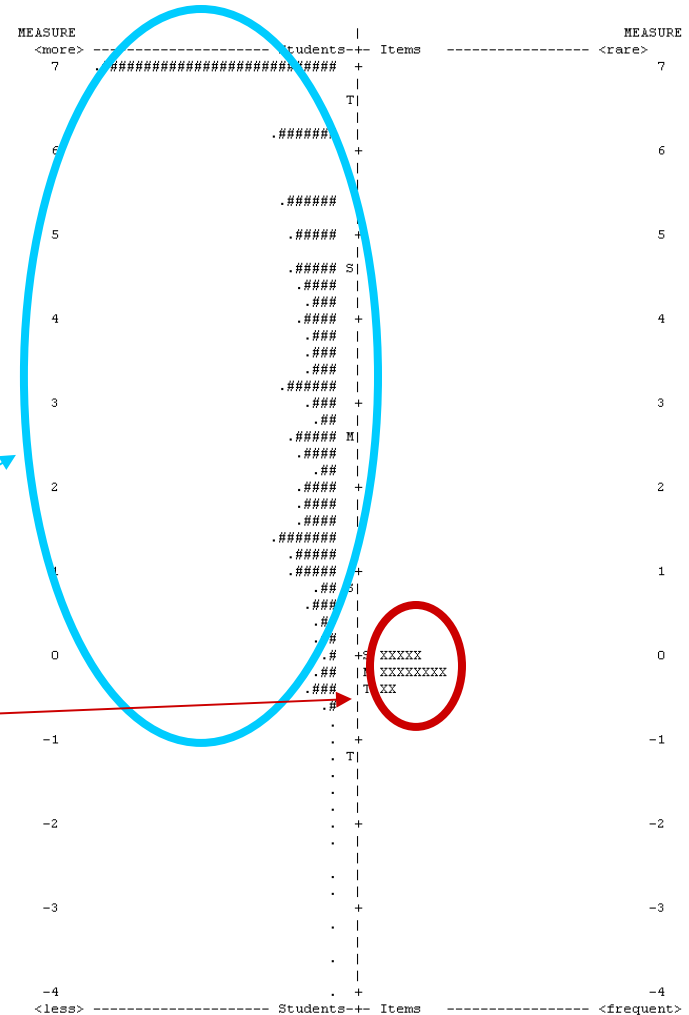
- Other subscales on the EOT Survey include Learner, Course, and there are some items that tap demographic information

OFFICE OF INSTITUTIONAL EFFECTIVENESS

DATA MANAGEMENT | RESEARCH | ASSESSMENT | ANALYTICS

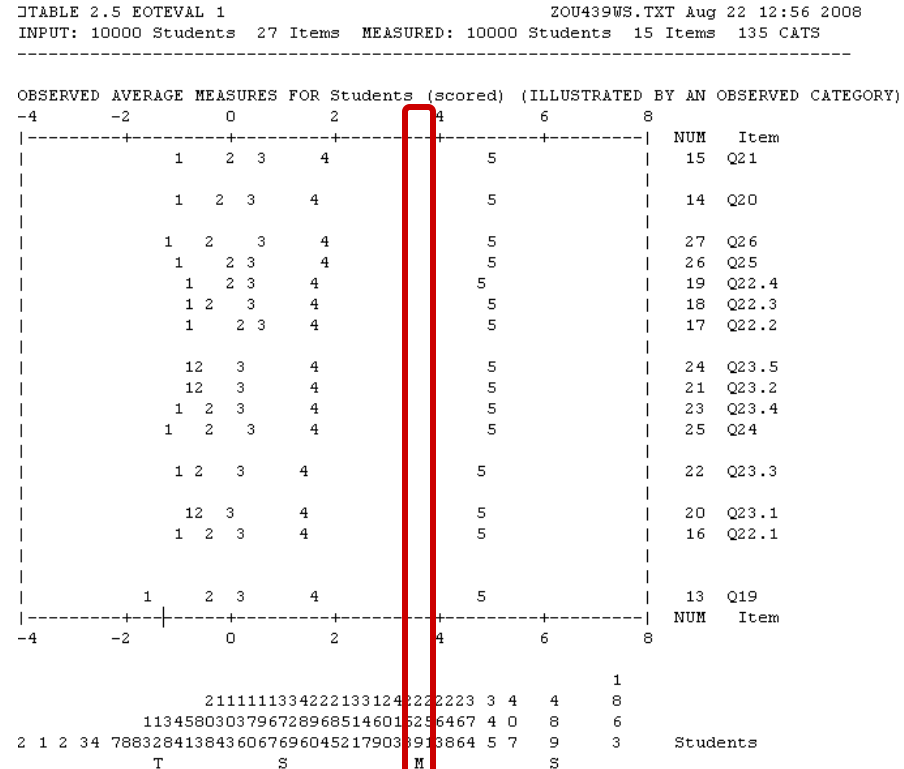
Example | Item & Person \times Logit

- The following examples are analyses of the 15-item teacher scale on the current undergraduate End of Term (EOT) Survey
- In a separate report, Cindy Hasemeier and I examined the dimensionality of the survey using exploratory factor analysis
- To the right, see the variable map from Facets for students and items
- Note that difficulty of **items** averaged at 0 logits, while **students** tended toward high (positive) ratings (aka highly skewed)
- Could suggest insufficient range for our teacher rating items



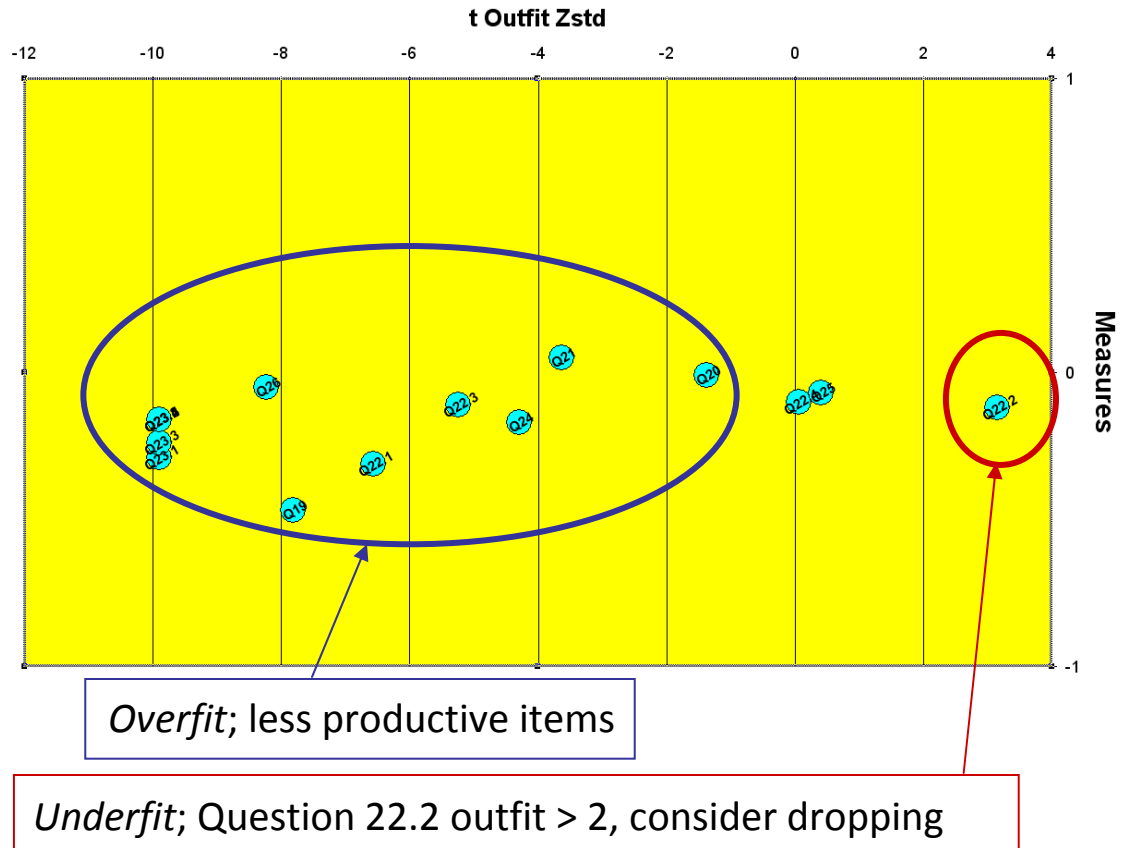
Example | Empirical Item-Category Measures

- Main figure shows distribution of the item categories by logits from mean for measure. Note little variation in distance from mean by item; remember, *all items fall at same level of difficulty*
- See bottom of figure; that's the distribution of persons. Mean for persons shown by "M", which is nearly 3 logits above mean for measure.



Example | Analysis of item misfit (here, outfit)

- Another indicator of possible problems with the Teacher subscale lies in examination of outfit statistics
- Outfit: Outlier sensitive fit statistic; high scores can suggest a distorted picture of the data
- Scores for *many* of the items are extreme
- *Perhaps time to rethink the scale*



Future Directions | Features of a Better Measure

- We are probably more concerned with dissatisfaction or poor quality – We need better item representation there
- We do not want all of our items to measure same point on satisfaction/quality dimension
 - This is true not only for stem, but also for any multiple choice options
- It may be easier to demonstrate desirable measurement properties if we use true-false items, or checklist items.

Future Directions | Possible items (varying in difficulty)

- The instructor was always well-prepared to teach.
- The instructor was usually very encouraging.
- The instructor succeeded in keeping our discussion on track.
- The instructor sometimes applied grading rules in an inconsistent manner.
- The instructor's train of thought wandered aimlessly during class.
- The instructor did not respond to my questions.

Bottom line: Don't make the multiple choice options do all the work!

Professional Development | Course and Statistics.com

- Statistics.com is professional development resource for statistics and measurement
 - Rasch Core Topics is the first of 3 sequenced courses on Rasch measurement theory and methods; there are additional Rasch-related courses of interest beyond this sequence.
 - The \$385 core topics course had 4 units (one week each) and took one month to complete; it was equivalent to about 1 traditional university credit
 - The course was entirely asynchronous and had about 32 students
 - The instructor was Mike Linacre, a leader in Rasch measurement theory, and a developer of some of its leading specialized software tools
- Good resource: Bond, T. G., & Fox, Christine M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Rasch Psychometric Analysis of a Course Satisfaction Survey

Presented by Andrew D. Carson,
Senior Research Analyst,
Office of Institutional Effectiveness,
Kaplan University