



EDUC 8720: Advanced Topics in Measurement Spring 2010

Mon 8:30-11:00 EDUC 330

Instructor

Derek Briggs

Office: EDUC 211 Tel: (303) 492-6320 E-mail: derekcbriggs@gmail.com

Office Hours: Tue 1:00-2:00 or by appointment

Course Overview

The focus of this course is on psychometric models for measurement and their applications in educational testing. An emphasis is placed on understanding and evaluating the utility of models from item response theory (IRT). This course is especially appropriate for students expecting to do research in which the score from a test instrument is the basis for evaluative conclusions about learning.

The best way to gain an understanding about measurement models is to apply and compare them in the context of simulated or empirical data sets. To this end, readings on various aspects of measurement models will first be presented and discussed in class. Next, the use of these models will be demonstrated by the instructor. Finally, students will be expected to apply the models using data sets provided by the instructor.

The focus of the first 2/3 of the semester is foundational topics necessary to help you understand IRT models at an acceptable level of depth.

1. Historical Context
2. Models for Dichotomous Items (BILOG)
3. Models for Polytomous Items (ConQuest)
4. Estimating the Parameters of an IRT Model
5. Evaluating Model Fit
6. Interpreting the IRT Score Scale

The focus last 1/3 of the course will be on three extensions and applications of IRT that represent areas that are both important, and areas where I have some professional experience.

1. Using IRT to Create a Developmental Score Scale: Vertical Scaling
2. Evaluating Parameter Invariance and Test Fairness: Differential Item Functioning
3. Dealing with Violations of Unidimensionality: Multidimensional Item Response Theory

Expectations and Objectives

Students enrolling in this course are expected to have previously taken EDUC 8710: Measurement in Survey Research, in which students are given an introduction to a variety of foundational measurement topics (e.g., survey instrument design, item formats, item analysis, reliability theory and validity theory). To understand *some* of the topics in this class at a deep level (e.g., parameter estimation) a background in calculus and matrix algebra is necessary (but this is not a requirement).

We are going to be dealing with some very complicated material in this course. How well you learn it will really depend on how much work you decide to put in. At a minimum you will need to do the required readings (of which there are a lot!), complete the assigned tasks, and laugh at my jokes when I give class presentations (or if you can't do that at least don't boo me or throw things). If you do all this, I expect you acquire some functional literacy when it comes to measurement models and their application in contexts where people are given standardized test instruments.

Functional literacy is nothing to sneeze at! It means you'll understand the jargon that gets used in the IRT literature, and if given a data set with item responses in the future, you'd be able to use your class notes to apply an IRT model to them and even explain the resulting output to someone that cares. Functional literacy also means that you will be a more critical consumer of measurement models than you were before you took this course.

But if you want to be more than functionally literate—if, for example, you would like to pursue a career in which psychometrics figures prominently, if you want to be able to think deeply about issues germane to a goal of “measuring the mind” as Borsboom puts it in his book—then you need to work even harder. That means reading the material multiple times, meeting outside of class to discuss it with classmates, visiting me in office hours to ask questions, and challenging yourself by taking on additional readings beyond those that have been explicitly required (what I refer to below as the “going deeper” readings). Perhaps more importantly, you will need to appreciate that it is not possible to become an expert in psychometrics in one (or even two) semesters! What this course should do is lay a foundation—what you do with that foundation is up to you.

Course Readings

Required Textbook

Embretson, S. E., and Reise, S. P. (2000) *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.

All required readings beyond these two books will be available at course web site: <http://culearn.colorado.edu>. All “going deeper” readings will be made available by request.

Please do the readings in the order in which they are listed within the course schedule.

Course Schedule

There will be no class meeting on the following dates:

January 18 (MLK Jr Holiday)
March 22 (Spring Break)

0. Overview of Course

1/11

- Introductions
- Recap from Measurement in Survey Research
- Why this Stuff Matters

1. Some Historical Context for Measurement Models

1/25

Required Readings:

Embretson & Reise [Chapter 1, 3-9]
Borsboom [Ch 1-2, 1-44]
Lord & Novick (1968) *Statistical theories of mental test scores*. [Ch. 1, 13-26]
Rasch (1960) *Probabilistic models for some intelligence and attainment tests*. [Preface and introduction, xx-xxiii, 3-12]

Going Deeper

Traub, R. E. (1997) Classical test theory in historical perspective. *Educational Measurement: Issues and Practice*. December 1997, 8-14.
Brennan, R. (1997) A perspective on the history of generalizability theory. *Educational Measurement: Issues and Practice*. December 1997, 14-20.
Bock, R. D. (1997) A brief history of item response theory. *Educational Measurement: Issues and Practice*. December 1997, 21-32.*
Wright, B. D (1997) A history of social science measurement. *Educational Measurement: Issues and Practice*. December 1997, 33-45.

Note: * indicates technical material requiring a strong background in statistics

Readings to do when you have more time

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*. Cambridge University Press.

2. The Promise & Potential of IRT (in a Nutshell)

2/1

Required Readings:

- Partchev, I. A Visual Guide to IRT by Ivailo Partchev <http://www.metheval.uni-jena.de/irt/VisualIRT.pdf> (Ch. 1-8, pp. 1-51)
Embretson & Reise [Chapter 3, 40-64]
Hambleton, R., & Jones, R. (1993) An NCME instructional module: Comparison of Classical Test Theory and Item Response Theory and their applications to test development. *Educational Measurement: Issues and Practice* 12 (3), 38–47.
Borsboom [Ch. 3, 49-81]

Going Deeper

- Lord & Novick (1968) [Ch. 16: Latent Traits and Item Characteristic Functions, 368-396]
Van der Linden, W. J., and Hambleton, R. K. (1997) *Handbook of modern item response theory*. New York, NY: Springer-Verlag.* [Preface, 1-28]

Readings to do when you have more time

- De Boeck, P. and Wilson, M., eds. (2004) *Explanatory item response models: a generalized linear and nonlinear approach*. New York: Springer.* [Ch. 1-2]
Holland, P. W. (1990) On the sampling theory foundations of item response theory models. *Psychometrika*, 55, 577-602.*

3: The Mechanics of Modeling Dichotomous Items (BILOG)

2/8

Required Readings

- Embretson & Reise (Chapter 4, 65-83)
Harris, D. (1989). Comparison of 1-, 2-, and 3-parameter IRT models. *Educational Measurement: Issues and Practice* 8 (1), 35–41.
Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms? In Smith & Smith (eds) *Introduction to Rasch Measurement*. JAM Press, 143-166.
BILOG User's Manual

Going Deeper:

- Birnbaum, A. (1968) Some latent trait models. In Lord, F. M and Novick, M. R. *Statistical theories of mental test scores*. * [Ch. 17, 397-424]
Thissen, D., & Orlando, M. in D. Thissen and H. Wainer, eds. (2001) *Test Scoring* Mahwah, NJ: Lawrence Erlbaum Associates. [Ch. 3, 73-98]

4. Estimating the Parameters of IRT Models (BILOG)

2/15, 2/22

Required Readings

Embretson & Reise [Chapters 7-8; 158-225]

Mislevy, R. & Stocking, M. (1989) A consumer's guide to LOGIST and BILOG.
Applied Psychological Measurement, 13(1), 57-75.

Going Deeper

Thissen & Orlando, Ch. 3 (98-137) in *Test Scoring*.

Thissen et. al, Ch. 4 (149-173) in *Test Scoring*.

Readings to do when you have more time

Baker, F. B. and Kim, S-H. (2004) *Item response theory: Parameter Estimation Techniques*. 2nd Edition. New York, NY: Marcel Dekker, Inc.*

5. Evaluating Model Fit (BILOG)

3/1

Required Readings

Embretson & Reise [Chapter 9, 226-248]

Hambleton, R. K. & Swaminathan, H. (1985) *Item response theory: principles and applications*. Boston: Kluwer-Nijhoff Publishing. [Chapters 8-9, 151-196]

Yen, W. (2006) Item response theory. In R. Brennan, (ed.) *Educational Measurement*, 4th ed, Westport, CT: American Council on Education/Praeger. [Ch. 4, 138-143]

Going Deeper

Smith, R. (2004). Fit analysis in latent trait measurement. In Smith & Smith (eds)
Introduction to Rasch Measurement. JAM Press, 73-92.

Karabatsos, G. (2000). A critique of Rasch residual fit statistics. *Journal of Applied Measurement*. 1(2), 152-176.

Readings to do when you have more time

Orlando, M. & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement*, 24(1), 50-64.

Glas, C., & Verhelst, N. (1989). Extensions of the partial credit model. *Psychometrika*, 54(4), 635-659.*

6: The Mechanics of Modeling Polytomous Items (ConQuest)

3/8

Required Readings

Embretson & Reise [Chapter 5, 95-124]
ConQuest 2.0 User's Manual

Going Deeper

Thissen et. al, Ch. 4 (141-186) in D. Thissen and H. Wainer, eds. (2001) *Test Scoring*
Mahwah, NJ: Lawrence Erlbaum Associates.
Adams, R., Wilson, M., Wang, W. The multidimensional random coefficients
multinomial logit model. *Applied Psychological Measurement*, 21, 1-23.*

Readings to do when you have more time

De Boeck, P. and Wilson, M., eds. (2004) *Explanatory item response models: a
generalized linear and nonlinear approach*. New York: Springer.* [Ch. 3]

7. Interpreting the Measurement Scale in IRT

3/15, 3/29

Required Readings

Embretson & Reise [Chapter 6, 125-157] 3/15
Borsboom [Chapters 4, 85-120] 3/15
Borsboom [Chapter 5, 121-145] 3/29

Going Deeper

Kyngdon, A. (2008) The Rasch model from the perspective of the representational theory
of measurement. *Theory & Psychology*. 18(1), 89-109.
Michell, J. (2008). Conjoint measurement and the Rasch model. *Theory & Psychology*.
18(1), 119-124.
Karabatsos, G. (2001). The Rasch model, additive conjoint measurement, and new
models of probabilistic measurement theory. *Journal of Applied Measurement*, 2,
389-423.

Readings to do when you have more time

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*.
Chicago: University of Chicago Press. [Ch. 5-6, 62-108]
Luce, R. D. and Tukey, J. W. (1964). Simultaneous conjoint measurement: a new type of
fundamental measurement. *Journal of Mathematical Psychology*, 1, 1-27.*

8. Using IRT to Create a Developmental Score Scale: Vertical Scaling 4/5

Required Readings

- Kolen, M (2006). Scaling and norming. In R. Brennan, (ed.) *Educational Measurement*, 4th ed, Westport, CT: American Council on Education/Praeger. [Ch. 5, 155-186]
- Briggs, D. C. & Weeks, J. P. (2009) The impact of vertical scaling decisions on growth interpretations. *Educational Measurement: Issues & Practice*, 28(4), 3-14.
- Ballou, D. (2009). Test scaling and value-added measurement. *Education Finance and Policy*, 4(4), 384-414.
- Briggs, D. C. & Betebenner, D. (2009) Is Growth in Student Achievement Scale Dependent? Paper presented at the invited symposium “Measuring and Evaluating Changes in Student Achievement” at the annual meeting of the National Council for Measurement in Education, San Diego, CA, April 14, 2009.

9. Evaluating Parameter Invariance: Differential Item Functioning 4/12

Required Readings

- Clauser, B. & Mazor, K. (1998) Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*. 17, 31-44.
- ConQuest 2.0 User’s Manual, Chapters 1-2, 8.
- Gierl, M. (2005) Using dimensionality-based DIF analyses to identify and interpret constructs that elicit group differences. *Educational Measurement: Issues and Practice*, 24, Spring 3-14.
- Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991) *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications. [Chapter 8, 109-122]

Going Deeper

Camilli, G. Test fairness. In R. Brennan (ed) *Educational Measurement*, 4th Edition.

Readings to do when you have more time

- Holland, P. & Thayer, D. (1988) Differential items and Maentel-Haenszel. In H. Wainer & H. Braun (eds.) *Test Validity*. Hillsdale, NJ: Erlbaum, 129-144.*
- Penfield, R. & Lam, T. (2000) Assessing differential item functioning in performance assessment: review and recommendations. *Educational Measurement: Issues and Practice*, 19, Fall 5-15.

10. Multidimensional Item Response Models

4/19, 4/26

Required Readings

- Ackerman, T., Gierl, M. J., and Walker, C. M. (2003) Using multidimensional IRT to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice* (37-53).
- Briggs, D. C. and Wilson, M. W. (2003) An introduction to multidimensional measurement using Rasch models. *Journal of Applied Measurement*, 4(1), 87-100.
- Kennedy, C. (2005). Constructing measurement models for MRCML estimation: A primer for using the BEAR scoring engine. BEAR Center Technical Report Series No. 2005-04-02

Going Deeper

- Reckase, M. (2009) Multidimensional IRT models in the book *Multidimensional Item Response Theory*, Springer. (Chapter 4, 79-112)*

Readings to do when you have more time

- Ackerman, T. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 29, 67-91.*
- Ackerman, T. (1994). Using multidimensional item response theory to understand what items and tests are measuring. *Applied Measurement in Education*, 7, 255-278.*
- Briggs, D. C. (2008) Using explanatory item response models to analyze group differences in science achievement. *Applied Measurement in Education*.
- Kupermintz, H., Ennis, M. M., Hamilton, L. S., Talbert, J. E., & Snow, R. E. (1995). Enhancing the Validity and Usefulness of Large-Scale Educational Assessments .1. NELS-88 Mathematics Achievement. *American Educational Research Journal*, 32(3), 525-554.
- Nandakumar, R., and Ackerman, T. (2004). Test modeling in D. Kaplan, ed., in *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks, CA: SAGE Publications. [93-106]
- Reckase, M. (2009). *Multidimensional Item Response Theory*. New York, NY: Springer-Verlag.*
- Thissen, D., and Wainer, H., eds. (2001) *Test Scoring*. Mahwah, NJ: Lawrence Erlbaum Associates. [Ch. 5-6]*
- Van der Linden, W. J., and Hambleton, R. K. (1997) *Handbook of modern item response theory*. New York, NY: Springer-Verlag. [Part III, 221-323]*
- Walker, C., Azen R., & Schmitt, T. (2006) Statistical Versus Substantive Dimensionality: The Effect of Distributional Differences on Dimensionality Assessment Using DIMTEST. *Educational and Psychological Measurement*, 66(5), 721-738.

BONUS TOPIC: Generalizability Theory

My plan is to conduct a special 2-3 session workshop on this topic at a time and date TBD. Attendance is optional.

Required Readings

Brennan, R. (1992) NCME Instructional Module on Generalizability Theory.
Thompson, B. (2003) A brief introduction to generalizability theory. In *Score Reliability*, B. Thompson, ed. Thousand Oaks, CA: SAGE Publications.

Going Deeper

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability Theory: A Primer*. Newbury Park: Sage Publications.

Readings to do when you have more time

Brennan, R. (2001). *Generalizability Theory*. New York: Springer-Verlag.*
Briggs, D. C., & Wilson, M. (2007) Generalizability in item response modeling. *Journal of Educational Measurement*, 44(2), 131-155.*
Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. New York: John Wiley.* [out of print]

Class Assignments

Weekly Discussions

You will be expected to team with one or two of your classmates on a regular basis to lead classroom discussions or activities each week. Each student team will meet with me to discuss their plan of action by the Friday prior to a Monday class.

Weekly “Problem Sets”

From week to week I will be giving you tasks to do to prepare yourself for what goes on in class. These tasks are meant for purely formative purposes—they will not be graded. However, it is very important that you work on these because if all you do is read the assigned materials, you will never come to an adequate understanding of the underlying topics.

Empirical Analysis

You will be given one data analysis assignment in which you will apply and/or compare IRT models for dichotomous items using the software BILOG and an empirical data set that I will provide. (Note: you are welcome to use your own data sets, provided they meet

certain requirements in terms of number of items and respondents.) The assignment writeup should be between 5 and 10 double-spaced pages, including figures (but not including references) You are required to turn in your writeup for feedback. This feedback will include suggestions for improving your analysis. The first draft will be due on March 1st.

Details on this assignment will be forthcoming as a separate handout.

Independent Project

As will become clear, there are many, many applications and extensions of IRT, especially in the context of large-scale standardized testing. For this project you will be given the flexibility to pick your own application/extension to pursue in detail. A good place to look for ideas for your independent project would be

1. Chapters 10-12 in the Embretson & Reise textbook.
2. Browsing through recent issues in research journals;

While I expect you to tailor this project to your specific interests in educational measurement, another approach would be to follow up on topics introduced in class at a deeper level through either a literature review and/or by conducting analyses with empirical or simulated data.

You will turn in a 1-2 page proposal for your independent project no later than March 8th. At that point I will give you feedback and help you to make a plan for what readings you should consult and data you should analyze as part of your project. The page limit for this project is 20 pages, double-spaced. The final draft of your project will be due by 10:00 on May 3rd.

Software

We will be using software for two programs as part of this course.

1. BILOG-MG 3.0. Item response models for dichotomous items.
2. ConQuest 2.0. Rasch family of item response models for dichotomous and polytomous items.

At the following web site <http://estore.e-academy.com/index.cfm?loc=IRT/main> you can try (for 15 days), or rent (for 6 months, \$40), the BILOG program. I'm trying to figure out a way to provide you with a free copy of ConQuest.

Neither of the programs above are especially intuitive. Figuring out how to use them effectively will be something we will work on together in class.

Student Assessment

Grades for this course will be based on the following criteria

Leading Topic Discussions & Class Participation	20%
Data Analysis	40%
Independent Project	40%

Reasonable Accommodation

If you qualify for accommodations because of a disability, please submit to me a letter from Disability Services in a timely manner so that your needs may be addressed. Disability Services determines accommodations based on documented disabilities. Contact: 303-492-8671, Willard 322, and www.Colorado.EDU/disabilityservices

Disability Services' letters for students with disabilities indicate legally mandated reasonable accommodations. The syllabus statements and answers to Frequently Asked Questions can be found at www.colorado.edu/disabilityservices

Religious Observances

I will make every effort to accommodate all students who, because of religious obligations, have conflicts with scheduled exams, assignments, or other required attendance, provided advance notification of the conflict is given. Whenever possible, students should give at least two weeks advance notice to request special accommodation. For additional information on this policy, see http://www.colorado.edu/policies/fac_religh.html

Classroom Behavior

Students and faculty each have responsibility for maintaining an appropriate learning environment. Students who fail to adhere to such behavioral standards may be subject to discipline. Faculty have the professional responsibility to treat all students with understanding, dignity and respect, to guide classroom discussion and to set reasonable limits on the manner in which they and their students express opinions. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender variance, and nationalities. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. See policies at <http://www.colorado.edu/policies/classbehavior.html> and at http://www.colorado.edu/studentaffairs/judicialaffairs/code.html#student_code

The University of Colorado at Boulder policy on Discrimination and Harassment (<http://www.colorado.edu/policies/discrimination.html>), the University of Colorado

policy on Sexual Harassment and the University of Colorado policy on Amorous Relationships applies to all students, staff and faculty. Any student, staff or faculty member who believes s/he has been the subject of discrimination or harassment based upon race, color, national origin, sex, age, disability, religion, sexual orientation, or veteran status should contact the Office of Discrimination and Harassment (ODH) at 303-492-2127 or the Office of Judicial Affairs at 303-492-5550. Information about the ODH and the campus resources available to assist individuals regarding discrimination or harassment can be obtained at <http://www.colorado.edu/odh>

Student Honor Code

All students of the University of Colorado at Boulder are responsible for knowing and adhering to the academic integrity policy of this institution. Violations of this policy may include: cheating, plagiarism, aid of academic dishonesty, fabrication, lying, bribery, and threatening behavior. All incidents of academic misconduct shall be reported to the Honor Code Council (honor@colorado.edu; 303-725-2273). Students who are found to be in violation of the academic integrity policy will be subject to both academic sanctions from the faculty member and non-academic sanctions (including but not limited to university probation, suspension, or expulsion). Other information on the Honor Code can be found at <http://www.colorado.edu/policies/honor.html> and at <http://www.colorado.edu/academics/honorcode/>

DEREK'S EVOLVING LIST OF RESOURCES AND READINGS FOR ADVANCED TOPICS IN MEASUREMENT

(* indicates technical material requiring a strong background in statistics)

Highly Recommended Textbooks

Brennan, R. L., ed. (2006) *Educational Measurement*, 4th ed. Westport, CT: American Council on Education/Praeger.

Lord, F. M & Novick, M. R. (1968; 2008) *Statistical theories of mental test scores*. Information Age Publishing Inc. *

Thissen, D., & Wainer, H., eds. (2001) *Test Scoring*. Mahwah, NJ: Lawrence Erlbaum Associates.

Other Recommended Books

Baker, Frank (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD. Available online for free at: <http://edres.org/irt/>

Baker, F. B. & Kim, S-H. (2004) *Item response theory: Parameter Estimation Techniques*. 2nd Edition. New York, NY: Marcel Dekker, Inc.*

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. New York: John Wiley.* [out of print]

De Boeck, P. & Wilson, M., Eds. (2004) *Explanatory item response models: a generalized linear and nonlinear approach*. New York: Springer.*

Fischer, G. & Molenaar, I., Eds. (1995) *Rasch models: foundations, recent developments and applications*. New York: Springer.*

Hambleton, R. K. & Swaminathan, H. (1985) *Item response theory: principles and applications*. Boston: Kluwer-Nijhoff Publishing.

Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991) *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications.

Kolen, M. J. & Brennan, R.L. (2004) *Test Equating, Scaling and Linking*. 2nd Edition. New York: Springer-Verlag.

Lord, F. M. (1980) *Applications of Item response theory to Practical Testing Problems*. Mahwah, NJ: Lawrence Erlbaum Associates.*

Linn, R.L., Ed. (1989) *Educational Measurement*. 3rd edition. Washington, D.C.: American Council on Education.

Smith, E. V., & Smith, R. M., eds. (2004) *Introduction to Rasch Measurement*. Maple Grove, MN: JAM Press.

Van der Linden, W. J., & Hambleton, R. K. (1997) *Handbook of modern item response theory*. New York, NY: Springer-Verlag.*

Wright, B. D., & Masters, G. M. (1982). *Rating scale analysis*. Chicago, IL: Mesa Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: Mesa Press.

Recommended Journals with a Focus on Psychometrics

(* indicates a journal that focuses primarily on technical modeling issues in measurement)

Applied Measurement in Education

Applied Psychological Measurement*

Educational Assessment

Educational Measurement: Issues and Practice

Educational and Psychological Measurement

Journal of Applied Measurement

Journal of Educational Measurement*

Measurement: Interdisciplinary Research and Perspectives

Psychometrika*

Psychological Bulletin

Internet Resources

National Council for Measurement in Education www.ncme.org

A wonderful feature at this site is the NCME instructional modules, some of which we will be using in this class: <http://www.ncme.org/pubs/items.cfm>

Rasch Measurement <http://www.rasch.org/rmt/index.htm>

Lots of short, informative articles about the application of the Rasch Model.

The Psychometric Society <http://www.psychometrika.org/>

IRT Tutorial from the University of Illinois at Urbana-Champaign

<http://work.psych.uiuc.edu/irt/tutorial.asp>

IRT at Wikipedia http://en.wikipedia.org/wiki/Item_response_theory

This is actually a pretty good summary, believe it or not.