**Psychology and Research in Education 932: Diagnostic Testing**

## Course Information:

| | |
|---|---|
| Course Meeting Time: | Wednesdays from 1:30-4:20pm |
| Course Location: | Room 245, Joseph R. Pearson Hall |
| Course Website: | http://wp.me/P3nkOf-qc |
| | |
| Instructor: | Jonathan Templin |
| Email: | jtemplin@ku.edu |
| Office: | Room 614, Joseph R. Pearson Hall |
| Office Hours: | Mondays from 1:00-4:00pm or by appointment |
| Office Phone: | (785) 864-5714 |

## Course Objectives, Materials, and Pre-Requisites:

There is a great demand for more useful, more actionable test scores. Traditional large-scale group administered tests do not provide this kind of information due to low reliabilities of, or high inter-correlations among, sub-scores. This course will explore approaches used by individually administered tests to provide diagnostic information, new psychometric models that hold promise of providing better diagnostic information, and implications for test design. A primary focus will be on how psychometric models can be used with diagnostic subscores that are more reliable and less correlated than traditional approaches.

### Required Textbook

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York: Guilford Press.

### Prerequisite

Prerequisite: PRE 922 or equivalent course.

### Statistical Computing

We will be using the FlexMIRT (Cai, 2013) package for estimating diagnostic models. We may also use some or all of the following programs: R, SAS, and Mplus.

FlexMIRT is available to you for free with an academic license. Details will be made available on the course website.

**Course Website/Technology**

As a hybrid course, video lectures are to be watched by students outside of class each week. Lectures will be made available on the course website. ***Please note that you will be expected to conduct approximately 9-12 hours outside of class each week.***

**This course will not use Blackboard for lectures or course announcements.** Instead, we will use freely available commercial software for communication and dissemination of course materials. Blackboard will be used for storing your individual grades.

***Audio/Video Recordings of Class***
I will be making a flash recording of each class, which will be posted on the website by the end of the day following class.

***Course Materials Over Dropbox***
Occasionally, files for the course will be disseminated over Dropbox (www.dropbox.com). **To gain access to the shared folder, please send me an email with the email address you use for dropbox.** You do not have to install the Dropbox application as you can download files from any web browser.

***Course Website***
Course lecture slides, links to course readings, lecture examples, flash video files, assignments, and information are available on the website. The website is http://wp.me/P3nkOf-qc. The website features the ability to post questions and comments about each lecture and homework where everyone can discuss the topic.

**Course Structure and Student Evaluation**

Student evaluation will be made on the basis of four components: (1) weekly quizzes on the readings, online lectures, and materials assigned each week, (2) simulation project performance, (3) empirical data analysis project performance, and (4) participation in weekly course discussions. **For all student work, please notify the instructor well in advance of any planned absences.**

***Weekly Quizzes*** (50% of course grade)
To ensure that assigned materials are read and watched, each week (except weeks one and two) we will have a weekly 5-item quiz that is administered in class. As the quiz accounts for 50% of your grade, please be prepared to answer questions on the materials assigned each week before coming to class. The lowest quiz grade will be dropped – and no quizzes can be made up if class is missed.

***Simulation Project*** (20% of course grade)
Throughout the semester, we will be working together on a simulation project that will become a paper to be submitted to a peer-reviewed journal. Information on this project will be disseminated in class. Tasks involved in the simulation will be assigned throughout the course and all students are expected to complete all tasks.

***Empirical Data Analysis Project*** (20% of course grade)
Throughout the semester, we will be working together on a data analysis project that will become a paper to be submitted to a peer-reviewed journal. Information on this project will be disseminated in class. Tasks involved in the simulation will be assigned throughout the course and all students are expected to complete all tasks.

***Class Participation*** (10% of course grade)
The final 10% of the course grade is for class participation. Students will earn these points for attending class and being active participants in course discussions each week.At the end of the semester, a weighted average percentage of points will be compiled and grades will be assigned according to the following system:

***Course Grading System***

| Percentage of Points | Grade |
|---|---|
| 100-93 | A |
| 92-90 | A- |
| 89-87 | B+ |
| 86-83 | B |
| 82-80 | B- |
| 79-77 | C+ |
| 76-73 | C |
| 72-70 | C- |
| 69-60 | D |
| Below 60 | F |

**Course Style and Content**

***Lecture Format***
As this is a hybrid course, each week students will watch online lectures outside of class. In class lectures will be roughly structured so that the first hour will be for discussing the online lectures and readings, the second hour will be for discussing the simulation project, and the third hour will be for discussing the empirical data analysis project. Time allocation may vary from week to week depending on course needs.

***Reading Assignments***
To be fully successful in this course, read the assigned papers and/or chapter(s) prior to the coming to class when we will cover the topic. Even if you have difficulty reading the material, exposure to the information prior to lecture will aid in your understanding of the course.

***How to Succeed in this Course***
- Read the assigned papers and chapters (even if it doesn't make sense to you – it will eventually)
- Come to class (and bring your questions about what you just read that week)
- Ask questions when you do not understand
- Come to office hours
- Be an active participant in all aspects of the course

**Tentative Schedule of Topics:**

Note: RTH refers to book by Rupp, Templin, and Henson (2010). Links to readings will be available on course website and will be accessible from KU campus networks.

| Date | # | Topic | Readings |
|---|---|---|---|
| 27-Aug-14 | 1 | Course Introduction and Psychometric Overview | None |
| 3-Sep-14 | 2 | Introduction to Diagnostic Models | RTH, Ch. 1 and Ch. 3<br>Templin, Bradshaw, and Paek (in press) |
| 10-Sep-14 | 3 | Attribute Specification | RTH, Ch. 4;<br>Gorin (2007); |
| 17-Sep-14 | 4 | The Loglinear Cognitive Diagnosis Model (or LCDM) | RTH, Ch. 6 - p. 112-115;<br>RTH, Ch. 7 - p. 144-158; Henson, Templin, and Willse (2009) |
| 24-Sep-14 | 5 | Diagnostic Structural Models | RTH, Ch. 8;<br>Templin and Henson (2006);<br>de la Torre and Douglas (2004) |
| 1-Oct-14 | 6 | The LCDM in Practice | Jurich and Bradshaw (2014); Bradshaw, Izsák, Templin, and Jacobson (2014) |
| 8-Oct-14 | 7 | Other Latent Class-Based DCMs | RTH, Ch. 5 and Ch. 6;<br>De La Torre (2011);<br>von Davier (2008) |
| 15-Oct-14 | 8 | Model Estimation in DCMs | RTH, Ch. 11;<br>Bartholomew and Knott (1999), Ch. 6;<br>de La Torre (2009) |
| 22-Oct-14 | 9 | Model Fit in DCMs | RTH, Ch. 12;<br>Maydeu-Olivares and Joe (2005);<br>Cai, Maydeu-Olivares, Coffman, and Thissen (2010);<br>Maydeu-Olivares & Joe (2014) |
| 29-Oct-14 | 10 | Attribute Estimation and Reliability | RTH, Ch. 10;<br>Templin and Bradshaw (2013) |
| 5-Nov-14 | 11 | DCM Item Discrimination and Computerized Adaptive Testing | RTH, Ch 13;<br>Cheng (2009);<br>Henson and Douglas (2005) |
| 12-Nov-14 | 12 | Extensions of the LCDM: Generalized Models for Differing Data Types | Bozard (2010);<br>Templin, Henson, Rupp, Jang, and Ahmed (2008)<br>Skrondal and Rabe-Hesketh (2004), Ch. 2 and Ch. 4 |

| | | | |
|---|---|---|---|
| 19-Nov-14 | 13 | Extensions of the LCDM:<br>Bifactor and Testlet DCMs | Bradshaw and Templin (in press);<br>Choi (2010);<br>Others TBA |
| 26-Nov-14 | | No Class - Thanksgiving | No Readings |
| 3-Dec-14 | 14 | Extensions of the LCDM: Attribute Hierarchies | Templin and Bradshaw (2014a);<br>von Davier and Haberman (2014);<br>Templin and Bradshaw (2014b) |
| 10-Dec-14 | 15 | Comparing DCMs and BINs: The DLM Psychometric Model | Wu (2013);<br>Others TBA |

## References


Bozard, J. L. (2010). *Invariance testing in diagnostic classification models*. Unpublished masters' thesis. The University of Georgia, Athens, GA.

Bradshaw, L., Izsák, A., Templin, J., & Jacobson, E. (2014). Diagnosing teachers' understandings of rational numbers: Building a multidimensional test within the diagnostic classification framework. *Educational measurement: Issues and practice*, *33*, 2-14.

Bradshaw, L. P., & Templin, J. (in press). Combining scaling and classification: A psychometric model for scaling ability and diagnosing misconceptions. *Psychometrika*.

Cai, L., Maydeu-Olivares, A., Coffman, D. L., & Thissen, D. (2006). Limited-information goodness-of-fit testing of item response theory models for sparse 2P tables. *British Journal of Mathematical and Statistical Psychology*, *59*, 173-194.

Cheng, Y. (2009). When cognitive diagnosis meets computerized adaptive testing: CD-CAT. *Psychometrika*, *74*, 619-632.

Choi, H. J. (2010). *A model that combines diagnostic classification assessment with mixture item response theory models* Unpublished doctoral dissertation, University of Georgia, Athens, GA.

de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, *34*, 115-130.

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*,*76*, 179-199.

de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, *69*, 333-353.

Gorin, J. S. (2007). Test construction and diagnostic testing. . In J. Leighton & M. J. Gierl (Eds.), *Cognitively diagnostic assessment for education: Theory and applications* (pp. 173-201). Cambridge, UK: Cambridge University Press.

Henson, R., & Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement, 29,* 262-277.

Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74,* 191-210.

Jurich, D. P., & Bradshaw, L. P. (2014). An illustration of diagnostic classification modeling in student learning outcomes assessment. *International Journal of Testing*, *14,* 49-72.

Maydeu-Olivares, A., & Joe, H. (2005). Limited-and full-information estimation and goodness-of-fit testing in 2 n contingency tables: a unified framework. *Journal of the American Statistical Association*, *100,* 1009-1020.

Maydeu-Olivares, A., & Joe, H. (2014). Assessing approximate fit in categorical data analysis. Multivariate Behavioral Research, *49,* 305-328.

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York: Guilford Press.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. New York: Chapman & Hall / CRC.

Templin, J., & Bradshaw, L. (2013). Measuring the reliability of diagnostic classification model examinee estimates. *Journal of Classification, 30,* 251-275.

Templin, J., & Bradshaw, L. (2014a). Hierarchical diagnostic classification models: a family of models for estimating and testing attribute hierarchies. *Psychometrika, 79*, 317-339.

Templin, J., & Bradshaw, L. (2014). The use and misuse of psychometric models. *Psychometrika, 79,* 347-354.

Templin, J., Bradshaw, L., & Paek, P. (in press). A comprehensive framework for integrating innovative psychometric methodology into educational research. In A. Izsák, J. T. Remmillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations.* Journal for Research in Mathematics Education monograph series. Reston, VA: National Council of Teachers of Mathematics.

Templin, J., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods, 11,* 287-305.

Templin, J., Henson, R., Rupp, A., Jang, E., & Ahmed, M. (2008). *Cognitive diagnosis models for nominal response data.* Paper presentation at the annual meeting of the National Council on Measurement in Education Society, New York, NY.

von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, *61*, 287-307.

von Davier, M., & Haberman, S. J. (2014). Hierarchical diagnostic classification models morphing into unidimensional 'diagnostic' classification models—a commentary *Psychometrika*, *79,* 340-346.

Wu, H. (2013). *A comparison of general diagnostic models (GDM) and Bayesian networks using a middle school mathematics test.* Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.