Instrument Design and Validation
Course Syllabus
FALL 2018

Course Prefix and Number: ORL 5524
Course Title: Instrument Design and Validation

Meets on: Mondays, 3:00-4:40 pm (ROOM: GDH 273A)

Instructor's Name, Address and Phone Number:
Madhabi Chatterji, Ph.D
Professor of Measurement, Evaluation and Education
Director, Assessment and Evaluation Research Initiative (AERI)
P.O. Box 6, Teachers College, Columbia University,
New York, NY 10027.
Office GDH 282
E-mail: mb1434@columbia.edu
Program Office: ZB 222; 212-678-3273

Office Hours:
Mondays 5-6:30 PM
Wednesdays 4:30-6:00 PM
Appointments preferred

Required Textbooks

1. Chatterji, M (2003). *Designing and Using Tools for Educational Assessment*. Boston, MA: Allyn & Bacon. *[EA*]
   THIS BOOK MAY BE PURCHASED VIA OUTLETS SUCH AS AMAZON.
   (NOTE: I may also share a few chapters from my forthcoming book, *Designing Assessments for Multidisciplinary Constructs and Diverse Populations: A User-centered Methodology*)

2. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, D.C.: Author. *[2014 Standards*]

This course will cover chapters 1-3, 5-9, 12-13 from *EA* (above) in depth, along with a selection of published studies as illustrative cases in instrument design and validation for class discussions and demonstrations.

Chapter 10 of EA, dealing with basic statistical concepts relevant to psychometric analyses, will not be treated in class. Please read it on your own to follow sessions dealing with particular measurement or psychometric techniques. Participants are encouraged to bring their own instrumentation needs, problems, examples or readings for discussion.

Recommended Supplementary Texts:

Depending on needs and student interests, I may refer certain segments of these two books.

Price, L.R.. (2018). *Psychometric Methods: Theory into Practice*. New York, NY: Guilford.

Bandalos, D. L.(2018). *Measurement Theory and Applications for the Social Sciences*. New York, NY: Guilford.

Targeted Audience:
Advanced graduate students (doctoral and 2nd year master's students) and professionals interested in designing assessments/ instruments for research and/or applied decision-making purposes.

Course Prerequisites:
Students should have completed ONE foundational course in measurement, quantitative research/evaluation methods, or statistics (a 4000 or 5000 level course at TC). Desirable: Completion of an intermediate level statistics

course, (e.g., multiple regression or general linear models).

Course Purposes:
The purpose of this course is to provide students with basic measurement theory, coupled with applied experiences in designing tests, survey-based scales, and other types of assessment instruments and validating construct measures they yield. Ideally, students should come to class with an idea for an instrument, test or scale that they would like to design tapping into educational, psychological, health, social or other constructs. The type of instrument can vary according to student needs and interests. In the past, participants have designed attitude scales; multi-construct surveys; behavior-based assessments; portfolio-based assessments; constructed response tasks or multiple choice assessments of cognitive abilities.

To view examples of tools designed by prior students, please follow the link to the computer module below. Look under PROCESS MODEL-APPLICATIONS under each User Path:
http://www.columbia.edu/~mb1434/EdAssess.htm

Course Objectives/Expected Learning Outcomes:
Following course completion, students/participants should be able to:

1. Specify the *assessment context* for given instrument design scenarios:
   a. identify construct(s) to be measured,
   b. identify target populations,
   c. identify the assessment purposes and inferences to be made from scores/construct measures.
2. Identify the defining features, advantages and disadvantages of different *instrument types* and assessment modalities:
   -Self-report instruments, survey-based scales and questionnaires
   -Structured and constructed response "paper and pencil" tests and assessments
   -Performance assessments (behavior-, product- or portfolio-based assessments)
   -Technology-based, technology-mediated and computer adaptive assessments
   -Other instrument types
3. Apply *key concepts from Classical Test Theory* (CTT) when designing and validating construct measures:
   -constructs, domains, domain specifications, domain sampling
   -item or task pool
   -test blueprints; assessment design specifications
   -levels of measurement; properties of nominal, ordinal, interval and ratio scales
   -raw scores, derived scores, composite scores, group(ed) scores
   -validity, reliability, and utility;
   -types of validity and reliability evidence; empirical methods for evaluating item quality, validity and reliability
   - the Process Model for assessment design and validation
   -"use" or "user"-centered assessment design with the Process Model.
4. Specify *construct domains* with observable indicators, drawing on appropriate literature and alternate data sources.
5. Apply established guidelines to *write/select items or tasks*, produce an item pool, develop scoring protocols and rubrics, and assemble an instrument to serve a specified need.
6. Develop a *validation plan* for assuring the quality of items, overall instrument and construct measures.
7. Gather, analyze and interpret relevant kinds of *evidence to evaluate the quality of items, overall instrument and construct measures* produced (validity, reliability and utility evaluations).
8. Follow academic guidelines to *report results of a small- scale pilot-testing and/or validation* of items, overall instrument, and construct measures. [If data from large scale validation studies are available, that may be used.]

Course Delivery:
ORL 5524 is an interactive lecture-seminar. As you develop skills in assessment design and validation, there will be many opportunities to get constructive, "friendly" feedback from colleagues and your instructor. **Please attend all class meetings and do the necessary background reading and homework. Your participation in class discussions and interactions are the key to making the experience a worthwhile one for you.** (We will take

attendance, and no more than three (3) excused absences are allowed without penalty).

Course Requirements: The course is offered for 3 credits.

Main Assignment:
Apply an iterative "Process Model" (introduced in the course) to develop and validate an instrument for meeting a specified research or practice-based need.

Interim products	Points
1. Specify the assessment context and construct domains to be measured	10
2. Develop an initial item pool and draft of the instrument	10
3. Content-validate and refine the domains and items iteratively	20
Mid-term product: Items 1-3 above after one revision/iteration	
4. Develop a validation research plan, showing short- and long term goals	10
5. Present preliminary validation results in class followed by final, technical report based on in-class feedback	40
(Report should include at least a content validation study or small scale pilot)	
Final report should synthesize Items 1-5 above.	
Mid-term and Final Report should be submitted by agreed-upon deadlines.	

NOTE: A scoring rubric will be distributed at an appropriate point to guide your work, with examples of reports prepared by previous students for refernce.

Grading Criteria:
The course grade (Max=100 points) will be based on the mid-term and final products, weighted as shown above. Within the broad categories, finer grade distinctions may be made based on students' performance distributions.

A	90-100%
B	80-89%
C	70-79%

Journals where you will find reports of instrument development and construct validation studies.
1. Educational and Psychological Measurement. (See Validity section.)
2. Psychological Assessment
3. Educational Assessment
4. Journal of Applied Measurement
5. Applied Measurement in Education
6. The International Journal of Educational and Psychological Assessment
7. Journals in substantive fields (e.g., Language Testing; Measurement in Counseling; Measurement in Nursing, Contemporary Educational Psychology, Academic Medicine) also publish instrumentation and validity studies.

I may share my own published studies on instrument development and validation projects as illustrations.

INDIVIDUALS WITH DISABILITIES

The College will make reasonable accommodations for persons with documented disabilities. Students are encouraged to contact the Office of Access and Services for Individuals with Disabilities for information about registration (166 Thorndike Hall). The College states that services are available only to students who are registered and submit the appropriate documentation. As your instructor I will be happy to discuss specific needs with you as well.

Religious Observances:

The College policy on honoring religious holidays will be followed. Please keep me informed.

We will follow the College's guidelines and no-tolerance policies on plagiarism.

Weekly Session Outline: See attached general outline, which may change slightly depending on the class' pace.

Session and Day		**Topics**	**Readings**
1 1/27		**Foundational Concepts and Terms**: Testing, measurement, assessment and evaluation "Constructs" , "Concepts" and Variables (What we measure) "Construct measures"-raw scores, derived scores, composite scores, grouped scores Assessment Operations, Operational definitions of constructs Types of assessment devices/ instruments Measurement levels: nominal, ordinal, interval , ratio scales A Process Model for instrument design and validation **Applying Phase 1 of the Process Model** **-Specifying the Assessment Context (Exercise/homework)**	Chapters 1- 5 (EA) Chapter 10-Scales of Measurement (EA, pp. 290-294, Table 10.1) *2014 Standards* pp.75-83, 95-100 Examples and assessment cases
2 2/3		Validation -A definition Evidence of Validity, Reliability and Utility Measure-based Inferences Assessment Purposes-Specific Uses, Assessment Users User-centered design--*tying the desired interpretations and uses of test scores or measures to the design/validation process and needs of users* Evaluating validity in relation to measure-based inferences and uses **Homework review: Phase 1-Specifying the Assessment Context** **Assignment 1: Specifying the assessment context for your own instrument (Posted on Canvas. Draft DUE next week for feedback)**	Chapters 3, 13 (EA) *2014 Standards* pp.75-83 95-101 Examples and assessment cases
3 2/10		**Validity and types of validity evidence** Trinitarian (old) vs. unified (current) view of validity and validation Outline of an instrument design and validation study Illustraton of validation plan-Study 1 (Chatterji & Lin) **Assignment 1 review: Specifying the assessment context for your own instrument–Iteration 1**	Chapters 3, 13 (EA) *2014 Standards* pp.11-22, 49-62 Chatterji & Lin (2018) Wyer & Chatterji (2013)
4 2/17		Validity and types of validity evidence (continued) Reliability and types of reliability evidence User-centered validation, consequential validity, argument-based validation Illustration of validation plan- Study 2 (Beaudreau et al) Illustration of a validation plan-Study 3 (student example) **Exercise/homework: Developing a validation plan (Posted on Canvas. Due next week for feedback)**	*2014 Standards*, pp. 11-22. 33-41 Chapters 3, 13, 14 (EA) [Enrichment: Kane, 2006] *Pittsburg Sleep Quality Inventory*- Beaudreau et al (2012) study
5 2/24		Reliability and types of reliability evidence (continued) Utility evidence **Class review and feedback: Developing a validation plan**	*2014 Standards*, pp. Chapters 3, 13, 14 (EA)
6 3/2		Applying Phase II of the Process Model: Specifying the construct domain—methods, data sources, taxonomies, guidelines/conventions Specifying the domain for a cognitive construct- (Illustrative student example) **Exercise/homework: Specifying the domains for measuring "driving proficiency" OR "attitude towards**	Chapters 5-6 (EA) [Enrichment: Chapter 4 of new book)

		safe driving" (automobiles)	
7 3/9		Applying Phase II of the Process Model Specifying the construct domain—methods, data sources, taxonomies, guidelines/conventions (continued) Specifying the domain for a non-cognitive and behavioral constructs (Illustrative student examples) Content-validating domains -Feedback and class discussion of homework **Assignment 2: Specifying the domain(s) for your chosen constructs-send in drafts by 3.22 for class feedback on 3/23** **Mid-term instructions and rubric posted on Canvas.**	Chapters 5-9 (EA)
8 3/16		SPRING BREAK	
9 3/23		Applying Phase III of the Process Model. Designing multiple choice items and constructed response tasks for cognitive domains Item writing, guidelines and examples, designing scoring protocols and keys, critiquing and content validation of items, instrument assembly **Item writing practice** **Feedback on domains and class discussion (Assignment 2-drafts)**	Chapters 5-9 (EA) (For enrichment: Chapters 5-7 of new book)
10 3/30		Applying Phase III of the Process Model. Designing structured self-report (survey) items for non-cognitive and behavioral domains Designing performance assessments Item/task writing, guidelines and examples, designing scoring protocols and rubrics, critiquing and content validation of items/tasks, instrument assembly **Mid-term DUE**	Chapters 5-9 (EA) (For enrichment: Chapters 5-7 of new book)
11 4/6		Applying Phase IV of the Process Model-Empirical Validation and Instrument Refinements Content validation studies Content Validity Indices (CVI) Validity evidence of response processes Conducting cognitive interviews to validate response processes	Chapters 5-9 (EA) Lynn (1986) paper Polit et al (2007) paper Carbone et al (2002) paper
12 4/13		Applying Phase IV of the Process Model-Empirical Validation and Instrument Refinements Item analysis studies Item statistics for norm-referenced and criterion-referenced tests Item statistics for self-report (survey) items	Chapter 12 (EA) Chatterji & Lin (2018)
13 4/20		**Formal review and feedback of your domains, items and item pool by peers in class [Remember to bring your work to class]** Possible: Guest Speaker—Illustrative study Begin pilot –testing or validation studies –on your own (AERA/NCME conference week-Prof. Chatterji will be away, but Xiaoxue Du will hold class)	

14 4/27		Applying Phase IV of the Process Model-Empirical Validation and Iterative Instrument Refinements (continued) Validity evidence of internal structure: Exploratory factor analysis Confirmatory Factor Analysis	Chapter 13 (EA) Benson (1998) Chatterji & Lin (2018)
15 5/4 (Last day of class)		Project presentations (required for all enrolled students) Discussants assigned to provide constructive feedback to individual presenters (Attendance required for all)	
16 5/8		**Final papers DUE by 5/8/20, 5:00 PM (non-negotiable deadline). Papers may be emailed to the instructor or Teaching Assistant via Canvas. Marked papers may be picked up from Prof. Chatterji or Xiaoxue Du by appointment after grades are posted. Thanks!**	