Course Overview

“İ often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be.”

“Whatever exists, exists in some amount.”   E. L. Thorndike.

Measurement in survey research is both art and science. While much of the science can be learned from textbooks and articles, the art can only be learned through experience. The principal objective of this course is to give students an introduction to fundamental concepts of measurement though a semester-long project in which students are expected to develop, pilot test, analyze and evaluate their own survey instruments. Though surveys can be used to measure both manifest variables (i.e., factual information) and latent constructs (i.e., information residing inside a person’s head), in this course the empirical focus will be on developing a survey instrument that can measure at least one latent construct.

This course emphasizes the process of developing, analyzing and validating a survey instrument. The concept of validity is central. The most comprehensive definition of the term validity in the context of measurement was established by Samuel Messick in the opening paragraph to his chapter (“Validity”) in the 3rd edition of the book Educational Measurement (1989):

Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy an appropriateness of inferences and actions based on test scores or other modes of assessment…the term test score is used generically here in the broadest sense to mean any observed consistency, not just on tests as ordinarily conceived but on any means of observing or documenting consistent behaviors or attributes. Broadly speaking, validity is an inductive summary of both the existing evidence for and the potential consequences of score interpretation and use.
The first half of the course focuses on the underappreciated task of developing a survey instrument with items that derive from a clearly delineated theory for the construct to be measured. By the ninth week of the course, students (working either alone or in small groups no larger than 3) should have a survey instrument ready for pilot testing with a minimum of 30 respondents. By the beginning of November students should have results from their pilot test. In the second half of the course, the focus shifts to the analysis of item responses. Students will be introduced to classical test theory and item response theory as two modeling approaches used for this purpose, with a particular emphasis placed on item response theory models that fall within the Rasch family of measurement models. The culmination of the course will be putting together a validity argument to support the proposed uses of the survey. In so doing, students will be introduced to some historical and recent developments in validity theory.

This course serves as prerequisite for the course “Advanced Topics in Measurement” (offered in the spring of 2010) which focuses more exclusively on analytical and applied issues related to psychometric item response modeling.

**Important Due Dates**

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/14</td>
<td>Proposal for Class Project Uploaded to CULearn</td>
</tr>
<tr>
<td>9/21</td>
<td>Target date for feedback on proposal</td>
</tr>
<tr>
<td>10/19 to 10/31</td>
<td>Draft of 1st Half of Project Report Uploaded to CULearn</td>
</tr>
<tr>
<td>11/1 to 11/14</td>
<td>Feedback on 1st Drafts</td>
</tr>
<tr>
<td>12/14</td>
<td>Full Project Draft Due</td>
</tr>
</tbody>
</table>
Course Schedule

WEEK 1: Introduction and Overview                         Aug 24

- Course Expectations and Guidelines
- A Framework for Survey Research
- Presentations of two projects from the Fall 2007 Class

Assignments:
1. Handout: What research question(s) will your survey be used to address?

2. Visit HRC website (http://www.colorado.edu/VCREsearch/HRC/index.html)
   If this is your first time, go to
   http://www.colorado.edu/VCREsearch/HRC/First_Timer.html
   Complete the CITI course. Otherwise, make sure you have read
   http://www.colorado.edu/VCREsearch/HRC/FAQs_for_Investigators.html

WEEK 2: General Issues in Survey Research                  Aug 31

- Examples of Large-Scale Survey Instruments
- Introducing the Concepts of Validity and Reliability
- Sources of Error: Sampling and Measurement
- Different Theories of Measurement
- An Overview of Wilson’s Building Blocks for Measurement
- Establishing the Research Context

Readings: Groves et al (Chapter 1); Sapsford (Chapter 2); Wilson (Ch. 1); Kane (1992); Michell (1986)

Note re. the Michell reading. This paper focuses on empirical and philosophical differences in paradigms for what is meant when we speak of “measurement.” But this will take some work to follow. Consider this optional, but recommended.

Assignments:
Complete any two of Exercises 5-7 on pp. 36-37 of the Groves et al chapter. [Due 8/31]
Proposal for Survey Project [Due Sep 21]

WEEK 3: Labor Day Holiday—NO CLASS                         Sep 7
WEEK 4: Sampling in Survey Research  
Sep 14

- Being Clear about the Population of Interest
- Developing a Sampling Frame
- Probability Samples, Pilot Samples, & Convenience Samples
- Sampling Weights

Readings: Groves et al (Chapters 2-3); Sapsford (Chapters 3-4); Dillman et al (2002)

Proposal for Class Project Due

WEEK 5: Construct Maps  
Sep 21

- Construct Maps as the Instantiation of a Theory
- Examples of Construct Maps
- Using Construct Maps to Develop a Learning Progression

Reading: Wilson, Ch. 2, Masters & Forster; Briggs et al, 2006; Stevens et al, 2009

Want to read more about recent work being done on learning progressions? Check out http://www.education.uiowa.edu/projects/leaps/proceedings/Default.aspx

Also see http://www.cpre.org/index.php?option=com_content&task=view&id=282&Itemid=149

Assignment:
Wilson, Ch. 2: Exercises 1-4 [Creating a first draft of a construct map: Due 9/28]

WEEK 6: You Gotta Know What You Want to Measure…  
Sep 28

- Instrument Design with Cognition in Mind
- Using a Construct Map to Inform Item Design
- The Demand for Cognitive Diagnostic Assessments
- Evidence-Centered Instrument Design

Readings: Mislevy, Steinberg & Almond, 2002; Gorin, 2006; Briggs & Alonzo, 2009a; Huff & Goodman, 2006

Wilson, Ch. 2: Exercises 1-4 Due
WEEK 7: Asking the Right Questions (i.e., Designing & Scoring Items)     Oct 5

• A Taxonomy of Item Formats
• Working with Likert Items
• From Construct Maps to Scoring Rubrics
• Phenomenographic approaches
• The SOLO Taxonomy

Required Readings: Wilson, Ch. 3-4; Fowler, (1995, Ch. 3), Patton, (1980); Thurstone (1928)

Conditional Readings:
If you are thinking about using Likert items, read
  • Oppenheim, (1992, pp. 174-209) and Likert (1932)
If you are thinking about using Multiple-Choice items, read
  • Nitko, (1983)
If you are thinking about using open-ended items, read
  • Matron (1988)
If you are trying to create a learning progression, read
  • Biggs & Collis (1982)

Note: If you are a student in the school of education in particular, I highly recommend reading the Biggs & Collis piece regardless of its immediate fit with your project. Very interesting stuff.

Assignment:
Wilson, Ch. 3: Exercises 1-2 [Designing items for your instrument: Due Oct 12]

WEEK 8: Critiquing Items: The Item Panel     Oct 12

• Think-alouds
• Cognitive Interviews
• Context Effects
• Running an Effective Item Panel

Readings: Sudman, Bradburn & Schwartz (1996), Ch. 2 & 4; Wills (2005), Ch 5

Conditional Readings:
If you are developing Likert items to measure an affective latent construct:
  • Sudman, Bradburn & Schwartz (1996), Ch. 5
If you are developing multiple-choice items to measure a cognitive latent construct:
  • Sudman, Bradburn & Schwartz (1996), Ch. 6

Wilson, Ch. 3: Exercises 1-2 Due
Assignments:
Wilson, Ch. 4, Exercises 1-2 [Relating item responses to the construct map: Due Oct 19]
Conduct Item Panel [i.e., Wilson, p. 59-61]
Strongly Recommended: Conduct Think-Aloud

WEEK 9: Classical Test Theory and Classical Item Analysis Oct 19

- Observed Scores, True Scores and Error
- The Concept of Reliability and How it is Estimated in Practice
- Item P values and point-biserials

Note: At this point in the course you should have a draft of your survey instrument that is ready to be pilot tested.

Readings:
Wainer & Thissen (2001), [23-34, 52-54, 57-59, 64-70]
Crocker & Algina (1986) [311-338]
Thompson, (2003)

**First Draft of First Half of Course Project Due between 10/19 and 10/31**

WEEK 10: Item Response Theory and The Rasch Model Oct 26

- Probability, odds and log odds (logits)
- Item Characteristic Curves
- The Wright Map

Readings: Wilson, Ch. 5-6; Hambleton, Swaminathan & Rogers (1991), Ch. 1-3

Assignment:
Wilson, Ch. 5, Exercises 2-3 [Calculating response probabilities from the Wright Map]

WEEK 11: Measurement Error & Model Fit Nov 2

- The Standard Error of Measurement
- Item and Test Information Curves
- Misfit

Readings: Wilson, Ch. 7; Bond & Fox, Ch. 12

Note: At this point in the course you should be ready to begin analyzing data from the pilot test of your group’s survey instrument.
WEEK 12: Using ConstructMap to Analyze Item Responses (Part 1)  Nov 9

- Importing item response data into ConstructMap
- Running a classical item analysis
- Running the Simple Rasch Model
- Running the Rating Scale Model & Partial Credit Model
- Producing a Wright Map

Readings: ConstructMap Lite User Manual

Above and beyond (optional): Read Andrich (1978) and Masters (1982) for the original presentations of the Rating Scale and Partial Credit Model extensions of the Rasch Model.

Assignment:
Produce classical item statistics for your data using ConstructMap and interpret the output

WEEK 13: Using ConstructMap to Analyze Item Responses (Part 2)  Nov 16

- Plotting Item Characteristic Curves
- Assessing Item and Person Fit
- Examining Item and Test Information Curves

Readings: Wilson, Ch. 8

Assignment:
Fit your data with the appropriate item response model (Simple Rasch, Rating Scale, Partial Credit) using ConstructMap and interpret the output

WEEK 14: FALL BREAK—NO CLASS  Nov 23

WEEK 15: Did the Results Support Your Hypothesis?  Nov 30

- Sources of evidence for a validity argument
- Integrating evidence into a coherent argument
- Connecting a Wright Map back to a Construct Map

Readings: Kennedy & Wilson (2007); Briggs & Alonzo (2009b)
WEEK 16: Validity Theory

- The evolution of validity since the 1950s.
- The Standards for Validity


FINAL PROJECT DUE 12/14 by 8:00 pm
COURSE READINGS

Books to be Purchased


Optional but Recommended


Readings Available Electronically (CULearn)


Survey Project: The Nitty Gritty Details

Preamble

All survey instruments contain items, and each item, or collection of items on the survey is intended to measure something. In some cases the thing to be measured is a manifest variable and a single item may suffice to measure it; for example, we can measure a person’s age by asking for his or her date of birth. In other cases the thing to be measured is a latent variable and this will necessitate a set of items to measure it. For example, we might want to measure a person’s ability to reason quantitatively by administering the collection of math items found on the SAT. Of course, in many cases surveys are designed to measure both manifest and latent variables. In social science research, it can be difficult to conceptualize an important research question that does not necessitate the measurement of a latent variable. Latent variables come in two flavors: affective and cognitive. Affective latent variables would include a person’s attitude toward some topic, their motivation, or their self-confidence. Cognitive latent variables focus on a person’s knowledge, skills and understandings within or across some particular content domain. The term “construct” is often used as a synonym for a latent variable, and this is how I will be using it throughout the course.

Conditions

Your task over the 16 week period of this course will be to design, develop, pilot test, analyze and evaluate a survey instrument. Here are some conditions:

1. Your survey instrument must measure at least one latent construct, either affective or cognitive in nature.
2. You will need to have a MINIMUM of 30 respondents participating in the pilot test of your survey. (The more the better.)
3. Working in groups is fine, but each student has to turn in an independent project report.

STAGE 1: Project Proposal (No more than 10 double-spaced pages)

Your project proposal should address the following issues:

1. Specifying the underlying use for the Survey

All surveys are developed with some use in mind, so the starting point for your project will be to specify the proposed use(s) of a survey designed to measure at least one latent construct. One good way to do this is to write down one or more research questions that your survey instrument could help you answer. Here are some examples:

• How well do students understand Newton’s 2nd law of motion by the end of a semester-long introductory physics class?
• To what extent does a student’s understanding on Newton’s 2\textsuperscript{nd} law of motion change over the course of a semester while enrolled in an introductory physics class?
• What is the effect of an innovative lab structure on a student’s understanding on Newton’s 2\textsuperscript{nd} law of motion?

Sometimes survey “use” is hard to conceptualize in terms of a research question. Here is a different approach:

• The results from this survey will be used by the teacher to diagnose levels of student understanding…
• The results from this survey will be used to help the teacher make changes to the curriculum…
• Each year two scholarships are awarded to our best students. This survey will be used to identify the two best students in class.

2. Why do you need to design a new survey instrument? If one already exists, why doesn’t it meet your needs?

Once you have specified the use or uses for your hypothetical survey instrument, you will need to explain why existing survey instruments will not suffice. This will entail some investigation on your part—it will seldom be the case that no one has ever tried to measure the same construct. How will you be improving upon what has been done in the past? One resource that might be helpful is the *Mental Measurements Yearbook*, available at the Chinook website. Online searches using Google and ERIC can also be helpful. You need to really beat the bushes here! The foundation for all research is a solid literature review, and this is no different here.

You may find that someone has developed what looks like the perfect survey to measure the latent construct you have in mind. In this case, it would be entirely reasonable for you to put together a project in which you attempt to validate a pre-existing survey with new data.

3. Is there a pre-existing theory (or theories) that inform(s) the construct you wish to measure?

In this project your construct of measurement will be some cognitive or affective latent variable. Imagine the variable in terms of a unidimensional continuum. Some survey respondents will be high on the continuum, some will be low, some will fall in between. Has anyone come up with a theory that explains why? Does anyone have a theory that explains how people get from a low point on the continuum to a high point? If your construct is embedded in the content of a curriculum, is there a theory that links the instructional activities of the curriculum to a student’s hypothesized development in terms of the construct of interest? Do pre-existing theories contradict the idea that your construct of measurement can be conceptualized as a unidimensional continuum?
4. Who are some experts that can/will help you with your survey development?

These should be people with experience and/or knowledge working in the content domain you have chosen for your construct of measurement. For example, if the domain was Chemistry, your experts should have a background in Chemistry.

5. How are you defining the target population for your survey? What is your sampling frame? What is your practical plan for getting a sample of at least 30 pilot test respondents?

**STAGE 2: Obtain HRC Approval for your Survey**

Each survey project for this class should qualify for an “Exempt Review” to the extent that it constitutes

- Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
  1. Information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects.
  2. AND any disclosure of the human subjects’ responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects’ financial standing, employability, or reputation.

You will still need to apply for an Exemption under CU’s HRC. Getting this in a timely manner should be no problem, but you will need to develop a consent form to be signed by all survey participants. For details, go to

[http://www.colorado.edu/VCResearch/HRC/index.html](http://www.colorado.edu/VCResearch/HRC/index.html)

**STAGE 3: Survey Development**

This stage corresponds with three of the four “building blocks of measurement” described by Mark Wilson in Chapter 2-4 of his textbook Constructing Measures:

- Developing a Construct Map (Chapter 2)
- Designing Survey Items (Chapter 3)
- Scoring Survey Items (Chapter 4)

The culmination of this stage will be holding an item review panel with your experts, and conducting think-alouds or cognitive interviews with a subsample of your target respondents.

At this point you should be ready to turn in the first half of your project report for feedback.

**STAGE 4: Pilot Testing of Survey Instrument**

This should take place no later than November 1st.
STAGE 5: Analysis of Survey Item Responses

This stage constitutes the fourth “building block of measurement.” In this stage you will be applying a Rasch Measurement Model to your items responses using the (free) software “ConstructMap” (formerly “GradeMap”).

STAGE 6: Building a Validity Argument for your Survey Instrument

You will be expected to use the evidence gathered in stages 1-5 above to build a preliminary argument to support the validity of the survey for its proposed use. In other words, you will need to evaluate your survey. There will, of course, probably be many holes in the validity argument at this point. The objective is to identify these weaknesses such that revisions to the survey can be made, and a new plan for gathering additional evidence to support a stronger validity argument can be established.

Software for Administering and Analyzing Survey Responses

Administering Surveys

In the old days, surveys were administered in hard copy form and item responses had to be transcribed by hand. These days, there are some wonderful internet based programs that can make the act of gathering items responses much more convenient. Here are two sources that students have used in the past:

http://www.questionpro.com/
http://www.zoomerang.com/

At the time of this writing I’m working on another high-end possibility that would probably be the most ideal (stay tuned)

www.qualtrics.com

Of course, you may find that an “old-school” paper and pencil survey is still the best way to go…

Analyzing Survey Responses

We will be using the software ConstructMap to analyze item responses using both classical item statistics, and output based on a calibration of responses using a Rasch Measurement Model. You can download the latest version of ConstructMap (along with a user’s manual) at http://bearcenter.berkeley.edu/GradeMap/

We will be spending two class sessions working with ConstructMap.
Grading Policy

Assessment for EDUC 8710 will consist of

(75%) A research project, which will consist of the development and initial evaluation (pilot study) of a survey instrument. An outline of the proposed project will be developed by the student in conjunction with the instructor. The final project incorporates numerous assignments from throughout the semester:

1. Developing the theory behind your survey instrument (i.e., defining the construct of measurement in terms of a construct map).
2. Designing and critiquing items.
3. Collecting data and analyzing the instrument’s measurement properties.

If you do all this well (I’ll be more explicit about what I mean by well later in the semester) you’ll get at least a B+ on the project. To get an A, I want an empirically-based answer to a deceivingly simple question: What needs to be done to improve this survey instrument for future use?

A general outline for the final report can be found on p. 84 of the Wilson text, but I will be handing out my own, more detailed outline.

(25%) Participation in classroom discussions, small-group assignments and activities. This will include minor presentations by class members throughout the semester. It is expected that all students will complete assigned readings and exercises before each class.

CULearn

I will try to make all presentations, handouts, data sets, etc. available to you on CULearn. You can access this by going to https://culearn.colorado.edu and logging in. I recommend that you check this site within about 15 minutes of the start of each class for any presentation slides I may have posted for your convenience.

When you upload assignments via CULearn please use the following naming conventions for your files: “lastname_assignment_date.doc”
(i.e., Briggs_projectproposal_101909.doc)

Missing Classes

Sometime life intrudes and you will be forced to miss a class. In this event, it is your responsibility to take the initiative to get back up to speed. Please don’t assume that I will seek you out electronically or in person to provide you with materials or announcements you may have missed.
**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_relig.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 polices 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/