07.709 Measurement and Evaluation

Time: Thursday 4:00 – 7:00  Room: As assigned

Instructor: Dr. James Carifio  Office: 527 O’leary Library
Phone: 978-934-4609  email: James_Carifio@uml.edu
Office Hours:  Mon, Weds, Thursday 3-4, after class and by appointment and email

Required Texts


4. Class handouts, articles, and FREE-ON-LINE books, articles, software and helpful web sites: You will need to email me at the following email address (James_Carifio709@yahoo.com), and I then email you back an electronic form of the list of on-line books and sites I will hand out in class, so you will not have to type them all in. Emailing me will start the process of your receiving electronic handouts from me and post-class follow-up memos as well as my answering your questions and your turning in certain assignments to me. The Yahoo address is the address for doing “course business” to keep my university email uncluttered.

Optional Texts


Some Helpful Texts

Cronbach, L. Essential of Psychological Testing, Harper-Row
Popham, J. Modern Educational Measurement, Pearson
Anderson and Krathwohl. A Taxonomy for Learning, Teaching and Assessing. Pearson
Haladyna, Thomas. Writing Test Items to Evaluate Higher Order Thinking. Allyn and Bacon
Stiggins, Richard. Student Involved Classroom Assessment. Merrill Prentice-Hall
Ferguson, Statistical Methods in Education and Psychology
Merler, C. and Vannatta, R. Advanced and Multivariate Statistical Methods, Pryczak Publishing.
A separate list of journals (there are many).
Several UML dissertations (see separate list).
Overview

This course will survey basic measurement and evaluation theories, models, techniques, protocols, and systems. The survey will include the assessment of achievement, attitudes, opinions, abilities, personality, skills, and traits, and various kinds of environmental, process and developmental variables, as well as macro variables such as programs, schools, communities, and state or federal assessment and testing systems. The topics surveyed and amount of time spent on each topic will be tailored to match the interest and needs of the student taking the course.

One of the course’s primary objective is to help doctoral students and educational practitioners and managers evaluate, choose, modify or and develop appropriate instrumentation for their dissertation, or research, evaluation or assessment project, whether it is at the individual, class, school, district, state or national level. Another focus of this course will be on how one develops a reliable and valid measure of any kind (qualitative or quantitative) for any purpose (individual, group, program, or policy) and various statistical and non-statistical techniques that may be used for examining, assessing, and improving the reliability and validity of any measure. Aspects of program and policy evaluation will also be examined as well as observational systems and dynamic and tailored measurement and assessment techniques and systems.

Classical measurement theory, as well as performance, criterion, authentic, portfolio, and Rasch-type (trait free) theories, will be examined as needed in the course, particularly relative to the assessment of upper level cognitions and other complex behaviors and productions. The course will also examine qualitative techniques for use with “qualitative research” and assessment, such as scoring rubrics, protocols, analytical rating scales, observational systems, projectives, interviews, concept maps, content analysis, logs, journals, essays and the analysis of existing records and documents of various kinds. Emphasis will be placed on the scoring of responses or materials and the aggregation (indexing), interpretation and analysis of these systematically derived and/or transformed measurements.

The actual and specific content and techniques included and focused on in detail in the course will be adapted to the needs and interests of the students enrolled in the course. The first seven sessions will be core theory and models and the last 8 sessions will be tailored to the selected learning objectives and topics of the class.

Assessment and Grading

Formal assessments and the weight they will have in determining your grade will be as follows:

1. An in-depth written review of at least two existing measurement procedures, instruments, or systems will be required (20%).

2. Each student will have to develop and revise one prototype instrument (15%).

3. Each student must provide reviews and formal criticisms to their classmates on their prototype or formal instrument during the development process or their in-depth paper (10%) 

4. Two short papers or take-home mini-projects/problems will also be given, one of which may be the analysis of existing data on an instrument or procedure or a review of another students prototype or formal instrument (20%).
5. A final project is required which may be an in-depth paper on a specific measurement topic, procedure or area, or a pilot testing of the prototype instrument developed and revised including statistical analysis and evaluation of the instrument. An alternative to piloting your prototype instrument is my making available to you the pilot data from an instrument that has been developed along with the instruments blueprint and your doing the statistical analyses that are needed to validate the instrument and assess its quality and writing up these analyses (30%).

6. A presentation of your final project to your peers and me in the last class is also required (5%).

Separate handouts on the details and scoring criteria for the above assignments will be given out in class.

Computations and Computers

Most computations will be done with Excel, SPSS, Stat-View or some similar type of software although many will be done by hand or calculator either for convenience or because the software does not do them and they are important to do to complete analyses. **At a minimum, I will expect you to be able to interpret the output (results) of certain analyses (such as ANOVA) from some software programs (such as SPSS) and know how to read basic tables (such as an ANOVA table) in research articles and to make such tables which will require you knowing certain aspects of APA (reporting) formats. I will also expect you to be able to summarize the results reported in tables in acceptable professional English.** It would be nice if you learned how to develop a data file in Excel or SPSS and to select the appropriate analyses to run from the available menus, but having this operational knowledge and not knowing and understanding the conceptual and theoretical content of the course would be more than self-defeating, particularly in the long run. Once you learn the conceptual and theoretical content of this course, you can learn how to use various software packages as well as use them more intelligently even if this learning happens after this course. So if it is a choice of where to spend your limited time in this course, spend it on the conceptual and theoretical content because (1) I am telling you that is the most important content to know and the knowledge with the longest half-life, and (2) what will be weighted most highly in my evaluation of your work. Many resources, however, will be provided for you to learn the rudiments of using the most commonly used software packages.

Course Ground Rules, Expectations and Your Responsibilities

There is a separate handout on these three items that we will go over and discuss during the first class so that each is explicit and clear to you. In general, you need to understand that this content is **highly cumulative** and the concepts, principles, ideas, facts and details are **highly interrelated** and adapted in numerous ways. You **cannot cram this content** and must learn it day by day in a systematic fashion and this approach will need to be your learning style in this course. Next, you **cannot just be ‘acquainted’ with this content** as you will not only misunderstand it and miss key and critical points, but you will not be able to apply the content or think critically using it. You, therefore, need to learn each concept, principle, technique and idea with **some depth and sophistication** (i.e., at the higher levels of Bloom’s taxonomy) and synthesize, extend and **transfer your understandings** to different contexts and situations. You will also need to be meticulous and learn the precise, specialized and nuanced meanings of many words and new terms. Further, you will need pay attention to details, and **what is not there or said in a test manual, article or study** as the “silences” are often more important than the “sounds” in research (just like poetry). If research is observing and reflecting, then you need to
be actively observant at all times and not tuned out or unengaged. I expect you to be an independent, self-directed, self-managing, adult learner, and I actually expect you to work at teaching yourself (and others) this content. I also expect you to do the reading and exercises and to come to class prepared (if you come, as I do not require attendance). All of these things are our ‘covenant’ in this course. If you do not break covenant with me, I will keep re-explaining things and explaining them in different ways until you understand them sufficiently and to hold you harmless for any and all misunderstandings and mistakes. As Einstein once said, “If we knew what we were doing, they wouldn’t call it research [or the synthetic-experimental approach or learning].” Given these points, one of the most efficient and effective ways of my assessing whether you are meeting your responsibilities in this course and the goals of this course is to use the Socratic teaching method. I will call on you directly in class and pepper you with questions, so please, do not take it personally or be overly frightened or feel personally put upon by me. The Socratic method is just one of the most efficient and dynamic ways of measuring and verifying your knowledge and understanding and the degree that you are keeping covenant with me.

**SYLLABUS**

**J31:** Overview of Measurement and Assessment approaches, models and theories: Objective, subjective, projective, performance, unobtrusive, authentic, standardized, clinical and field-based tests and procedures. Tests, instruments and procedures as TECHNOLOGY and what that implies. Purposes of Measurement and testing. The importance of purporting and clear delimiting, and the key concepts of differentiated as opposed to universal reliability and validity. Readings: Kaplan Ch1 (or similar chapters from Renoylds or Popham), handouts.

**F7:** Classical Test and Psychometric Theory and Development Technologies. Theory, operational definitions, efficiency, costs, accuracy, generalizability, norm-referenced versus criterion-referenced, extraneous factors, reliability, validity, and interpretability. Domain and Standard-Referenced tests and protocols and new definitions of validity. Readings: Kaplan, ch3-6, (or similar chapters from Renoylds or Popham), handouts, Optional: Kerlinger ch25-27. Library/web exploration assignment due

**J14:** Assessment Domains and their general characteristics and typical measurement techniques and approaches. Measurement of groups and versus individuals; measuring programs and environments versus/and groups and individuals. Loci of accountabilities. Readings:, Kaplan, ch3-5, (or similar chapters from Renoylds or Popham), handouts, Optional: Initial chapters in Cronbach, Anastasi, or/and Nunnally or Kerlinger, ch29, scan any one (or more) of the dissertations by Nasser, Karna, McBride, Allen,, Kelley, Everitt, Eymero, or Perla.

**F21:** Taxonomies, Blueprints, Formats, Purporting, Test and “Item” Development Processes and Models. Readings: Kaplan, ch 6-7, (or similar chapters from Renoylds or Popham), handouts. Optional, Anderson and Krathwohl, Haladanya, or the dissertations of Nasser, Allen, McBride, Eymero, Everitt, Kelley, Perla, or Karna. Written evaluations of two professional prepared instruments due.

**F28:** Simple Item-Analysis Procedures and Criteria and their difficulties in alternative measurements contexts and models. Readings: handouts and item-Analysis sections/chapters in Kaplan, Renoylds, or Popham or Ferguson, chapter 23,24. Optional: appropriate chapters in Nunnally or Cronbach.

**FIRST SHORT PAPER/PROBLEM/PROJECT DUE.**

M20: Intersession

M27: Measuring Achievement: High-stakes, Low-Stakes; lower order and higher order cognition; performance and authentic assessment; content versus curricula validity; local, State, national and international achievement assessment programs. Readings: Kaplan Ch 10-11, handouts. Optional: remainder of Haladyna; **PROTOTYPE INSTRUMENT CONSTRUCTION DUE**.

A3: Measuring Achievement and capabilities continued. Individual differences, different populations, diversity, differing goals, and differing dreams. Readings: Kaplan 12-13 (or appropriate chapters in Reynolds or Popham), handouts. Optional: appropriate chapters in Nunnally or Osterlind. **FEEDBACK TO PEER ON PROTOTYPE INSTRUMENT DUE**

A10: Measuring traits, abilities, talents and interests; differing models and theories and specialized assessments and assessment procedures. Readings: Kaplan Ch 14-16, handouts; Optional: appropriate chapters in Cronbach and/or Nunnally; Rhodes, Eymero, or Allen Dissertations.

A17: Measuring Attitudes, Opinions, Beliefs, Emotions, Feelings, Practices and similar variables: Surveys, interviews, protocols, inventories, artifacts and unobtrusive techniques. Readings: Kaplan, Ch17-18, handouts. Optional: Nunnally, Kelinger and Chronbach chapters; Allen, Rhodes and Pearson Dissertations. **FEEDBACK TO PEER DUE ON PAPER TOPIC OR INSTRUMENT CONSTRUCTION PROJECT.**

A24: Measuring complex behaviors, processes and productions: Rubrics, protocols, analytical rating scales, panels, and content and interaction analysis and observational systems; dynamic assessment techniques.Readings: Arten, Handouts; Optional: Skelley, Kelley, Jackson, Kanalis, McBride, Karna, Everitt, or Perla Dissertations.

M1: Measurement and Assessment of Problem Solving Skills and similar higher order and complex behaviors.: Readings, Kaplan 10-11; handouts; Optional: McBride, Nasser, and Allen Dissertations, particularly their chapters on what is a problem and what is problem solving. **SECOND SHORT PAPER/PROBLEM/PROJECT DUE.**

M8: Measuring Environments and Process variables; Programs and Policies. Handouts, optional Weiss, dissertations by Allen, Kelley, Jackson, or Everitt, **List of all discovered valuable websites and journals due.**


M22: **Class Presentations.**