

Improving Information-Seeking Behaviors and Knowledge Bases for  
Thesis Students in the Applied Behavior Analysis Graduate Program and  
Students in the Rehabilitation and Mental Health Counseling Graduate Program

#### **Video 4: Finding Assessments, Tests, and Measures**

During the life of any research project, the day will come when you may need to find a test or measure. There are literally tens of thousands of published and unpublished tests and measures. Published tests are commercially available for purchase; unpublished tests are not commercially available for purchase. This is a brief guide to understanding the world of tests and measures, also known as psychometrics.

Psychometrics is the field of study concerned with the theory and technique of measurement of traits, abilities, attitudes, and knowledge. It deals with the construction of instruments and procedures for measurement; and the development and refinement of theoretical approaches to measurement. Psychometrics is applied widely in educational and psychological assessment.

Two key concepts are reliability and validity. A reliable measure is measuring something consistently, while a valid measure is measuring what it is supposed to measure. There are three major categories of reliability for most instruments: test-retest, equivalent form, and internal consistency. There are also three general categories of instrument validity: face validity, predictive and concurrent validity, and discriminant and convergent validity. The *Standards for Educational and Psychological Testing* (1999), established by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, provide a comprehensive basis for evaluating tests. These terms and standards are part of the evaluative component, which we will talk about shortly.

Your first question may be, How do I find a test for my research project?

The first place to start any search for tests and measures is at your local library. Not only does the library own numerous print and online resources that can help you to find tests, there are also librarians who can guide you in your search through these resources. These reference sources provide comprehensive and useful information on tests.

There are two major types of sources you will use to find and to evaluate the appropriateness of a test or measure. The first is a test review, which describes and evaluates a specific measure. These are outside reviews written by an expert in the field. Databases and books of test reviews, such as *Tests in Print (TIP)*, *the Mental Measurement Yearbook (MMY)*, and *Test Critiques*, are good places to start. Depending upon the specific resource, there are subject indexes, title indexes, and author indexes available to help you narrow down your search.

The second resource is journal articles written by researchers who have conducted validation studies on a specific test or who have used the specific test in their own study. These are found in journal databases, such as PubMed, PsycINFO, CINAHL, and ERIC.

You will need to spend some time determining what kind of test you need. Is it a personality test? Behavior assessment? Educational assessment? Who are you assessing? Children, adolescents, adults? A cultural or linguistic group? Is it gender-specific? The type of test for the specific population helps determine where you will start.

#### Who publishes the test?

In addition to listings in *Tests in Print (TIP)*, *the Mental Measurements Yearbook (MMY)*, *Tests (pro-Ed)*, *Test Critiques*, and other resources, the Test Collection at ETS (Educational Testing Service) and Pro-Ed test [publishers are two of the many test publishers. ETS offers a free brochure of major U. S. test publishers, which is yours for the time of a phone call \[1-609-734-5667.\] All major publishers have product catalogues \(online or print\) that you may request for free or for a small fee. As an individual purchaser of a test, you may be required to fill out a purchaser qualification form to ensure you are qualified and competent to administer and interpret the test.](#)

#### What is a test review?

A test review should cover the seven areas we just discussed. Minimum information includes the authors of the test, title and acronym, purpose, and publication date. In addition, the review should describe the population, either by years of age or other distinguishing characteristics, such as gender. It will also provide information on scoring, how the test is administered, the length of time to administer, and often publisher and price information. In the commentary, the reviewer offers an interpretation of the test, which addresses at least three issues regarding the instrument. The reviewer provides a description, how it was developed, technical information on how it was validated, and the reviewer's summary of the test. At a minimum, the review will include a reference to the original article describing the test, and often refer to other validation studies on the test.

#### Where can I find test reviews?

One of the best known is Buros' **Mental Measurements Yearbook (MMY)**. The Buros Institute of Mental Measurements at the University of Nebraska-Lincoln has been supporting advancements in test practices since 1938. [<http://www.unl.edu/buros/>]

The *MMY* is available online and in print at the USF Libraries. MMY divides the world of tests and measures into 18 test categories. In the MMY, tests are alphabetical by title. There are seven indexes that arrange tests by title, print status (in or out), acronym, classified subject, publisher, name (author, reviewer), and score. These are useful to locate tests about which you have limited information or to access groups of tests that fit into the same category, such as which tests yield the same score or which tests all measure depression. If you use the MMY online, it only covers volumes 9 to the current volume, which means you are missing out on 8 possible places where your test review may reside.

**Health and psychosocial instruments (HAPI)** provides access to approximately 15,000 tests and measures in the fields of health and psychosocial sciences. The database started indexing tests in 1985. One very valuable feature in HAPI is that you can limit your search to primary and secondary sources. A primary source is the original article where the test originally was published or a resource that contains the full-text of the test. Secondary sources do not include the entire test. The database can include articles that describe how the test was used, outcomes, and possibly sample questions. The database also allows you to search for translations of a particular test.

The **Test Collection at ETS** (Educational Testing Service) [[www.ets.org/testcoll/](http://www.ets.org/testcoll/)] is a database of more than 25,000 tests and other measurement devices. There are basic and advanced search options, and you can select for downloadable tests (after the fee is paid). Each record includes author, title and acronym, first article published, major and minor descriptors (subject headings), abstract and the security level of the test (whether it can be purchased). The descriptors are based on the *Thesaurus of ERIC Descriptors*, a controlled vocabulary—a carefully selected list of education-related words and phrases. Since searching by keywords requires matching the exact words found in a record, searching by Descriptors allows you to locate records indexed by subject, regardless of the terminology the author may have used. The *Thesaurus* is available online through the USF Libraries databases (choose Education Resources Information Center ERIC) as well as a free internet resource [<http://purl.access.gpo.gov/GPO/LPS5430>].

The next sources are print resources: ***Tests and Test Critiques***, published by Pro-Ed, Inc. ***Tests*** is a bibliographic encyclopedia containing information on thousands of testing instruments in psychology, education, and business. It contains concise descriptions of tests, however, it does not contain evaluative critiques or data on reliability, validity, or norms. If you have the name of the test, you can find evaluative criteria on your test in *Test Critiques*. In each of the three sections (psychology, education, and business), the tests are arranged alphabetically. ***Tests*** also has seven indexes: the test title index, author index, foreign language availability index, and a

publisher/distributor index, and three indexes of tests suitability for special populations: the hearing impaired, visually impaired, and physically impaired. *Tests* also lists out-of-print tests and has a publishers directory.

*Test Critiques* is a companion to *Tests* and contains psychometric information (reliability, validity, and norm development). Each test entry includes an Introduction, Practical Applications/Uses, and Technical Aspects, followed by a critique. The Introduction describes the test in detail, including authors, publisher, and the purpose of the test. Practical Applications/Uses describes the intended population, administration, scoring, and interpretation procedures. Technical Aspects includes citations from reliability and validity studies and opinions from experts regarding the technical adequacy of the test. A bonus feature of *Test Critiques* is that it explains methods and technical terms used in psychometrics. Each volume contains a table of contents with tests listed alphabetically. Each volume also contains cumulative indexes organized by test title, test publisher, test author and reviewer, and subject.

There are numerous stand-alone books that also carry test and measure information. Find them in the USF Libraries catalogue.

Secondary sources, such as journals databases are also a good place to search. We recommend **PsycINFO**, **PubMed**, **CINAHL**, and **ERIC**. Here are quick tips to ensure you pull up studies addressing the psychometric quality of a test.

Limit searches in **PsycINFO** using the classification codes. PsycINFO has a predetermined LIMIT option entitled 'TESTS & MEASURES' (2200 in the classification codes).

In **PubMed**, use your subject term and combine it with Questionnaires AND (Psychometrics OR Reproducibility of Results OR Validation Studies).

Search **CINAHL** (Cumulative Index to Nursing and Allied Health) two ways. Select the field "Instrumentation" from the drop-down box and search either by topic or name of test. Also, search by topic and limit by Publication Type to "Research Instrument" or "Questionnaire/Scale".

To search in **ERIC**, select Tests/Evaluation Instruments from the "Publication Type" limit box and add a keyword or subject search.

Now you have test reviews and articles on how the test was used in studies similar to your own proposed research project. Next is the hard part -- evaluating the test.

[How do I assess journal articles about my test or measure?](#)

In addition to reading test reviews, you will also be reading articles written to assess the validity and reliability of tests as well as articles of other studies that used a test in which you are interested. The *Standards* mentioned earlier provide a comprehensive basis for evaluating tests. Here are the seven areas that should be part of any evaluation of a test or measure.

1. Test coverage and usage
  - a. Intended uses of the test scores; appropriate score interpretations; limitations or restrictions of the interpretations.
  - b. Test population
  - c. Content/coverage decisions; development of the items and selection criteria
2. Appropriate sample for test validity evidence
  - a. Sampling methods
  - b. Size of sample and subgroups
  - c. Basis for validity and for norming; sufficient variation to distinguish participants (e.g., those with and without symptoms)
  - d. Representativeness of the sample (e.g., age, geographical distribution, cultural background, and disabilities)
3. Validity (Construct validity is the only benchmark in this category)
  - a. Content related evidence (exploratory or confirmatory factor analysis, evaluation of test content and test specifications, test item creation)
  - b. Relationship to other tests (criterion measures; score distributions, predictive accuracy, and differential prediction)
  - c. Other forms of validity evidence (experimental studies using the measure, use of a multi-trait multi-method matrix; construct similarity between different tests)
4. Reliability
  - a. Computation of reliability estimates, appropriateness of statistical methods
  - b. Test-retest time intervals

- c. Reliability estimates provided for all scores where interpretations are indicated.
- 5. Test administration
  - a. Is the administration information clearly understood and qualifications spelled out?
  - b. replication of test administration same conditions under which the test was validated and normed
- 6. Test reporting
  - a. How are test results reported, scales for reporting results are appropriate
  - b. Other resources and materials for use in interpretation
- 7. Test and item bias
  - a. Minimal cultural dependency, explicit statement of inappropriateness of tests for specific groups
  - b. Differential validity across groups and analytic review
  - c. Language proficiency of native and non-native speakers of the majority language.

Now that that's all clear, you are ready to start searching for a test or measure for your research project, find some possible measures, evaluate them, and choose the perfect test. We hope this has been a good guide to navigating the new world of tests and measures.

For more tips on how to search in specific databases, email us at [fmhlibrary@usf.edu](mailto:fmhlibrary@usf.edu) to set up an appointment or to talk through a search.

ABOUT THIS VIDEO SERIES: This video is one of 15 created with a grant from the University of South Florida CTE21. With this grant, we were able to create targeted content for graduate students in the Applied Behavior Analysis program and the Department of Rehabilitation and Mental Health Counseling in the College of Behavioral and Community Sciences, University of South Florida.