|  |
| --- |
| **What are Assessment Instruments**? |
| When educators want to know a student's current level of understanding in some domain of knowledge, they typically administer some kind of achievement test. Similarly, when educators want to understand a student's current level on some developmental continuum, they administer a psychological test. Both these types of tests will be referred to as assessment tools or instruments. They are used to quantify a student's level of knowledge or development. Change in the scores on these instruments is considered an indication of learning or growth in the student. |
| **How Do I Select an Instrument**? |
| Deciding on how to measure the learning/development that occurs in a program requires considering several factors, weighing the pros and cons of each available option, and deciding on an option that is most appropriate for a given situation. There is unlikely to be an ideal option, so be prepared to make compromises.  To ensure that you've made the best choice possible, make sure you have identified all possible options [(see section on finding existing instruments)](http://www.jmu.edu/assessment/pass/assmntresources/instruments.htm#ExistingInstruments), and have collected the necessary information about each instrument [(see section on judging instrument quality)](http://www.jmu.edu/assessment/pass/assmntresources/instruments.htm#Quality).  Factors to consider:  **1. Alignment to program objectives**  The most important factor to consider when selecting a measurement tool is how well the instrument aligns to the programs' learning objective. Alignment is crucial because it relates directly to the utility of the information gained from assessment. The more aligned the instrument is to the program's objectives, the more direct the inferences to the effectiveness of the program. When alignment is poor, the results of the assessment provide weak or limited information about the program.  **2. Cost/Resources**  Some instruments need to be purchased from a test publisher or author. Other instruments are available free of charge. Also, scoring or analyzing data from certain instruments will require time and energy, or employing others for assistance. For each potential instrument, be sure to consider what resources would be necessary to fully implement using the instrument for a program's assessment. Even if an instrument is available free of charge, it is still a good idea to contact the test author to request the use of their scale for assessment purposes.  **3. Selected vs. constructed response instruments**  There are two general types of instruments that can be used to measure learning and development: selected-response and constructed-response. Selected-response instruments include those in which the participant chooses a response from a list of options provided. Some examples of selected-response instruments are the common multiple-choice items on standardized test and the 'strongly agree' to 'strongly disagree' response scale options common on surveys and opinion measures. Constructed-response instruments include any type of assessment in which the answer is not provided and must be performed or generated. Examples of constructed-response instruments include essay tests, oral presentations, or reports/papers products.  The advantage of selected-response instruments is that they can be administered quickly and easily to large numbers of students and cover a broad range of topics. The advantage of constructed-response instruments is that they can provide a deeper, richer picture of what a student knows and can do. They also can be more diagnostic and formative if designed well. The disadvantage of constructed-response instruments is that they are costly to score because they must be rated by trained raters. Also, they are time-consuming for students so they typically only cover a few topics. Thus, a constructed response test generally represents less content/depth than a selected response  instrument. On the other hand, selected-response instruments are often criticized because they provide little to no  diagnostic utility and tend to focus more on recall and other low-level cognitive skills.  **4. Using an existing instrument or creating a new instrument**  Many measurement tools have already been created. The benefit of using an existing instrument is convenience. If you use an existing instrument you can save a lot of time in instrument development. On the other hand, it is hard to find an existing instrument that aligns exactly with a specific program's objectives. Thus, the advantage of creating a new instrument is that you can customize the instrument to achieve very strong alignment. Deciding about whether to use an existing instrument or create a new one is a matter of balancing convenience with alignment. Also, keep in mind that instrument development is a resource-intensive endeavor, but in the end you will own the instrument. Alternatively, if you use an existing instrument, you will likely have to pay the creator of the instrument, but will save yourself the development time.  **5. Instrument Quality**  Like most things, instruments can be of either high or low quality. The better the quality of the instrument, the more trustworthy the scores that are obtained from the instrument. It is imperative that high quality instruments are used for assessment work. The higher the quality the instrument, the more useful the results from the assessment will be. If you don't trust the scores from an instrument, you won't be willing to make decisions about a program based on the scores of the instrument. Before deciding on an instrument, the quality of the instrument should be evaluated (see section on [judging instrument quality](http://www.jmu.edu/assessment/pass/assmntresources/instruments.htm#Quality)). |
| **How Do I Judge the Quality of an Instrument?** |
| Evidence about the quality of an instrument is typically called validity evidence, and there are usually several different types of this evidence.  **Content validity**   * To know what a test measures, don't just rely on the title of the instrument! Look at the items. Find out what theory was used to guide item writing. * If an instrument is good, items were developed for each different aspect of the topic being covered. * If an instrument is good, experts in the field, who are unfamiliar with the items, agree on what the items are measuring.   **Internal-consistency validity**   * In a high-quality instrument, we should see patterns of item-responses we expect based on what we think the items are measuring. * If items are measuring the same thing, students should respond to the items similarly. * If items are measuring different things, students' responses to the items should differ at least slightly. * Often internal-consistency validity is called *reliability*. Reliability refers to how consistently an instrument measures something. * Reliability indices range from 0 to 1.0 with higher values indicating that the measure is more reliable. Ideally, for program assessment we look for reliability values of .70 or higher.   **External validity**   * External validity is comprised of convergent and discriminate validity. * Convergent validity: Scores on the instrument should relate highly to scores on other instruments that measure similar or highly related things. * Discriminate validity: Scores on the instrument should relate weakly to scores on other instruments that measure different or opposite things. |
| **How Do I Create Instruments?** |
| Instrument development can be time-consuming and complex. It is best to work alongside an assessment or measurement professional. Below is a general outline of the process to give you an idea what is involved.  http://www.jmu.edu/assessment/pass/images/instrumentdevelopment.gif |
| **How Do I Find Existing Instruments?** |
| In order to make the best decision possible about how to measure learning and development in a program, you'll want to find all possible options and compare them. There are lots of different places you can look to find instruments that may be useful for a program's assessment:   * **Mental Measurement Yearbook** - reviews commercially available tests * **Health and Psychosocial Instruments** - indexes non-commercially available tests * **Online reference databases** (PsycINFO and ERIC) - key words: tests, measures * **Tests in Print** - reference book available in most libraries * **Pro-Ed Test locator website** - <http://www.proedinc.com/customer/productLists.aspx?brandid=2> * **ETS Test Directory** - Test link: http://www.ets.org/ * **American Psychological Association**: http://www.apa.org/science/faq-findtests.html * Books about measurement in a specific area - sometimes there are books about how to measure the topic of interest (e.g. measuring diversity attitudes). * Conference presentations/papers that involved measuring a similar construct - (this will inform you of how researchers in the field are currently measuring the topic). |