

## **Steps in Instrument Development for this Course** **Mickie Swisher, January 2020**

This document provides some general procedures that we will use in this course to improve reliability, validity and discriminatory power. Use this document in conjunction with required readings for the class, the various documents of mine linked at the course website, and the additional resources provided each week.

**Step 1: Define all constructs, including those that are not the focus of your research, fully before you attempt to create an instrument.** Define *all of the constructs, not just those you will use and identify all of the dimensions for each construct*. You need to define all of them to make sure that there is no overlap or ambiguity in the definitions. Go to more than one source of information about the theory because there may be some differences in definitions even within the same theoretical perspective. Write down the definition you use. Include a one or two sentence definition of the construct *and* a description of the major differences in how the construct is defined and the decisions you had to make. For example, you might find that one theorist argues that there are three dimensions and another four. You have to make a decision about which to use. This is critical when you have to write your dissertation or thesis and when you publish. You need a written record of the decisions you made that could affect the outcomes of the study.

**Step 2: Identify variables that will represent the constructs.** You may find useful information in the literature that will help with this, but remember that the variables are specific to your study topic and the context (place, time, population) in which you will conduct the study. Make sure the variables represent all dimensions in the constructs of interest for your research. It is better to have too many variables than too few early in the process. You may want to generate measures for two or three variables representing the same construct because you can often merge the scores later if you decide to reduce the number of variables. You can also use this redundancy in the assessment of the validity of your measurements because multiple variables representing the same construct should have similar patterns of response. If you have multiple variables per construct, you can assess this *convergence* in scores. Multiple variables per construct are *recommended*.

**Step 3: Develop the specific items you will use.** It is always valuable to look for existing instruments and use them if they are appropriate to the context for your study and using an existing instrument – if it will work for you – is always preferable to creating your own. You must purchase or get the consent of the creator of an existing instrument to use it in your research. Failure to do so constitutes plagiarism. You can also sometimes “borrow” individual items from instruments others have developed. However, you need to cite the original work when you do this. If you are going to use more than one or two items, you need to get permission from the original creator to use the items. Even when whole instruments or items are available that seem like they will work for you, you need to test them thoroughly because context is critical to wording, content, and even response formats for items. In this class, I want you practice developing your own instruments and your own items. Therefore, I expect you to rely *primarily* on items that you development. Do not use existing instruments. **No more than 25% of the items in any of the instruments you develop should be from the literature – 75% are your creations.**

**Start with many redundant items.** You will run multiple tests on your item banks to determine which items yield reliable, valid, and discriminatory results. Redundant means that you start with many items for each variable, items that differ significantly so that they capture the full meaning of the construct. You will eliminate many (maybe most) of them. You may also have two, three or more versions of an item that differ in wording or construction (for example, a reverse-scored version of an item) where such differences could influence how the respondent understands

what you are asking. However, focus on redundancy in **content, not wording**. Taking an example of an instrument using closed responses (check the box), if I start with 30 items, I find that only 8 or 10 or them remain after all this testing. My rule of thumb is to start with 30-40 items **per variable** for instruments like an index, scale or test. For other instruments like an interview protocol with an open, narrative response, the key is to start with many different *kinds* of questions for each topic. For example, a leading question might start the conversation about a topic, followed by some probing questions (confirming and disconfirming) and end with some summary questions. We will look at the range of types of questions you can use in our exercise to develop an interview protocol. You would have several of the trails of questions for each variable just as you would have several items for a “check the box” set of responses. The principle is the same. **Start with many items in all assignments.**

**Step 4: Revise to correct problems in wording, response structure, etc. (technical issues).** Do not waste other people’s time looking at your instrument until you have done your best to correct technical problems. Do this before anyone else sees your instrument.

**Step 5: Expert Panel Review.** After you have completed your review and correction, you need to get an expert panel to review your work. An expert panel consists of people who have expertise either in methods of data collection *or* in the topic of your research and/or your theoretical approach. This is **not a check of wording, etc.** Experts may find some of those problems, but it is your responsibility to fix those problems. The expert panel normally consists of your colleagues and for students the chair and members of your supervisory committee. Expert panel review is what your committee does when you defend your research proposal. In this class, you will largely use each other as members of the expert review panel. The work of the expert panel is largely conceptual. Typically, you would ask them to do three things for you. (1) The panel should provide an assessment of whether your instrument(s) capture the full meaning of the construct. (2) They should identify any aspects of your instrument that are **not** relevant to the theoretical construct **or** that are not appropriate for the topic/context of your study. (3) The panel should assess the adequacy of the format and structure of the items.

**WARNING**

**You CANNOT USE the Likert-type response format for this class unless there are sound reasons why some other approach will not work.**

**The common use of the Likert-type statement (sentence or phrase) with several categories of agreement or disagreement (strongly agree to strongly disagree in most cases) is often an example of an inappropriate item-response format.** I once saw a Likert item-response format used with statements like “My department provides training for teaching assistants.” This is a simple “yes/no” item – if is a correct statement of fact or it is not. In any case, the Likert format of agree to disagree creates a great deal of mental work for the respondent who must complete at least four mental steps (1) read a statement, (2) decipher/interpret the meaning of the statement, (3) determine one’s basic agreement (yes/no), and (4) decide which specific category of agreement or disagreement best reflects one’s assessment. This mental work is valuable when topics are emotionally charged because it gives the respondent a way to create emotional distance from potentially disturbing ideas expressed in the statements. If the topic is not one that is apt to create emotional turmoil for the respondent, **do not use the Likert-type statement and agree/disagree response format.**

## WARNING

**You will be penalized (LOSE POINTS) for using statements rather than questions when there is no research imperative to use any of the response formats that rely on statements.**

**Questions are generally preferable to statements because they elicit a more straightforward thinking process on the part of the respondent.** They are easier for the respondent to process. Here is an example. Question: How often do you host dinner parties in your home? Select one. Responses: 1-3 times per year, 4-6 times per year, 7-9 times per year, 10-12 times per year, more than once a month. Statements; I rarely host dinner parties in my home; I occasionally host dinner parties in my home; etc. Just ask the question.

**Provide the panel members (again, that may be just one or two people) with the research questions, a brief description of your theoretical approach, and a full definition of each construct you are trying to operationalize.** They do NOT need to see the standard items for demographics and such. It is also useful to draw their attention to areas where you are uncertain and feel you need their input. Do everything you can to make this as easy as possible for them to help you. Unless otherwise instructed, always conduct a cognitive review with members of this class, faculty members, or other graduate students with the requisite expertise for all assignments in this class.

**Step 6: Cognitive Testing.** I regard cognitive testing as nearly equal in importance to expert panel review in the process of instrument development. Like expert panel review, you can reap major improvements from the process at relatively little expenditure of time and effort. You can and should use this technique with all types of data collection. Also like expert panel review, you **do not ask people to answer the questions in most instances – you ask them to tell you how they go about deriving an answer to the question.** E.g., you are asking people to explain the cognitive processes they *would use to answer the questions*. **Cognitive testing must be done with members of the target population for the study or with individuals who are “very much like” the members of the target population with regard to characteristics or traits that can affect how well they can respond to your questions.** I have provided several resources about the use of cognitive testing (Collins, 2003; Castillo-Diaz & Padilla, 2013; Priede & Farrall, 2011). I expect you to use cognitive testing in virtually every assignment in this class.

**Step 7: Pilot Testing.** **Pilot testing** is the step when you ask members of the target population to respond to your items. These are your first data points. I recommend that you treat the test as the first phase of data collection. If you do that, and your instrument performs well, you may be able to use the data as part of your final data base. The document “Procedures for Operationalization” discusses various techniques for pilot testing instruments. It covers techniques that I felt you need to know about **and use in this class**. We will discuss additional techniques as we cover specific methods of data collection during the semester.