**Item Writing**

In class, we discussed several principles for good item writing. Each member of the class submitted 5 items measuring “Satisfaction with Graduate School.” In small groups (3-4), discuss and critique as many of the items as you can in the time allowed for lab.

Remember that we said good items are:

* #1. Short.
* #2. Readable.
* #3. Grammatically correct.
  + Avoid double negatives.
    - Or perhaps avoid negative words altogether.
  + Use adverbs & adjectives correctly.
  + Read for unintended alternative interpretations.
  + Complete sentences are often best.
* #4. Consistent.
* #5. Straightforward.
  + Not double-barreled.
  + Related unambiguously to the construct.

Also keep in mind Burisch’s principles of communicability and simplicity. Even a well-written set of items will not yield a good *scale* if the items don’t clearly represent the construct you are trying to measure (in the judgment of an expert, not necessarily a test-taker), or if they don’t cover all of the construct domain, or if they don’t address the full *range* or spectrum of the construct you are trying to measure. Are some aspects of the construct or some levels (e.g., very dissatisfied) better represented than others in these items?

**Lab Exercise:**

Please answer the following questions and submit your answers as a Word document in Canvas. You can and should work in small groups to critique the items, but each person should turn in their own (distinct) lab assignment and answers to the questions.

1. List five items that the entire group agreed were good items, and explain why each is a good item.
2. List five items that the group thought were poor, and show how they could be revised to be good items. Give the original item, the revised item, and an explanation of why the changes address the problem.
3. Of the full set of items, how many were good, how many needed revision, and how many would you recommend dropping entirely?
4. Considering the set of items as a whole, do you think they measured some portion(s) of the construct spectrum better than others, or were they evenly distributed? For example, would these items tell us most about whether people were pretty satisfied vs. very satisfied, and not much about people who had really negative experiences? Or would they measure across all levels of the construct?
5. Are there any aspects of the construct that were not well addressed by this set of items? That is, if we really wanted to measure satisfaction with graduate school, are there areas about which we would need more items?
6. Based on your experience with this exercise, which characteristic of good item writing seems to be the hardest to achieve? Why do you think this is?