**04b: Questionnaire Selection**

**1. Define Purpose of Study and Measurement**

How does one decide which questionnaire or scale, among many possible, is best for a given situation? Below are a few steps to help in the process of questionnaire selection.

* What must be measured for study to be successful? Identify each variable that must be in the study.
* How are variables defined? (e.g., test anxiety is…)
* How are variables operationalized (e.g., SES = income only or SES = income, education, occupational prestige)?
* What is target population of study?
* How will variables be scored?
* Will variable scores be suitable for data analysis plans?

|  |
| --- |
| **Example Study**  **1. Define Purpose of Study and Measurement**  **Purpose**  Study will be designed to learn whether high school students’ academic performance can be predicted by their self-efficacy for learning, self-regulated learning behaviors, and autonomy support in the classroom.  **Variables Loosely Defined**  Self-efficacy for Learning  Confidence and belief that one can learn and perform well on the task at hand.  Self-regulated Learning  Degree to which students think about how they think and learn, and how they monitor and act on their behaviors to help them learn.  Autonomy Support  Enabling students to have a role in classroom planning and procedures, choice of activities, and recognizing and supporting their needs and desires for learning.  Academic Performance  Composite measure of student achievement across several tests and performance assessments.  Figure 1: Path Diagram of Planned Study |
| **Operationalized**  Goal is to find scales for each construct that include as few items as possible to measure each construct with valid and reliable scores. Relevant dimensions of each construct should be sampled by items (indicators) used to measure each construct.  **Scoring of Variables**  Scales should provide numeric scores that are ranked or sorted showing differences by degree or level for each construct.  **Data Analysis Plans**  Use either structural equation modeling with latent variables or regression with composite variables. |

**2. Locate Existing Instruments or Scales**

* Do instruments exist to measure relevant variables?
  + More than one instrument can be used
  + Can select parts or components from instruments and combined to form new instrument
  + Look for multiple scales and constructs; sometimes other relevant constructs will appear during search
* Google Scholar (scholar.google.com) useful search tool; GSU library online search is excellent
* Search terms – consider including some of the following terms
  + (Alpha OR Cronbach OR Raykov)
    - These will limit search to studies that include scales and report Cronbach’s alpha (a measure of internal consistency reliability)
  + Target population, e.g., “college” or “work” or “teachers” etc.
  + (Scale OR instrument)
* Locate sources that provide scale item wording
  + Many published studies will not present complete scales, instead they provide a few example items
  + Sometimes wording is provided in
    - Text of the Instrumentation or Measures section of Method
    - Tables in Method or Results section
    - Appendix
  + Do not rely on scale title to judge relevance
* Scale wording – adequate?
  + Read scale items carefully to ensure those items are relevant to your construct definition or operational definition
  + If wording match is inadequate, look for another scale or consider how to revise wording
  + If scale wording is offensive or invasive of privacy unnecessarily, consider another scale – lower response rates will be expected for scales that respondents find distasteful
* Complexity, Length, Administration, and Scoring
  + Does the scale have conditional questions? For example: If you answer “Yes” to Question 2, go to Question 5
  + Such complexity can create confusion and reduce likelihood that respondents will complete the instrument or complete it successfully
  + Find scales that are short as possible; shorter instruments tend to produce higher response rates
  + Does the scale require specially trained individuals to administer it to participants?
  + Does the scale require specially trained individuals to score it?
* Availability
  + Is it available for use freely or copyrighted?
  + Must you gain permission for use from authors?
  + If use requires a fee, find another scale.

|  |
| --- |
| **Example Study**  **2. Locate Existing Instruments or Scales**  **Google Scholar**  Used this search phrasing in Google Scholar  “autonomy support" classroom achievement (Alpha OR Cronbach OR Raykov) (Scale OR instrument)  **Search Results**  Sometimes one will have to look through 5 or more pages of results, or edit the search string, to find usable scale wording. It can be a slow process. Below are the first three links obtained.    **Find Scale Wording**  The 1st listing seems to have a scale that might work (Greene, Miller, Crowson, & Duke, 2004). Wording for the autonomy support scale is provided in the appendix (p. 478).    The response scale for these items is described (p. 469). |

**3. Reliability and Validity Evidence**

* If scale wording is appropriate, next move to assessment of reliability and validity evidence
* First, how well did the scale perform in prior research?
  + Did the scale produce scores that behaved as predicted (e.g., correlate with other variables as expected, or show group differences as expected)?
  + Did scores from the scale demonstrate adequate variability – were the scores grouped at the high or low end of the scale (i.e., floor or ceiling effects)?
  + Was there excessive missing data suggesting respondents were uninterested in completing the scale?
* Reliability - most common types listed below
  + Cronbach’s alpha (α)
    - most commonly used index of reliability
    - α should be .70 or higher; .60 to .70 acceptable if no other scales available
  + Raykov rho – similar to alpha; sometimes reported for latent variables
  + Both alpha and rho measure internal consistency
  + Test-retest
  + Intraclass correlation, ICC
* Validity – seek evidence of
  + Content/Face Validity (i.e., logical validity) – did author discuss
    - How items were developed
    - Theoretical fit of items to construct
    - Dimensions of construct
    - Expert review of items
    - Field test or pilot study of items
  + Structural or Factor Validity – did authors present or discuss
    - Factor analysis results of scale
    - Show factor scores and indicate a good fit was obtained
    - Discuss measurement model fit of scale and indicate good fit found
    - Discuss or show any empirical test for dimensionality of scale
  + Construct or Criterion Validity – did authors present or discuss
    - How scores from scale behaved in a predictable manner
    - How scale scores correlated as expected with other variables
    - Whether mean scale scores differed, as expected, across groups

|  |
| --- |
| **Example Study**  **3. Reliability and Validity Evidence**  Greene et al. (2004) provided several bits of evidence to support reliability and validity. First, for validity they explain (p. 469) that the autonomy support scale is part of instrument that was validated by Blackburn (1998).    Next, for structural validity (p. 470) they explain that they performed a confirmatory factor analysis of items from this instrument to test whether scores from each scale would behave as expected and form unique scale clusters as expected. The numbers they present suggest the model fit is satisfactory.    They present reliability evidence, Cronbach’s alpha (p. 471) in Table 2. The alpha of .65 is a bit below the value of .70; still usable, but lower than hoped. Might be worth seeking other scales of autonomy support and use scale is something better is not found. |

**4. Adapting Items**

* Common for researchers to edit items so they better fit their study
* Indicate in Instrumentation section of Method what changes were made to items and why
* Changes to items can cause validity and reliability to change too
* Important, if possible, to provide evidence for reliability and validity of scores obtained in your study, even more so if you make changes to items; remember, reliability and validity evidence is sample specific so you should plan to always provide this evidence for each study

|  |
| --- |
| **Example Study**  **4. Adapting Items**  Meluso et al. (2012) reported that they adapted a measure of self-efficacy to fit their study (p. 500). Unfortunately, they did not explain what adaptations were made to the items.    Below are the items they used (p. 502).    These focus on science, but could easily be adapted to other content areas such as mathematics, reading, etc. For example:  1. I am sure that I can learn ~~science~~ statistics.  2. I can get a good grade in ~~science~~ statistics.  etc. |

**5. Reading Factor Scores (to be added)**

**References**

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students’ cognitive

engagement and achievement: Contributions of classroom perceptions and motivation. Contemporary Educational

Psychology, 29, 462 – 482.

Kimberlin CL, & Winterstein AG. (2008) Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm. 65, 2276–2284.

Meluso, A., Zheng, M., Spires, H,A., Lester, J.(2012). Enhancing 5th graders’ science content knowledge

and self-efficacy through game-based learning. Computers & Education, 59, 497-504.