

Chapter 4 Constructing Questionnaires and Indexes

When one is interested in conducting primary research, he or she most often needs to either purchase or construct a measuring device called a measure or instrument. Instruments are often called a questionnaire, index, or scale; these terms are frequently used interchangeably.

A measuring instrument is composed of different types of item formats which are designed to produce data which are then transformed, via statistical and/or logical analysis, into useful information and then interpreted (with reference to the relevant scientific and professional literature) to answer a research question(s) or test a hypothesis. Instruments using the construction process and item formats discussed in this chapter are most often used in survey research.

The planning and instrument construction process is described in this chapter. The item formats, discussed will enable you to construct a survey instrument, questionnaire, scale, or index. The item formats, described in this chapter are “mixed and matched” to produce data which will ultimately answer a research question or test a hypothesis; so, it's not uncommon to have variety of item formats on the same data collection instrument. Presented in Appendices 4.1-4.5 are various measuring instruments used (or written primarily for illustrative purposes) by the authors to illustrate the item formats discussed within the chapter. The item formats presented in this chapter are routinely used to measure attitudes; for a description of an Attitude Taxonomy, see Appendix 4.8.

I. Constructing Instruments (e.g., Questionnaires, Indexes, or Scales)

A. Decisions to Make Before Constructing a Measure

1. Recall our earlier discussion that measurement instruments are typically administered to describe, scale, and/or classify and must be appropriately reliable and valid.
2. There are six (6) decisions that must be made when planning an instrument (see also Cooper & Schindler, 2013). These are:
 - a. The researcher must decide what and whom to study, i.e., respondent characteristics within some context, or respondent's opinions about what is presented to them, e.g., political issues or candidates, satisfaction (employee or customer), products or services.
 - b. The researcher must decide how respondents are to respond to the items presented to them. Typically, responses to items on measures are either:
 - (1) Rating items are used when respondents provide a score about, but don't directly compare an attribute, attitude, value, intention, object (e.g., product packaging), or behavior, etc.
 - (2) Ranking items require the respondent to compare two or more attributes, attitudes (The Attitude Taxonomy is presented in Appendix 4.8), values, intentions, objects (e.g., breakfast products), or behaviors, etc. Respondents may be asked to select which pair of glasses is more attractive or be asked to rank order the importance of color, style, fit, tint, and cost of glasses.
 - (3) The researcher must ensure that respondents sort an attribute, attitude, value, intention, object, or behavior, etc. into groups or categories, by the manner in which the individual items are constructed. Categories might include

demographic characteristics, income levels, preferences, disagreement or agreement, etc.

- c. The researcher must determine the number of dimensions to be measured.
- (1) A measurement scale or index may be unidimensional. A unidimensional measure (also called a scale or index) is designed to measure one attribute or characteristic. Consider personality which has many attributes. Now, suppose we wanted to measure only shyness then we would write items designed to measure only shyness and we would employ a unidimensional measure as we excluded the other dimensions. For example, see Appendix 3.1.
 - (2) An instrument (scale or index) designed to measure more than one attribute is said to be multidimensional. Such a complex construct as personality has many attributes or dimensions including shyness, intelligence, locus of control, self-concept, etc. Thus, if we were to measure personality, the measure would need at least four unidimensional scales or subtests. The more fully the construct, personality, is defined or described there is greater likelihood that even more subtests would be needed. Decisions about dimensionality are critical; we must have a clear, complete understanding of what it is we are attempting to measure.
 - (a) For an example, The Academic Credit Participation Index (ACPI in Appendix 4.7) is based on the Cross' Chain of Response Model (Cross, 1981, p. 124); see Figure 4.1. The COR model consists of 7 parts or dimensions. The model or theory attempts to explain how someone is motivated to return to school to earn an associates, bachelors or graduate degree. The process goes like this.
 - [1] A prospective student conducts a self-assessment of his or her academic skills (Part A) which is influenced by prior experience with formal, organized learning (Part B).
 - [2] Assuming the prospect self-evaluates highly, he or she decides whether or not his or her goals will be met by enrolling in a degree program (Point C) but this decision is influenced by life events (Part D).
 - [3] Assuming, the prospect has decided to enroll in formal academic credit learning, he or she determines whether or not enrollment is possible given presenting opportunities and barriers (Point E). Next, he or she learns about opportunities and how to overcome barriers by collecting information (Point F).
 - [4] Once a prospective student works his or her way through Points A-E, he or she participates in academic credit courses (i.e., a degree program).
 - (b) In Table 4.1 each dimension of the Cross Model, based on its operational definition, is aligned with a specific ACPI subtest. See, also, Appendix 4.7. The operational definition provides guidance for writing subtest items to measure the dimension. The operational definition needs to be as clear and as complete as possible so that the dimension is accurately and completely measured. Since the COR Model has 7 dimension, the ACPI has 7 subtests, one for each dimension.

Table 4.1

COR & ACPI Alignment

COR Dimension	ACPI Subtest
A	A
B	B
C	C
D	D
E	E
F	F
G	G

- d. The researcher must determine the types of data (i.e., level of data) to collect.
- (1) It is required that we know the type of data (nominal, ordinal, interval, or ratio) that each item on the measure will produce. This will enable us to select the most appropriate statistical indices and/or tests when data are tabulated, summarized, analyzed, interpreted, and reported.
 - (2) Once we know what level of data we need, given our reporting purposes, then we can construct items, subtests, etc. which will provide those data.
 - (3) The initial data analysis plan is written at this point; however, it will be revised based from what was learned during pilot testing.
- e. The researcher must determine what type of scale or index to construct.
- (1) While Cooper and Schindler (2013) outlined five types of scales or indexes, we will consider only two: arbitrary and factoring.
 - (2) All scales or measures are constructed for a purpose; so in a sense all scales or indexes are arbitrary. There are well developed instruments designed to measure constructs such as leadership style, problem solving skills, sales potential, teaching effectiveness, etc. Well-developed arbitrary scales and indexes have established reliability and validity; but, many “arbitrary” scales don’t and should be used with caution. Appendices 4.1, 4.2, 4.3, 4.4 and 4.5 are examples of the arbitrary approach.
 - (3) Some scales are constructed using a powerful statistical procedure called factor analysis. Recall our prior discussion concerning construct validity. Most scales based on enduring constructs (e.g., organizational culture or personality tests) are constructed using factor analysis. The Academic Credit Participation Index (Appendix 4.7) was constructed using the factoring approach.
 - (4) Many scales or measures are constructed to assess a particular organization, customer pool, or some specific attribute and are only used once or twice. These are often referred to as Questionnaires. It is unlikely that these measures were constructed using factor analysis. In order to have confidence in making decisions based on study results, all instruments must be at least content valid and internally consistent.

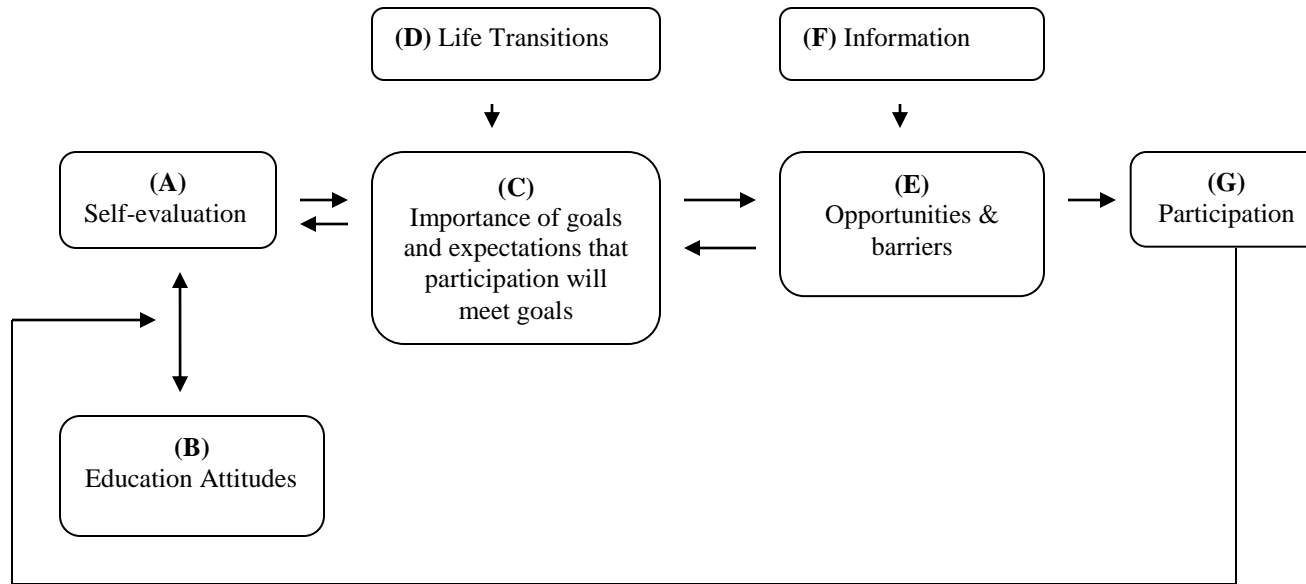


Figure 4.1 The COR Model (Cross, 1981, p. 124).

- f. Establish a quality control panel (QCP).
 - (1) The QCP is composed of subject matter experts and research or evaluation methodologists who advise the study principal or co-investigators (the folks who actually plan and conduct the study) on study design and execution.
 - (2) A QCP subcommittee might be composed of advisors who are similar to intended subjects or respondents, to ensure study instrumentation, design, and management doesn't introduce biases which may contaminate the study.
 - (3) The QCP can also be used to establish the measuring tool's content and/or construct validity. If an instrument is truly valid, it's also likely to be reliable.

B. Process for Constructing an Instrument

1. First, the research or evaluation study must be planned.
 - a. Every study is conducted for a reason. Usually, there is a management dilemma, a decision to be made, curiosity to be satisfied, or knowledge and understanding to be gained.
 - b. The reasons or situation for the study must be adequately described.
 - c. Next, research questions or hypotheses are framed to guide the construction of the measure to ensure that the data collected will contribute to answering the research questions or test the hypothesis.
 - d. Often, it is necessary to develop sub-questions for each of the research questions or hypotheses to further clarify what data needs to be collected and from where; this is an iterative process.
 - e. Some recommend that it is often necessary to disguise a study's purpose and sponsorship. While there may be good reasons for this in a clinical environment, the authors see no reason to disguise a study's purpose or its sponsorship in a management, educational, or training environment.
2. Second, the research methodology (e.g., interview, survey, focus group, database research, etc.) is determined.
3. Third, actually write the items to be included on instrument. For each item, the following questions should be answered.
 - a. Ensure the item related to the research question or hypothesis. While it is nice to know interesting information, ask only what you need to know.
 - b. Ensure the item is focused and complete.
 - (1) Prioritize the information you need from most critical to least critical. Focus first on constructing items which will provide you the most critical information. Repeat this process until you have items designed to generate the information you need or until the measure is just too long. It is always better to design focused items which will provide you the needed information.
 - (2) Avoid double-barreled items, which is one item asking two questions. An example is "What is your height and weight?" These items are confusing and frustrating. Each item should require the respondent to provide a single complete response.

- (3) Is the item absolutely precise? Each item should be phrased using the most precise wording so as to elicit the needed information. Item writing is a time consuming, labor intensive process. Spend whatever time is necessary to phrase precise items. Have a knowledgeable colleague or two to review and edit your best efforts.
- c. Ensure the intended respondent able to answer the question.
- (1) Respondents will need time to answer the item.
 - (a) As an instrument designer, you need to be familiar with the nature of the respondents, the estimated amount of time for responding, and the research or evaluation context.
 - (b) The key question you will need to ask is, “Does the respondent have enough time to think of a response?” Close-ended items are very efficient as the respondent doesn’t need to write a response as required by an open-ended item.
 - (2) Some respondents will try to participate in the study whether they are eligible or not. To separate out ineligible respondents, use filtering or qualifying items to determine eligibility. These qualifying items should be sufficient in number so as to ensure only eligible subjects/respondents participate.
 - (3) Avoid constructing leading items. A leading item is one where the subject responds to an implicit or explicit prompt. This is one reason precise wording is needed. When selecting the words to comprise an item, choose the most objective word possible.
 - (4) The item writer must strike a balance between generality and specificity in writing items. This is best achieved by knowing exactly what information the item is intended to produce and then selecting the most precise words for that item. Constructing an item to be general or specific neither is good nor bad, just appropriate or inappropriate. The criterion is the information the item is intended to generate to answer the research question or test the hypothesis.
 - (5) Put the most sensitive items towards the end of the instrument. Use categories (e.g., \$10,000 to \$19,999) when asking about income or other intimate information. There is debate about where relevant demographic items should be placed. The authors often use demographic items to “break the ice” at the first part of a survey instrument, provided we’ve no evidence or suspicion that doing so will “put-off” respondents.
 - (6) Avoid items which require respondents to recall information from the distant past. The longer the time frame a respondent must remember back to, the greater the chances of error, false memories, and inability to respond. To avoid recall and memory decay, keep the actual event and study as close to “real time” as possible.
- d. Select an efficient item response strategy.
- (1) The study’s purpose must drive the selection of response strategies (i.e., how the respondent or subject answers each item on the instrument). Key considerations in making this selection are respondents’ educational attainment, reading level, knowledge, motivation, writing and vocabulary skills, etc. Select response options with these in mind.

- (2) Close-ended items tend to require the least motivation, eliminate the need for high level writing skills, and generate data which can be used to compute a measure's reliability coefficient. However, close-ended items are labor intensive and time consuming to write. Instruments composed largely of close-ended items should be pilot-tested for readability, meaning, and completeness with a sample of intended subjects, who are later not included in the actual study. Close-ended items are excellent for assessing demographic characteristics, sentiments, and behaviors, provided the item stem is sufficiently well developed. Response options should always be mutually exclusive.
 - (3) Open-ended items are useful for assessing levels of knowledge (but there are other more appropriate and efficient strategies for assessing learning and knowledge), opinions, response frame of references, and respondent communication skills, (e.g., vocabulary, spelling, complex writing, reasoning skills, etc.).
- e. Write the Items.
- (1) Items should be framed in the vocabulary of subjects. Profane, insulting, or trigger (e.g., politically incorrect) language is avoided. If the item writer doesn't share the vocabulary of intended subjects, then an advisory group of prospective subjects should be convened. This group will not only assist the item writer in constructing items, but also help avoid biased wording, scenarios, etc. The items are written to generate specific information and must be in a language the intended subjects understand.
 - (2) Items can be written based on assumptions (frequently inaccurate) about intended subjects. Don't assume subject characteristics and then base item construction on those assumptions. Know your subjects, and base item writing on knowledge.
 - (3) Know the subjects' or respondents' frame of reference. Anticipating answers will help devise response options for close-ended items and for framing highly focused open-ended items. If the researcher is unsure as to the subject's frame of reference, use open-ended questions at least on the pilot-test.
 - (4) If you are using close-ended items, ensure that all reasonably possible alternatives are included in your response options. One way to do this is to provide "Don't Know or Refused" and "Other: _____" response options. What is adequate is determined by the intent of the item.
 - (5) Write "scorable" items.
 - (a) For quantitative instruments (e.g., surveys, tests, etc.) ensure the items are scorable, i.e., can produce an item subtest score and total score. See Appendix 4.9 for more information.
 - (b) For examples of quantitative instruments, see Appendices 3.1, 4.1, 4.2, 4.3, 4.4, 4.5, and 4.7.

- f. Compile the items in an instrument.
 - (1) Write a clear, brief introduction directly on the instrument to be administered or read.
 - (a) State the purpose of the study; generally, what the respondent will be asked to do; and stress the importance of participation as well as its nature: anonymous, confidential, or public.
 - (b) If the measure is self- or group-administered, ensure that the directions are clear and complete; provide an example of how to respond to the items.
 - (2) Next, you transition to your qualifying items. Qualifying items ensure you get the subjects or respondents needed for the study.
 - (a) Examples of qualifying items are:
 - [1] Do you live in Pasco County? (County residents)
 - [2] Are you over 18 years old? (County resident adults)
 - [3] Did you register to vote in the last election? (Registered voters)
 - [4] Did you vote in the last election? (Likely voters)
 - [5] Will you vote in the next election? (Intending voters)
 - (b) These five (5) items are intended to “qualify” respondents as the survey targets likely registered, adult voters, who have a history of and the intention to vote in the next election.
 - (c) If the intended subject qualifies, continue by introducing the next set of items. If not, politely terminate the interview or survey.
 - (3) Since the response strategies (i.e., how the items will be answered) have already been determined, the issue of item sequencing is raised. Follow these general guidelines:
 - (a) Use an “ice breaker” item to attract interest and motivate.
 - (b) Move from general items to increasingly specific ones.
 - (c) Place the more sensitive items towards the end of the measure. This also applies to open-ended items which require more time, thought, and effort than close-ended items.
 - (d) Ensure skip patterns are complete and accurate.
 - (e) It is a good idea to group similar items together and to explain to the subject what the next few items are about. This explicit help assists the subject in adjusting his or her response frame of reference, which leads to more accurate and useful information.
 - (4) At the end of the instrument, thank the respondent for participating and add a sentence or two about how important the participation was and how much it is appreciated.
4. Fourth, pilot test the instrument.
 - a. Unless the instrument is very short or constructed by a major deity, pilot-testing (or field-testing) is needed.
 - b. Select a group of intended subjects and administer the instrument to them and check for item wording, sensitivity, and meaning. Evaluate the instrument for clarity of directions; item continuity and flow; skip pattern accuracy; and compliance with the development plan; and ability to motivate respondents.

- c. If there are enough correctly completed instruments, implement your data analysis plan. This will give you the opportunity to determine whether or not the measure produces the information needed to answer the research questions and test the research hypothesis.
 - d. Provide space on the form for pilot-test respondents to recommend changes in item wording, to record any thoughts, confusions, or frustrations, they might experience.
 - e. Determine whether or not the instrument provides the needed information to answer the research question(s) or test the research hypothesis.
5. Fifth, revise the instrument items based on what was learned in the pilot test. Make one final pass through the quality control panel. Once the panel agrees the instrument is in final form, the study is ready to begin.
6. Sixth, using pilot test data develop a data analysis plan to ensure that the data collection tools (e.g., tests, surveys, etc.) will give you the needed information to answer your evaluation research question(s) or test your hypothesis.
- a. First, describe your data set (i.e., data distribution) using descriptive statistics. Descriptive indices will help you more fully understand your data.
 - b. Second, if you are testing for significant differences or associations between groups, apply the correct statistical test or tests to your data based on your research design. For statistical analysis options, see statistical textbooks, or a statistical consultant. It's important that the correct analysis be done; so interpretations are accurate.
 - c. Report your results to knowledgeable colleagues who understand the study's context and your results; they should be able to help you determine if your data collection instrument(s) and data analysis plan will provide the information needed to answer the research question or test the hypothesis.

C. Determining the Practicality of an Instrument

1. Determining the practicality of a measure involves assessing its economy, convenience, and interpretability.
 - a. Economy
 - (1) The instrument must be economical in terms of administration and completion time, associated costs, and data collection method(s), and scoring.
 - (2) Economy can be achieved by selecting those instruments which are self-administered, machine scored, and as short in length as possible.
 - (3) Factors which influence ease of administration are clear, concise directions; clean, tight item layout; and convenient response mechanism.
 - (4) No instrument should take more than 30 minutes (preferably 12-15) to complete. Field-test the instrument for understanding, readability, completion time, and item suitability.
 - b. Convenience
 - (1) The measure must be conveniently readable, set with clear type and uncluttered in format.
 - (2) Spelling, syntax, and punctuation must be correct.

- (3) Items, either questions or statements, must be explicitly clear in language so that the respondent understands what is being requested. Sensitive items should be towards the end of the form as should any demographic items.
 - (4) Skip-patterns across items should be kept to an absolute minimum. See item 6 in Appendix 4.7.
- c. Interpretability
- (1) Interpretation may become an issue if someone else other than the test or instrument designer does the data interpretation. Select an appropriate measurement scale that most will agree with and this potential problem will most likely be avoided.
 - (2) Whoever interprets the results must be qualified in terms of knowledge and skills and also have the interpretative tools necessary.
2. A measure's practicality involves a balance between economy, convenience, and interpretability. The researcher or evaluator makes this determination.

II. Selecting and Writing Measurement Items

- A. Data are collected directly (primary data collection) or indirectly (secondary data collection). The instruments in Appendices 4.1 to 4.5 and 4.7 are primary data collection instruments or tools. Secondary research is the analysis of previously collected primary data. Sources of secondary data include professional or trade journals, data mining from big datasets, organizational or historical documents, etc.
1. When collecting primary data, we must first determine what demographic and socioeconomic data (e.g., age, gender, occupation, education, income, race and ethnicity, marital status, social class, etc.) are important for the study. Demographic socioeconomic characteristics (also called, variables) are usually nominal data don't suit themselves to traditional measures of validity and reliability, but do contribute to describing the population or sample studied.
 - a. Often these "variables" are used as sorting keys to examine ordinal, interval, or ratio data to describe or compare groups on variables of interest.
 - (1) Examine items 1, 2, 3, and 4 in Appendix 4.1. These are demographic characteristics for this survey. These items provide respondents with specific options from which to select one response. (e.g., using Appendix 4.1, Item 1 [Payroll classification] to see if GTA-M's differed from GLA-D's in types of assignments given [Item 5] or testing strategies applied [Item 6] or use of testing devices [Item 6]. Demographic and socioeconomic data (variables) lets investigators (researchers) compare groups.
 - (2) It was noticed that master's or doctoral level graduate teaching assistants (GTA-M or GTA-D, respectively) used fewer student learning assessment strategies than either assistant or associate professors. Armed with this information, training programs were prepared for the GTA-M or GTA-D students to expand their use of different assessment strategies.
 - b. In Appendix 4.4 Item 11, "type of course for which tutoring was received" is a demographic variable and was indicated by placing a "√" in the corresponding blank. Next, the investigators used Subtest A responses to draw a profile of each

tutoring subject area (e.g., ENG 121, ENG 122, Science, etc.). Based on these profiles, changes were made as needed.

2. Personality and lifestyle Characteristics are routinely measured to help explain or predict behavior or decision-making.
 - a. Examine Items 1 to 10 in Appendix 4.2. These Likert scale items are used to measure supervisor communication characteristics. From these data, a communications consultant can explain communications style or predict a supervisor's style before meeting him or her.
 - b. In Appendix 4.5 are fifteen items (actually pairs of contrasting adjectives) which require a respondent to check along a seven point continuum. This is called a semantic differential scale, the higher the score, the higher, the self-concept.
3. Attitudes, Opinions, Intentions, and Motivation are also collected to assist in documenting intended behavior, to explain present behavior, or predict future behavior. The Attitude Taxonomy is found in Appendix 4.8.
 - a. Attitudes reflect a person's preference or feelings about a particular issue or phenomena; whereas opinions are articulated (verbally or written) attitudes. Motives are internal urges, needs, drives, wants, etc. that influence behavior. Intentions are expressions of anticipated behavior.
 - b. Items 1 to 17 in Appendix 4.3, ask respondents to indicate their degree of disagreement or agreement with specific statements regarding a freshman introduction to college course; the higher the score (highest is 85 or $17 * 5$), the greater the degree of agreement or satisfaction.
4. Knowledge and/or awareness are often terms which are used interchangeably. However, awareness is a rather shallow version of knowledge.
 - a. Having an awareness of how to drive doesn't mean that a person knows how to drive. While there are many definitions of knowledge, most cognitive psychologists define knowledge as a set of intellectual skills. For a manager, it is vital that subordinates possess the required intellectual skills to perform their jobs.
 - b. Knowledge and skill performance are most often assessed via a classroom, competency, or licensure test or skill demonstration. The process for constructing tools to measure knowledge and skill are discussed in Chapter 5.
5. Behavior is most often researched using the self-report strategy. While physical in nature, behavior should be described in terms of activity, time, location, circumstances, actors, and the actors' role(s).
 - a. Items 5, 6, and 7 in Appendix 4.1 measure behavior using an activity Likert Style item format.
 - b. Items 22, 27, and 28 in Appendix 4.3 describe an actors' behavior as does Item 4 and Items 8-9 in Appendix 4.4. Items 5, 6, and 7 in Appendix 4.1 ask a respondent to describe or measure his or her behavior with respect to strategies to assess student learning.

6. Next, we will examine two types of data collection tools (i.e., instruments or measures): Rating Scales and Ranking Scales.

B. Rating Scales

1. Rating scales are used when respondents don't directly compare attributes, characteristics or attitudes; they report their individual attitudes, emotions, or knowledge.
 - a. Five standard item formats are presented. An actual rating scale may be composed of any, one or more of these types of items.
 - b. Individual scores are computed by adding responses to each item on the scale. See the scoring directions for Appendices 3.1 and 4.5 as examples.
2. Single Category or Dichotomous Scale
 - a. There are only two response options available to respondents. Examples are: male/female, agree/disagree, yes/no, etc. See Appendix 5.5.
 - b. These types of items produce nominal level data. The most appropriate Measure of Central Tendency (MCT) is the mode (see Chapter 9). The number and percentage of respondents endorsing (i.e., selecting) each response option is reported.
3. Multiple Choice—Single or Multiple Response(s)
 - a. When this item format is applied to measuring attitudes, subjects will elect any or all options as is appropriate.
 - b. Multiple choice items typically produce nominal level data. The most appropriate MCT is the mode (see Chapter 9). The number and percentage of respondents endorsing each response option is reported.
 - c. Examples are Appendix 4.1, Items 1-4 and Appendix 4.4, Item 11.
4. The “Likert Scale”
 - a. This item format is typically used to measure attitudes, self-report behaviors, preferences, values, etc. See Appendix 4.6 for examples.
 - b. This item format is easy to construct and is expressed as complete statements (Appendix 4.3, Items 1–34 or Appendix 4.4, Items 1-10), sentence fragments (Appendix 4.3, Items 1-34), and may be very specific (Appendix 4.2, Items 1-10) or more general (Appendix 4.4, Item 10).
 - c. Likert originally recommended a five point scale with equivocation, such as “Strongly Disagree”, “Disagree”, “Neutral or No Opinion”, “Agree” and “Strongly Agree”. “Neutral or No Opinion” is the equivocation option.
 - (1) If respondents endorse one of these sets of words, then ordinal data are produced which require the use of non-parametric or distribution-free statistics.
 - (2) If numerical Likert Scales are used reliability coefficients can be computed as well as an item mean, median and mode (see Chapter 9). The number and percentage of respondents endorsing (i.e., selecting) each response option is reported.

- d. If the “Strongly Disagree”, “Disagree”, “Neutral or No Opinion”, “Agree” and “Strongly Agree” continuum is formatted with numbers representing the word groupings (Appendix 4.3, Items 1-35), then interval data are produced and parametric statistics are applied. In this case, any or all of the MCT’s are appropriate; and Measures of Variation (e.g., the range and standard deviation) are typically reported for each item and the total score for the instrument or measure. The number and percentage of respondents endorsing each item response option is also reported.
- e. To maximize variance, which contributes to scale or subtest reliability, a six point continuum, with no equivocation (i.e., no “Neutral or No Opinion”) is optimal. However, such a response continuum may not be practical.
- f. Indexes, scales or subtests, composed of sets items in the Likert scale format, lend themselves to traditional notions of validity and reliability. We should ensure that the appropriate type(s) of reliability and validity indices are documented.

5. Stapel Scale

- a. The format is an alternative to the Semantic Differential scale (Appendix 4.5).
- b. A construct is identified, (e.g., brand image), then attributes or characteristics of that construct are identified (e.g., high quality products, highly trusted, well respected, well known, etc.)
- c. Positive and negative rating continua are placed next to each characteristic as below:

+5	+5	+5
+4	+4	+4
+3	+3	+3
+2	+2	+2
+1	+1	+1
Highly Trusted	Well Respected	Well Known
-1	-1	-1
-2	-2	-2
-3	-3	-3
-4	-4	-4
-5	-5	-5

- d. The more the respondent thinks or feels the characteristic describes the brand, the higher will be the endorsement (+1 to +5). The less the respondent feels the characteristic describes the brand, the lower the endorsement (-1 to -5).
- e. Individual item descriptive statistics can be computed as this item format produces interval level data. In this case, any or all of the MCT’s are appropriate; and Measures of Variation (e.g., the range and standard deviation) are typically reported (see Chapter 9). The number and percentage of respondents endorsing each response option is also reported.

6. Semantic Differential Scale

- a. The semantic differential scale is commonly used in measuring attitudes, preferences, or values. See Appendix 4.5 for an example.
- b. Bipolar adjectives are identified in such numbers that a complete description of the attitude, preference, or value is fully described.

- c. Expert judges are used to sort the best bipolar pairs into “piles” ranging from the least to most descriptive. Those bipolar pairs which are most descriptive are then organized into a scale, with equidistant intervals separating each anchor of the bipolar pair. A seven point continuum is most commonly used.
- d. To reduce response set (i.e., a subject responding in a consistent manner that does not represent his or her “real” sentiments) several score weights are reversed. For example:
 - (a) Eager (7)..... Indifferent (1)
 - (b) Useless (1)..... Useful (7)
- e. This item format produces interval level data and parametric statistics may be applied. In this case, any or all of the MCT’s are appropriate; and Measures of Variation (e.g., the range and standard deviation) are typically reported. The number and percentage of respondents endorsing each response option is also reported.

C. Ranking Scales

1. Ranking scales require the respondent to compare two or more attributes, characteristics, or attitudes. Respondents may be asked to select which pair of glasses is more attractive or be asked to rank order the importance of color, style, fit, tint, and cost of glasses. Three standard item formats are presented.
 - a. In ranking scales, the respondent compares two or more options and makes a preferred choice, or rank orders his or her preference.
 - b. The median and mode are the most appropriate MCT’s (see Chapter 9). The number and percentage of respondents endorsing each response option is reported for these item formats.
 - c. These item formats produce ordinal data and requires the application of non-parametric or distribution-free statistical procedures. While ordinal data can be transformed into interval data, by using the standard normal curve, it is often easier to apply non-parametric statistics.
2. Paired-Comparison Scale
 - a. The subject compares specified objects, (e.g., similar products or services) and selects from each paring the option most preferred. If three objects are to be compared, then $[(n)(n-1)/2]$ is the number of pairs needed, where n is the number of stimuli or objects (Cooper & Schindler, 2013). Using the formula: $[(3)(2)/2] = 3$ paired comparisons are needed. For example:

We are interested in knowing which types of soup you prefer to be served in the cafeteria. We plan to offer two soups daily during the winter months. One regular soup will be offered at all times along with an alternate. Place an “x” in the blank by the soup you most prefer if you had to choose.

- | | | | |
|--------------------------|----------------|--------------------------|-----------|
| <input type="checkbox"/> | Tomato | <input type="checkbox"/> | Vegetable |
| <input type="checkbox"/> | Chicken Noodle | <input type="checkbox"/> | Tomato |
| <input type="checkbox"/> | Vegetable | | |
| <input type="checkbox"/> | Chicken Noodle | | |

- b. To avoid, respondents tiring, keep the number of paired-comparisons between a maximum of six to ten. One other limitation is that ties between response options (tomato, chicken noodle, or vegetable soups) do happen.

3. Forced Ranking Scale

- a. Subjects are given a list of response options and asked to rank order each against the others based on some order of preference. No more than five to 10 options are recommended. For example:

Presented below are characteristics prospective students consider when selecting a college or university from which to earn an MBA. Please rank each characteristic from 1 (least important) to 5 (most important) that you considered when selecting this college or university.

- a. Academic Reputation of the MBA program
- b. Cost of Attendance (e.g., tuition, books, etc.)
- c. Convenience to home or work
- d. Week-end Class
- e. Number of months required to earn the MBA

- b. Items in this format are relatively easy to construct and subjects are usually fairly motivated to answer.

4. Comparative Scale

- a. A new product, service, or sentiment is compared against a standard in this item format. For example:

Think back to your senior year in college. In terms of performance expectations, how does the rigor of your MBA course work compare to that of your senior year? Circle the number that represents your choice.

Less Rigor		About the Same Rigor		Much Higher Rigor	
1	2	3	4	5	6

- b. This item format requires a known comparison standard and could be used for benchmarking best practices. In this case the comparison standard is the degree of rigor measured from 1 (Low Rigor) to 6 (High Rigor).

D. Open-ended items and Selected Qualitative Strategies

1. Open-ended Items
 - a. Each of the item formats presented above are also called close-ended items. The respondent is required to select from an array of pre-determined response options. Virtually no variability in subject responses is allowed.
 - b. Open-ended items require focused, but unstructured responses by respondents. (Appendix 4.3, items 36-38.) Open-ended items may be used with either rating or ranking scales.
 - c. These items should be highly focused and extensively field-tested to ensure that the desired topical response is provided by the subject.
2. Selected Qualitative Strategies
 - a. These include logs, diaries, or journals.
 - b. Described are activities, experiences, and/or feelings written during participation in a course, program, intervention or experience. These consist of running entries which are written at specific intervals (daily, weekly, etc.). Entries in journals are typically longer than those in logs. Logs and journals typically are employed to report on others. Diaries report primarily report on the writer.
3. Comments Common to both Approaches
 - a. While allowing for greater variability in responses, these strategies should be content analyzed and coded which is time and labor intensive. Another strategy is to list, verbatim, the response(s) and present them in an appendix to the report.
 - b. These types of items are very useful in revealing patterns, providing context, generating research ideas, stimulating close-ended item development, and encouraging respondents to express themselves.

Review Questions

Directions. Read each item carefully; either fill-in-the-blank or circle letter associated with the term that best answers the item.

1. Which item below is open-ended?
 - a. What do you think of your manager's leadership style?
 - b. Have you ever "called in sick" and not be?
 - c. What computer operating system do you prefer—PC or Mac?
 - d. All are open-ended items.
2. _____ should be described in terms of activity, time, location, circumstances, actors, or actors' role.

a. Socioeconomic characteristics	c. Lifestyle characteristics
b. Intentions	d. Behaviors

3. These measures require a respondent to compare two or more attributes or characteristics
 - a. Rating scale
 - b. Ranking scale
 - c. Categorization scale
 - d. Open-ended indexes
4. A scale or index which measures a single attribute or characteristics is called _____.
 - a. Unidimensional
 - b. Multidimensional
 - c. Triangulated
 - d. Open-ended
5. Which one of the following statements concerning critical characteristics of a measure is not accurate?
 - a. Conveniently readable, set with clear type and uncluttered in format.
 - b. It should be as brief as possible with correct spelling, syntax, and punctuation.
 - c. No instrument should take more than 40 minutes to complete.
 - d. Field-test the instrument for understanding, readability and item suitability.
6. Which one of the following statements concerning critical characteristics of a measure is not accurate?
 - a. Items, either questions or statements, must be explicitly clear in language so that the respondent understands what is being requested.
 - b. Sensitive items should be towards the front of the form, as should any demographic items.
 - c. Skip-patterns across items should be kept to an absolute minimum.
 - d. Extensive prior preparation is required before launching any survey.
7. Knowing the type of data (nominal, ordinal, interval, or ratio) an item is intended to produce is important. Which one of the following statements is not true?
 - a. Knowing the type of data that each item on the measure will produce enables one to select the most appropriate statistical indices and/or tests.
 - b. Once we know type of data need, we can construct items, subtests, etc. which will provide those data.
 - c. Demographic variables which are nominal or ordinal are well suited to traditional measures of validity and reliability.
 - d. Demographic variables contribute to documenting the representativeness of the sample.
8. Which one of the following statements about constructing a measure is not true?
 - a. Every study is conducted for a reason.
 - b. Research questions or hypotheses are framed to guide the construction of the measure.
 - c. If the research question or hypotheses is well phrased, it is rarely necessary to develop sub-questions for each research question.
 - d. Once data needs have been established, the communication strategy (e.g., interview, survey, focus group, database research, etc.) between researcher and respondent (or other data sources) is determined.
9. Which general item format is most efficient?
 - a. Closed-end
 - b. Open-ended

10. Which one of the responses listed below best speaks to an item's precision?
- Collect only information one needs to know
 - First collect only the most critical information
 - Avoid double-barreled items
 - Item writing is a time consuming, labor intensive process
11. Which response option presents the first four steps in constructing items for a measure in the correct order?
- "Does the question need to be asked?"; "frame the item"; "select the response category"; and "Can the item be answered?"
 - "Does the question need to be asked?"; "select the response category"; "frame the item"; and "Can the item be answered?"
 - "Does the question need to be asked?"; "frame the item"; "Can the item be answered?" and "select the response category."
 - "Does the question need to be asked?"; "Can the item be answered?"; "frame the item"; "select the response category."
12. When evaluating the practicality of a measure, each of the following are critical elements, except:
- Economy
 - Precision
 - Interpretability
 - Convenience
13. Which one of the following is not a rating scale?
- Likert scale
 - Stapel scale
 - Semantic differential scale
 - Paired-Comparison scale
14. The scale composed of bipolar adjectives is called?
- Likert scale
 - Stapel scale
 - Semantic differential scale
 - Paired-Comparison scale
15. The scale where subjects order response options based on preference is called?
- Forced choice scale
 - Forced ranking scale
 - Pair-Comparison scale
 - Comparative scale
16. Which one of the following statements is not accurate?
- In responding to close-ended items, respondents have variability in responses.
 - Open-ended items may be used with either rating or ranking scales.
 - Open-ended items require focused, but unstructured responses by respondents.
 - Open-ended items should be highly focused.

17. Regarding qualitative approaches to measuring attitudes which one of the following statements is not accurate?
- These include logs, diaries, or journals.
 - These (logs, diaries, or journals) consist of running entries which are written at specific intervals (daily, weekly, etc.).
 - Entries in journals are typically longer than those in logs.
 - Logs, journals, and diaries typically are employed to report on others.

Answers: 1. a, 2. d, 3. b, 4. a, 5. c, 6. b, 7. c, 8. c, 9. a, 10. d, 11. d, 12. b, 13. d, 14. c, 15. b, 16. a, 17. D

Reference

Cooper, D. R. & Schindler, P. S. (2013). *Business research methods* (12th ed.). Columbus, OH: McGraw-Hill Irwin.

Appendices

Appendix 4.1 is an instrument to document faculty strategies used to assess student academic performance at the University of Georgia in 1996. It is a complex, multidimensional (3 dimensions) measure. There are four demographic items which were treated as independent variables. The three subtests in Part B (one per dimension) were to assess aspects of the operationally defined dependent variable, student assessment. Scores were summed, described, and compared. You will note that there were several nominal categories within each of the four demographic items. Once data collection was completed on this project, we found that there were not enough data to use all categories. We dropped items 2 to 4 and collapsed item one into three categories, graduate assistant, adjunct faculty, and full-time faculty. We found no differences between groups. The instrument was reviewed by three experts to ensure content validity. Since it is not possible to compute reliability indices for nominal or ordinal data, Cronbach's alpha was applied only to the three subtests in Part B with a range of 0.70 to 0.79.

Appendix 4.2 is a unidimensional index to measure a supervisor's communication effectiveness. The maximum score is 50; there is no classification schema. Note the brief interpretation guidelines. Clinical measures are typically more complex with complicated scoring procedures. Don't treat a simple index like this as a clinical instrument. We have no knowledge of the theory or research, if any, upon which this measure is based.

Appendix 4.3 is a measure to evaluate a freshman transition program whose chief element was a four credit introduction to college course, using a combination of Likert scale items grouped to assess program outcomes, course characteristics, advising, and orientation. Likert items of the form presented are useful as responses can be summed to a subtest (i.e., A, B, and C) or a total score. Scores are necessary if descriptive or inferential statistics are to be applied to the data. Open-ended questions and statements were used to encourage honest unstructured responses, as it was impossible to anticipate all probable responses to the three open-ended items. There were no demographic items as there was no intention to sort responses for comparison purposes.

Appendix 4.4 is a survey instrument designed to measure student satisfaction with tutoring services provided by a university's academic support program. Items one to ten are intended to tap student perceptions about the tutoring environment and to render an overall effectiveness judgment. Item 11 is a demographic item whose purpose was to sort and compare scores from items one to 10 to ascertain whether or not one set of tutoring services was perceived to be more effective than another. Item 12's purpose was to compare scores based on service intensity as defined by use. It is always a good idea, space permitting, to include an open-ended item that encourages the respondent to communicate whatever else he or she desires; use a focused question or statement.

Appendix 4.5 is a unidimensional self-concept scale using semantic differential scale. Note the scoring methodology.

Appendix 4.6 presents a compilation of response options which can be used for Likert style items.

Appendix 4.7 presents the Adult Academic Credit Participation Index which is based on a theory of adult education persistence (i.e., studying in University to complete academic goals).

Appendix 4.8 presents the Attitude Taxonomy which provides guidance on measuring particular types of attitudes.

Appendix 4.9 presents a discussion on scoring and reporting scorable items for quantitative data collection tools (e.g., surveys, attitudinal indexes or scales, etc.).

Appendix 4.1 Strategies Used for Assessing Student Learning Survey

Dear Colleague:

The University of Georgia requires a periodic review of all of its programs. The Georgia Center and the Department of Evening Classes are currently under going their required review. The Department is collecting data, as required, to describe how student learning is assessed by its teaching faculty. Accordingly, you are requested to complete this short, anonymous survey form. **Please return the completed form to JRL 304 or, via campus mail in the enclosed envelope, to Dr. Charles D. Hale, Rm. 196 in the Georgia Center.**

Part A: Please answer the few demographic items presented below.

1. What is your Fall 1996 payroll classification? _____ ()

- | | | |
|----------|-------------------------|------------------------|
| 1. GTA-M | 3. GLA-D | 7. Assistant Professor |
| 2. GTA-D | 4. PTI | 8. Associate Professor |
| 3. GLA-M | 6. Full-time Instructor | 9. Full Professor |

(GTA-M=Masters degree seeking graduate teaching assistant; GTA-D=doctoral degree seeking graduate teaching assistant; GLA-M=masters degree seeking graduate lab assistant; GLA-D=doctoral degree seeking graduate lab assistant; PTI=part-time instructor, not a full-time UGA faculty member.)

2. Considering all of your teaching, regardless of location and level (e.g., public school, college, or university), how many years full-time experience do you have? _____ ()

- | | | |
|-------------------|------------------|---------------------|
| 1. 1 year or less | 3. 6 - 9 years | 5. 13 - 15 years |
| 2. 2 - 5 years | 4. 10 - 12 years | 6. 16 or more years |

3. What is the course level of the course or lab(s) you are teaching during the 1995 fall quarter? _____ ()

- | | | |
|----------------------|--------------|--------------|
| 1. ACA or UNV course | 3. 200 level | 5. 400 level |
| 2. 100 level | 4. 300 level | |

4. In which of the following disciplines would you place this course? _____ ()

- | | | |
|----------------------------|--------------------------|---------------------|
| 1. Agricultural Sciences | 4. Education | 7. Natural Sciences |
| 2. Business Administration | 5. English or Literature | 8. Social Sciences |
| 3. Consumer & Food Sci. | 6. Foreign Languages | 9. Other: _____ |

Part B: The following questions ask about the types of assignments, examination strategies, and testing devices you use to assess student learning.

Important Definitions: An assignment is defined as academic work completed by a student who often views such as "homework" or preparation for an examination or test. An examination or test is defined as an announced point(s) within the course when a student is required to demonstrate his or her knowledge, skill, or understanding under conditions which are generally construed as testing by both students and faculty.

5. How frequently do you use or plan to use any of the following types of assignments or exercises (not tests) to assess student learning in this course, using this scale:

Never (**N**), circle 1
Rarely (**R**), circle 2
Occasionally (**O**), circle 3

Often (**OF**), circle 4
Frequently (**F**), circle 5
Very Frequently (**VF**), circle 6

	N	R	O	OF	F	VF	
a. Turned in computational problem sets	1	2	3	4	5	6	()
b. Turned in individual or group project(s)	1	2	3	4	5	6	()
c. Short essay or theme style papers	1	2	3	4	5	6	()
d. Turned in student lab books or logs	1	2	3	4	5	6	()
e. Individual or group presentation(s)	1	2	3	4	5	6	()
f. Unannounced quizzes	1	2	3	4	5	6	()
g. Individual or group student conferences	1	2	3	4	5	6	()

6. Using the same response scale in Question Five, how frequently do you use or plan to use any of the following testing strategies to assess student learning in this course?

	N	R	O	OF	F	VF	
a. Take-home examinations	1	2	3	4	5	6	()
b. In-class closed book examinations	1	2	3	4	5	6	()
c. In-class open-book examinations	1	2	3	3	5	6	()
d. Objective (e.g., true/false) examinations	1	2	3	4	5	6	()
e. Subjective (e.g., essay) examinations	1	2	3	4	5	6	()
f. Unit or mid-term examination(s)	1	2	3	4	5	6	()
g. Cumulative final examination	1	2	3	4	5	6	()
h. Non-cumulative final examination	1	2	3	4	5	6	()
i. Individual student examinations	1	2	3	4	5	6	()
j. Group examinations	1	2	3	4	5	6	()
k. Individual or group portfolios of student work	1	2	3	4	5	6	()
l. Individual term paper or project	1	2	3	4	5	6	()
m. Group term paper or project	1	2	3	4	5	6	()
n. Other: _____	1	2	3	4	5	6	()

7. Using the same response scale in Question Five, how frequently do you use or plan to use any of the following testing devices to assess student learning in this course?

	N	R	O	OF	F	VF	
a. Multiple choice test/quiz items	1	2	3	4	5	6	()
b. True-False test/quiz items	1	2	3	4	5	6	()
c. Matching test/quiz items	1	2	3	3	5	6	()
d. Short Answer test/quiz items	1	2	3	4	5	6	()
e. Computational problem sets for tests/quizzes	1	2	3	4	5	6	()
f. Essay test/quiz items	1	2	3	4	5	6	()
g. Oral tests or quizzes	1	2	3	4	5	6	()
h. Individual performance check lists	1	2	3	4	5	6	()
i. Group performance check lists	1	2	3	4	5	6	()
j. Other: _____	1	2	3	4	5	6	()

Appendix 4.2 Supervisor Communication Effectiveness

Purpose: The purpose of the instrument is to determine perceptions of a supervisor's effectiveness as a communicator. Communication effectiveness is the ability of the supervisor to communicate clearly to employees or teachers about work matters.

Directions: Read each statement carefully and then circle the degree to which your supervisor models the stated behavior. Circle the number to the right of the statement which represents your rating using the following scale:

1 = Almost Never
2 = Sometimes

3 = Often

4 = Frequently
5 = Almost Always

- | | | | | | |
|---|---|---|---|---|---|
| 1. My supervisor provides a clear vision of what our company is all about. | 1 | 2 | 3 | 4 | 5 |
| 2. My supervisor conducts formal discussions concerning the improvement of our products and services. | 1 | 2 | 3 | 4 | 5 |
| 3. My supervisor conducts formal discussions concerning employee productivity. | 1 | 2 | 3 | 4 | 5 |
| 4. Improved productivity results from discussion with my supervisor. | 1 | 2 | 3 | 4 | 5 |
| 5. My supervisor provides me with information on current job-related topics. | 1 | 2 | 3 | 4 | 5 |
| 6. My supervisor facilitates my participation in useful training opportunities. | 1 | 2 | 3 | 4 | 5 |
| 7. My supervisor promotes an ongoing review of job processes. | 1 | 2 | 3 | 4 | 5 |
| 8. My supervisor uses clearly established criteria for judging my performance on the job. | 1 | 2 | 3 | 4 | 5 |
| 9. My supervisor provides frequent feedback concerning my performance. | 1 | 2 | 3 | 4 | 5 |
| 10. My supervisor is available to address my job-related concerns. | 1 | 2 | 3 | 4 | 5 |

Scoring: Add the numbers representing your rating for each item. The closer to 50, the more effective your supervisor is perceived. This instrument isn't validated for diagnostic or evaluative purposes; so don't use it for such. Its purpose is to just stimulate discussion.

Appendix 4.3

FRM 100 Foundations of Learning and Knowing Student Survey

The purpose of this brief survey is for you to describe your experiences in this course, freshman advising, and freshman orientation. Your opinions are very important. Your responses will be grouped with other students completing this survey. The information will then be used to improve the course. Your responses are anonymous. Your professor will collect your completed surveys and return them to the University's Academic Assessment Program office. Thank you!

Please read each statement carefully and then circle the number that represents your degree of agreement or disagreement with the statement using the following options:

Strongly Disagree = 1
Disagree = 2

No opinion = 3

Agree = 4
Strongly Agree = 5

A. These first few items ask you to describe your FRM 100 Foundations of Learning and Knowing experience.

My experience in FRM 100 has helped me to:	SD	D	N	A	SA
1. Understand the purpose and role of universities in generating new knowledge and transmitting what is already known.	1	2	3	4	5
2. Understand the role of higher education in enabling students to learn and to become creative, critical thinkers.	1	2	3	4	5
3. Know the mission of Saint Leo University and the Catholic and Benedictine values that guide its behavior.	1	2	3	4	5
4. Establish personal goals and know how to plan to attain them.	1	2	3	4	5
5. Identify and understand different learning styles.	1	2	3	4	5
6. Determine what is <u>my</u> preferred learning style.	1	2	3	4	5
7. Be an active participant in both classroom learning activities as well as those outside the classroom.	1	2	3	4	5
8. Effectively manage my schedule so that I have enough time to study.	1	2	3	4	5
9. Efficiently take notes and tests.	1	2	3	4	5
10. Effectively use print and electronic tools to gather information for research papers and other related assignments.	1	2	3	4	5
11. Critically evaluate different perspectives on issues.	1	2	3	4	5
12. Critically evaluate differing views on what knowledge is and how humans learn.	1	2	3	4	5
13. Understand how values (e.g., respect for others, diversity, and responsible stewardship, etc.) guide behavior in an academic community such as Saint Leo.	1	2	3	4	5
14. Participate in discussions in a positive, respectful manner.	1	2	3	4	5
15. Work effectively in small learning groups.	1	2	3	4	5
16. Make good quality presentations.	1	2	3	4	5
17. Write more clearly, concisely, and effectively.	1	2	3	4	5

B. The next few items ask you to generally describe other selected aspects of your FRM course. Please use the same response options presented above.

	SD	D	N	A	SA
18. The purpose of this course was clear to me.	1	2	3	4	5
19. As far as you are concerned, the course met its stated purpose.	1	2	3	4	5
20. In general, the assigned readings were interesting.	1	2	3	4	5
21. In general, the out-of-class assignments were worthwhile given the purpose of the course.	1	2	3	4	5
22. Generally, the professor followed the syllabus.	1	2	3	4	5
23. This course helped me to adjust to college life.	1	2	3	4	5

	SD	D	N	A	SA
24. The joint class experiences were valuable in helping me learn.	1	2	3	4	5
25. The small group project on university values was successful in enabling me to learn.	1	2	3	4	5
26. The mid-term examination adequately covered material we had in class and/or worked on in assignments.	1	2	3	4	5
27. My FRM 100 professor/advisor helped me adjust to college life.	1	2	3	4	5
28. My FRM 100 professor/advisor has been effective in helping me learn given the course objectives.	1	2	3	4	5
29. Overall, I would rate <u>my</u> FRM 100 experience as excellent.	1	2	3	4	5

C. These last few items ask you describe your experiences with freshman advising and freshman orientation.

30. I like how my freshman courses are scheduled.	1	2	3	4	5
31. The new student orientation program adequately prepared me for what I experienced during my first semester at Saint Leo.	1	2	3	4	5
32. I know students with an eligible disability can receive an academic accommodation.	1	2	3	4	5
33. Overall, I would rate <u>my</u> freshman advising as excellent.	1	2	3	4	5
34. Overall, I would rate <u>my</u> orientation experience as excellent.	1	2	3	4	5
35. How many times did you meet with your freshman mentor/advisor? _____ (Please fill in the blank.)					

D. These last few items ask you to make specific recommendations to improve FRM 100, freshman course scheduling, the freshman orientation program, and any other service provided by the University. Please write neatly. Use more than one sheet if necessary.

36. What specific recommendation(s) do you have to improve FRM 100?
-
-
-
-
-
-
-
37. How would you improve freshman advising, class scheduling, and orientation?
-
-
-
-
-
-
-
38. Please describe how you would improve any other service provided by the University so that you may be better served.

Appendix 4.4 Student Tutoring Satisfaction Survey

The purpose of this brief survey is for you to describe your experience(s) with the tutoring services provided to you by the University's Learning Resource Center. The information you provide will be used to improve the tutoring services.

Your responses are anonymous, as they will be grouped with other students for reporting purposes. Please complete this brief survey and return it to the University's Academic Assessment Program office in the enclosed envelope. Thank you!

Please read each statement carefully and then circle the number that represents your degree of agreement or disagreement with the statement using the following options:

Strongly Disagree = 1
Disagree = 2

No opinion = 3

Agree = 4
Strongly Agree = 5

A. These first few items ask you to describe your tutoring experience(s)

	SD	D	N	A	SA
1. My tutoring was provided at a convenient time.	1	2	3	4	5
2. My tutor(s) had access to the facilities needed to assist me.	1	2	3	4	5
3. My tutor(s) had access to the equipment needed to assist me.	1	2	3	4	5
4. My tutor or tutors was/were on time for my appointment(s).	1	2	3	4	5
5. I was not hurried through my tutoring session(s)	1	2	3	4	5
6. The tutoring <u>environment</u> helped me learn.	1	2	3	4	5
6. My tutor worked with me so that I was able to learn what was needed to be successful in my class(es).	1	2	3	4	5
7. The procedure for obtaining tutoring was reasonably "hassle" free.	1	2	3	4	5
	SD	D	N	A	SA
8. University staff at the Learning Resource Center were polite.	1	2	3	4	5
9. University staff at the Learning Resource Center were helpful.	1	2	3	4	5
10. Overall, I found that the tutoring services provided to me were effective.	1	2	3	4	5

B. These final few items ask you to further describe your tutoring experience. Please write or "check" your response in the blank or space provided.

11. In what general subject or skill areas did you receive tutoring? (Check all that apply):

English 121 ___ Science_____ Math_____ Writing_____

English 122_____ Business_____ Reading_____ Other_____

12. How many tutoring sessions did you attend? _____

13. Please make specific recommendations to improve the effectiveness of the tutoring services provided by the University. Please continue on the back of this sheet if necessary.

**Appendix 4.5
Self-Concept Scale**

Purpose: Designed to measure an individual’s perception of “myself as a person.” It is intended as a global measure of self.

Directions: Reach each pair of adjectives carefully. Next, consider where on the continuum between each set of adjectives you see your-self most often. Indicate your position by marking an “X” in the blank provided.

1. Interested	_____	_____	_____	_____	_____	_____	_____	Apathetic
2. Inactive	_____	_____	_____	_____	_____	_____	_____	Active
3. Unbending	_____	_____	_____	_____	_____	_____	_____	Elastic
4. Take part	_____	_____	_____	_____	_____	_____	_____	Shun
5. Sluggish	_____	_____	_____	_____	_____	_____	_____	Vigorous
6. Authoritative	_____	_____	_____	_____	_____	_____	_____	Weak
7. Negative	_____	_____	_____	_____	_____	_____	_____	Positive
8. Diligent	_____	_____	_____	_____	_____	_____	_____	Idle
9. Disgusting	_____	_____	_____	_____	_____	_____	_____	Beautiful
10. Intelligent	_____	_____	_____	_____	_____	_____	_____	Unintelligent
11. Disagreeable	_____	_____	_____	_____	_____	_____	_____	Agreeable
12. Ineffectual	_____	_____	_____	_____	_____	_____	_____	Effective
13. Glad	_____	_____	_____	_____	_____	_____	_____	Sad
14. Gloomy	_____	_____	_____	_____	_____	_____	_____	Upbeat
15. Unsightly	_____	_____	_____	_____	_____	_____	_____	Attractive

Scoring: Going from left to right score using 1, 3, 3, 4, 5, 6, & 7. For eager, participating, powerful, hardworking, sharp and happy, score from right to left 1, 2, 3, 4, 5, 6, & 7. The higher the total score, the higher is your self-concept. Do not use this instrument for diagnostic or evaluative purposes. It is not validated for such.

Appendix 4.6 Likert Scale Item Response Formats

The response options presented below are commonly employed in the construction of measures using Likert scale style items. It is recommended that two or three point options are used with younger respondents or respondents who may find more response options confusing or who might not be willing to commit the time to complete a more complex instrument.

It is essential that the response options match the purpose of the item. It is not uncommon to find five point Likert scales associated with items requiring a binary “yes” or “no” response.

Six point response scales with no equivocation (i.e., lacking a neutral or undecided option) have been found to generate the most variance. Higher levels of variance are associated with higher reliability coefficients. Using scales which lack equivocation is a researcher’s choice; however, not allowing respondents a “neutral” or “undecided” option is a risk if there is a logical probability that such selections are likely.

It should also be noted that if the researcher is to construct his or her own semantic differential scale, polar binary options are a good place to start. Make sure that the polar binary options are logically related to the purpose of the semantic differential. A word-processing thesaurus is a good place to start.

A. Dichotomous or Binary Options

1. Fair...Unfair
2. Agree...Disagree
3. Yes...No
4. True...False
5. Good...Bad
6. Positive...Negative

B. Three Point Options

1. Exceed Expectations...Meet Expectations...Do Not Meet Expectations
2. Too Much...About Right...Too Little
3. Too Strict...About Right...Too Lax

C. Four Point Options

1. Most of the Time...Some of the Time...Seldom...Very Seldom
2. Strongly Disagree...Disagree...Agree...Strongly Agree
3. Exceeded...Met...Nearly Met...Not At All Met

D. Five Point Options

1. Almost Never...Sometimes...Often...Frequently...Almost Always
2. Strongly Disagree...Disagree...Neutral...Agree...Strongly Agree
3. Very High...Above Average...Average...Below Average...Very Low
4. Very Good...Good...Fair...Poor...Very Poor
5. Excellent...Above Average...Average...Below Average...Very Poor
6. Very Satisfied...Satisfied...Neutral...Dissatisfied...Very Dissatisfied

7. Completely Satisfied... Very Satisfied... Satisfied... Somewhat Dissatisfied... Very Dissatisfied
8. Extremely []... Very []... Moderately []... Slightly []... Not at all []
9. Very Inconsistently... Inconsistently... Neutral... Consistently... Very Consistently
10. Very Favorable... Favorable... Neutral... Unfavorable... Very Unfavorable
11. Met Few Expectations... Met Some Expectations... Met Most Expectations... Met All Expectations... Exceeded All Expectations

E. Six Point Options

1. Never... Rarely... Occasionally... Often... Frequently... Very Frequently
2. Very Strongly Disagree... Strongly Disagree... Disagree... Agree... Strongly Agree... Very Strongly Agree
3. Very Fast... Fast... Average Speed... Slow... Very Slow... Slowest
4. Highest Quality... High Quality... Good Quality... Average Quality... Low Quality... Lowest Quality
5. Highly Likely... Likely... Somewhat Likely... Somewhat Unlikely... Unlikely... Very Unlikely

F. Seven Point Options

1. Very Dissatisfied... Moderately Dissatisfied... Slightly Dissatisfied... Neutral... Slightly Satisfied... Moderately Satisfied... Very Satisfied
2. Very Poor... Poor... Fair... Good... Very Good... Excellent... Exceptional
3. Extremely Favorable... Favorable... Somewhat Favorable... Neutral... Somewhat Unfavorable... Unfavorable... Very Unfavorable
4. Fastest... Very Fast... Fast... Average Speed... Slow... Very Slow... Slowest
5. Extremely Important... Important... Somewhat Important... Neutral... Somewhat Unimportant... Unimportant... Extremely Unimportant

Appendix 4.7 Academic Credit Participation Index

Presented in Appendix 4.7 are a complete survey form, codebook, and report. In 1995, one of the authors conducted a study at the University of Georgia. The purpose of the survey was to describe students taking classes through the Department of Evening Classes (EVCL) in the Georgia Center for Continuing Education, using Cross' Chain of Response Model of Adult Learner Participation (1981, p. 125). Data generated by the survey was used to improve services and assist faculty and staff to better understand the students they served. Its purpose of here is to provide you a complete survey research case in the hopes that you will find it useful and instructive.

Student Participation Profile Department of Evening Classes

PART I Directions. These first few questions ask you to tell us a little about yourself. Please read each question carefully. You may record your answer by writing the number that represents your choice of answer in the blank provided, unless otherwise directed.

1. What is your **present** age in years? _____ (1)
2. What is your gender? _____ (2)
3. What is your ethnic origin? _____ (3)
 1. White
 2. Black, African American
 3. Hispanic
 4. Asian or Pacific Islander
 5. Native American
 6. Multiracial
4. What is your **current** marital status? _____ (4)
 1. Single
 2. Married
 3. Divorced
 4. Widowed
5. Do you have a dependent spouse or parent(s) in your care? _____ (5)
 1. No
 2. Yes
6. Do you have dependent children in your care? _____ (6)
 1. No (*Skip to Question 8.*)
 2. Yes (*Next question.*)
7. If you have dependent children in your care, how many do you have in each age category?

1. Under 1 year _____ (7)	4. 6-11 years _____ (10)
2. 1 - 2 years _____ (8)	5. 12-17 years _____ (11)
3. 3 - 5 years _____ (9)	6. 18 + years _____ (12)
8. Which one of the following best describes your **current employment** status? _____ (13)

(The \leq sign means equal to or less than; \geq equal to or more than.)

 1. Employed full-time, attending school part-time (≤ 11 hours)
 2. Employed part-time, attending school part-time (≤ 11 hours)
 3. Employed full-time, attending school full-time (≥ 12 hours)
 4. Employed part-time, attending school full-time (≥ 12 hours)
 5. Unemployed, attending school part-time, (≤ 11 hours) (*Skip to Question 12.*)
 6. Unemployed, attending school full-time (≥ 12 hours) (*Skip to Question 12.*)
9. What type of job do you **currently** have (e.g., cook, clerk, accountant, etc.)? _____ (14)
10. Approximately, how many hours per week do you currently work at a job for which you are paid or volunteer? _____ (15)

11. Which one of the following best describes your primary place of employment, paid or volunteer? _____ (16)
- | | |
|---|---|
| 1. Do not work or volunteer | 6. Food Service |
| 2. Hospitality (e.g., hotel) | 7. Educational Institution |
| 3. Health care (e.g., hospital) | 8. Retail or Other Sales |
| 4. Agriculture or Manufacturing | 9. Business Services (e.g., banking or insurance) |
| 5. Personal Services (e.g., house cleaning) | 10. Other (Specify: _____) |
12. Does your current work or personal circumstances require you to take courses in the evening or on weekends? _____ (17)
- | | |
|-------|--------|
| 1. No | 2. Yes |
|-------|--------|
13. Did you transfer academic credits or hours from another college or university to UGA? _____ (18)
- | | |
|-------|--------|
| 1. No | 2. Yes |
|-------|--------|
14. What is your **current** college classification? _____ (19)
- | | |
|----------------------------|-------------------------|
| 1. Freshman (0-45 hours) | 4. Senior (135 + hours) |
| 2. Sophomore (46-89 hours) | 5. Irregular/Transient |
| 3. Junior (90-134 hours) | 6. Don't Know |
15. What is your current UGA grade point average (GPA)? _____ (If not sure, give best estimate.) (20)
- | | |
|--------------|--------------------|
| 1. 4.0 | 5. 2.00-2.49 |
| 2. 3.50-3.99 | 6. 1.50-1.99 |
| 3. 3.00-3.49 | 7. 1.00-1.50 |
| 4. 2.50-2.99 | 8. Not Established |
16. On average, about how many miles do you travel, round trip, to attend this class? _____ (21)
17. On average, how many courses do you take each quarter? _____ (22)
18. On average, how many total clock hours each week do you study the class(es) you are taking this quarter? _____ (23)

Academic Credit Participation Index

Part II Directions: Please read each question carefully and select only one (1) answer for each question, using the following scale:

- | | |
|---|---|
| Very Strongly Disagree (VSD), circle 1 | Agree (A), circle 4 |
| Strongly Disagree (SD), circle 2 | Strongly Agree (SA), circle 5 |
| Disagree (D), circle 3 | Very Strongly Agree (VSA), circle 6. |

If you feel that an item is not applicable, please circle either 1, 2, or 3 depending on the degree of non-applicability.

A. I would rate myself in the top quarter of my classes (currently or when enrolled) for academic credit in:

	VSD	SD	D	A	SA	VSA	
1. Completing reading assignments	1	2	3	4	5	6	(24)
2. Completing writing assignments	1	2	3	4	5	6	(25)
3. Participating in discussions	1	2	3	4	5	6	(26)
4. Earning good grades	1	2	3	4	5	6	(27)
5. Working jointly on projects	1	2	3	4	5	6	(28)
6. Conducting library research	1	2	3	4	5	6	(29)
7. Making a class presentation	1	2	3	4	5	6	(30)
8. Participating in a group presentation	1	2	3	4	5	6	(31)
9. Taking essay tests	1	2	3	4	5	6	(32)
10. Taking multiple choice or similar tests	1	2	3	4	5	6	(33)

B. I was (or am) an above average learner in:

	VSD	SD	D	A	SA	VSA	
1. Primary school (e.g., grades 1 - 5)	1	2	3	4	5	6	(34)
2. Middle school (e.g., grades 6 - 8)	1	2	3	4	5	6	(35)
3. High School	1	2	3	4	5	6	(36)
4. Any job training at work	1	2	3	4	5	6	(37)
5. Learning on my own (e.g., reading a book)	1	2	3	4	5	6	(38)
6. Professional conferences	1	2	3	4	5	6	(39)
7. Non-academic credit courses or classes	1	2	3	4	5	6	(40)
8. Academic credit courses or classes	1	2	3	4	5	6	(41)
9. Correspondence study for academic credit	1	2	3	4	5	6	(42)
10. All other learning activities in which I engage	1	2	3	4	5	6	(43)

C. I engage in academic credit learning activities (e.g., classes or correspondence study) for the following reasons:

	VSD	SD	D	A	SA	VSA	
1. To improve my job performance	1	2	3	4	5	6	(44)
2. To prepare for a career	1	2	3	4	5	6	(45)
3. To advance my career (e.g., degree = promotion)	1	2	3	4	5	6	(46)
4. To increase earnings ability	1	2	3	4	5	6	(47)
5. To increase career options	1	2	3	4	5	6	(48)
6. To achieve academic goals	1	2	3	4	5	6	(49)
7. To achieve personal goals and/or satisfaction	1	2	3	4	5	6	(50)
8. To improve status at home	1	2	3	4	5	6	(51)
9. To improve status at work	1	2	3	4	5	6	(52)

D. The following events or circumstances have (or recently have had) prevented my participation in academic credit learning activities (e.g., classes or correspondence study):

	VSD	SD	D	A	SA	VSA	
1. Starting a new job	1	2	3	4	5	6	(53)
2. Advancing in a job	1	2	3	4	5	6	(54)
3. Losing a job (involuntary)	1	2	3	4	5	6	(55)
4. Starting in or changing my occupation	1	2	3	4	5	6	(56)
5. Starting a close personal relationship	1	2	3	4	5	6	(57)
6. Ending a close personal relationship	1	2	3	4	5	6	(58)
7. Community volunteer involvement	1	2	3	4	5	6	(59)
8. Personal health concerns or changes	1	2	3	4	5	6	(60)
9. <u>Current</u> parenting or care giving responsibilities	1	2	3	4	5	6	(61)
10. <u>Change</u> in current parenting or care giving responsibilities	1	2	3	4	5	6	(62)
11. Starting or continuing a hobby	1	2	3	4	5	6	(63)

E. The following events or circumstances have prevented me from participating or continuing to participate, as I would like, in academic credit learning activities (e.g., classes or correspondence study):

	VSD	SD	D	A	SA	VSA	
1. Home responsibilities	1	2	3	4	5	6	(64)
2. Job responsibilities	1	2	3	4	5	6	(65)
3. Lack of a place to study	1	2	3	4	5	6	(66)
4. Lack of time to study	1	2	3	4	5	6	(67)
5. Opinions of family or friends	1	2	3	4	5	6	(68)
6. Lack of transportation	1	2	3	4	5	6	(69)
7. Time and location of courses	1	2	3	4	5	6	(70)
8. Amount of time to complete degree, other requirements, or my goals	1	2	3	4	5	6	(71)
9. Lack of access to academic support services	1	2	3	4	5	6	(72)
10. Feeling too old or too young to learn	1	2	3	4	5	6	(73)
11. Lack of confidence in doing well on assignments or tests	1	2	3	4	5	6	(74)
12. Costs (e.g., tuition, books, transportation, etc.)	1	2	3	4	5	6	(75)
13. Too tired for learning, given other responsibilities	1	2	3	4	5	6	(76)

F. Each of the following are important sources of information about learning opportunities related to my academic credit goals:

	VSD	SD	D	A	SA	VSA	
1. Newspaper stories or ads	1	2	3	4	5	6	(77)
2. Radio stories or ads	1	2	3	4	5	6	(78)
3. Television stories or ads	1	2	3	4	5	6	(79)
4. Friends or family members	1	2	3	4	5	6	(80)
5. Other students	1	2	3	4	5	6	(81)
6. Co-workers or supervisor(s)	1	2	3	4	5	6	(82)
7. Posters or pamphlets	1	2	3	4	5	6	(83)
8. Course schedule(s)	1	2	3	4	5	6	(84)
9. Newsletters or other mailings	1	2	3	4	5	6	(85)
10. Volunteer activities (e.g., church, YMCA, etc.)	1	2	3	4	5	6	(86)

G. Below are listed several types of learning activities. How frequently do you engage in each of these, using this scale?

Never (N), circle 1
 Rarely (R), circle 2
 Occasionally (O), circle 3

Often (OF), circle 4
 Frequently (F), circle 5
 Very Frequently (VF), circle 6

	N	R	O	OF	F	VF	
1. Reading books to gain knowledge (Not textbooks)	1	2	3	4	5	6	(87)
2. Watching TV to gain knowledge (e.g., TV courses, documentaries, etc.)	1	2	3	4	5	6	(88)
3. Listening to the radio to gain knowledge	1	2	3	3	5	6	(89)

4. Taking a course for academic credit	1	2	3	4	5	6	(90)
5. Taking a non-credit course	1	2	3	4	5	6	(91)
6. Using a computer network	1	2	3	4	5	6	(92)
7. Conducting library research, but not for a class	1	2	3	4	5	6	(93)
8. Attending professional conferences	1	2	3	4	5	6	(94)
9. Attending a conference not related to my professional or academic goals	1	2	3	4	5	6	(95)
10. Consulting an expert	1	2	3	4	5	6	(96)
11. Reading pamphlets, reports, etc.	1	2	3	4	5	6	(97)
12. Correspondence study	1	2	3	4	5	6	(98)

EVCL Participation Profile & ACPI

Survey Codebook

<u>Column</u>	<u>Variable</u>	<u>Description</u>	<u>Range</u>	<u>Item</u>
1	V00001	Age	18-80	1
2	V00002	Gender	1=Male, 2=Female	2
3	V00003	Ethnic Origin	1=White, 4=Asian 2=Black, 5=Nat. Amer. 3=Hisp., 6=Multiracial	3
4	V00004	Marital Status	1=Single, 3=Divorced 2=Married, 4=Widowed	4
5	V00005	Dependent Spouse	1=No, 2=Yes	5
6	V00006	Kids	1=No, 2=Yes	6
7	V00007	<1 Years	0-10	7
8	V00009	3-5	0-10	
10	V00010	6-11	0-10	
11	V00011	12-17	0-10	
12	V00012	18+	0-10	
13	V00013	Employment Status	1-6(Appendix A)	8
14	V00014	Job Type	(Appendix A)	9
15	V00015	# Hours	0-60	10
16	V00016	Employ Place	1-10(Appendix A)	11
17	V00017	Non-Traditional	1=No, 2=Yes	12
18	V00018	Transfer Students	1=No, 2=Yes	13
19	V00019	Classification	1-6(Appendix A)	14
20	V00020	UGA GPA	1-8(Appendix A)	15
21	V00021	Miles Travel	1-999	16
22	V00022	Course Load	1-5	17
23	V00023	Hours Study	000-160	18
24	V00024	Readings	1-6	A1
25	V00025	Writing	1-6	A2
26	V00026	Discussions	1-6	A3
27	V00027	Good Grades	1-6	A4
28	V00028	Projects	1-6	A5

<u>Column</u>	<u>Variable</u>	<u>Description</u>	<u>Range</u>	<u>Item</u>
29	V00029	Library	1-6	A6
30	V00030	Presentation	1-6	A7
31	V00031	Group Presentation	1-6	A8
32	V00032	Essay Tests	1-6	A9
33	V00033	Objective Tests	1-6	A10
34	V00034	Primary School	1-6	B1
35	V00035	Middle Sch.	1-6	B2
36	V00036	High Sch.	1-6	B3
37	V00037	Job	1-6	B4
38	V00038	Own Learning	1-6	B5
39	V00039	Conferences	1-6	B6
40	V00040	Non-Credit	1-6	B7
41	V00041	Academic Credit	1-6	B8
42	V00042	Correspondence	1-6	B9
43	V00043	All Other	1-6	B10
44	V00044	Job Performance	1-6	C1
45	V00045	Prepare Career	1-6	C2
46	V00046	Advance Career	1-6	C3
47	V00047	Increase & Ability	1-6	C4
48	V00048	Increase Career Options	1-6	C5
49	V00049	Do Academic Goals	1-6	C6
50	V00050	Do Personal Goals	1-6	C7
51	V00051	Home Status	1-6	C8
52	V00052	Work Status	1-6	C9
53	V00053	Starting Job	1-6	D1
54	V00054	Advancing in Job	1-6	D2
55	V00055	Losing Job	1-6	D3
56	V00056	Occupation Change	1-6	D4
57	V00057	Starting Relationship	1-6	D5
58	V00058	Ending Relationship	1-6	D6
59	V00059	Volunteering	1-6	D7
60	V00060	Health Issues	1-6	D8
61	V00061	Current Parenting	1-6	D9
62	V00062	Change Parenting	1-6	D10
63	V00063	Hobby	1-6	D11
64	V00064	Home Resp.	1-6	E1
65	V00065	Job Resp.	1-6	E2
66	V00066	No Study Place	1-6	E3
67	V00067	No Study Time	1-6	E4
68	V00068	Opinions	1-6	E5
69	V00069	No Transportation	1-6	E6
70	V00070	Course Time	1-6	E7
71	V00071	Time to Complete	1-6	E8
72	V00072	Lack Access to Support Services	1-6	E9
73	V00073	Too Old Learn	1-6	E10
74	V00074	No Self-Confidence	1-6	E11
75	V00075	Costs	1-6	E12
76	V00076	Too Tired	1-6	E13
77	V00077	Newspaper	1-6	F1

<u>Column</u>	<u>Variable</u>	<u>Description</u>	<u>Range</u>	<u>Item</u>
78	V00078	Radio	1-6	F2
79	V00079	TV	1-6	F3
80	V00080	Friends	1-6	F4
81	V00081	Other Students	1-6	F5
82	V00082	Co-Workers	1-6	F6
83	V00083	Posters	1-6	F7
84	V00084	Course Schedule	1-6	F8
85	V00085	Newsletters	1-6	F9
86	V00086	Volunteering	1-6	F10
87	V00087	Reading Books	1-6	G1
88	V00088	Watching TV	1-6	G2
89	V00089	Listening to Radio	1-6	G3
90	V00090	Taking AC Course	1-6	G4
91	V00091	Noncredit Course	1-6	G5
92	V00092	Computers	1-6	G6
93	V00093	Library Research	1-6	G7
94	V00094	Professional Conference	1-6	G8
95	V00095	Unrelated Professional	1-6	G9
96	V00096	Expert	1-6	G10
97	V00097	Reading Pamphlets	1-6	G11
98	V00098	Correspondence Study	1-6	G12
99	Subtest A	Score	0-60	
100	Subtest B	Score	0-60	
101	Subtest C	Score	0-54	
102	Subtest D	Score	0-66	
103	Subtest E	Score	0-78	
104	Subtest F	Score	0-60	
105	Subtest G	Score	0-72	
106	Participation	Level	1=low; 2=medium; 3=high	

Supplemental Codes

Item 2: Gender Male 1 Female 2

Item 9: Job Codes

Accounting Services	1	Lifeguard	21
Bank Teller	2	Manager, Office	22
Bartender	3	Manager, Other	23
Bus Driver	4	Manager, Restaurant	24
Cashier	5	Manager, Sales	25
Child Care Worker	6	Newspaper Worker	26
Clerk (General)	7	Painter	27
Construction Worker	8	Physical Plant Worker	28
Cook	9	Police Services	29
Customer Servs. Rep.	10	Radio Worker	30
Delivery Driver	11	Sales Rep.	31
Dental Servs. Worker	12	Secretary	32
Executive Assistant	13	Social Services Worker	33
Gofer	14	Teacher	34
Health Services Worker	15	Telephone Services Worker	35
Homemaker	16	Test Servs. Worker	36

Instructor	17	Trucking Svs. Worker	37
Intern	18	Waiter, Waitress	38
Landscape Svs. Worker	19	Waiter, Waitress, & Other Role	39
Library Svs. Worker	20	Other	40
		Animal Svs. Worker	41
<u>Item 11, Employment Codes</u>		Counselor, Personal	42
Auto Repair/Svs.	11	Technician	43
Communications	12	Information Svs	44
Construction	13	Sports Worker	45
Government Svs.	14	Musician	46
Civic Org.	15	Mechanic	47
Animal Health Svs.	16	Missing	99
Leisure Svs.	17		
Legal Svs.	18		

Appendix 4.8: The Attitude Taxonomy

Attitudes influence learning, as attitude predisposes the mind. Students, who are positive about learning, like school; feel safe, physically and emotionally; and when nurtured, learn more readily and deeply than those who don't. Instructors, who are positive about their subject matter, believe students can learn, and who exhibit reasonable student affirming behavior, are the most effective knowledge transmitters and learning facilitators. Assessing attitudes and then constructively acting on that data, is critical to creating a productive learning environment.

Krathwohl, Bloom, and Masia (1964, pp. 35-37) advanced the Affective Domain Taxonomy, in a similar manner as Bloom's (1956) intellectual skill taxonomy. The Krathwohl, et al. taxonomy serves as our application framework in this chapter. First, we will examine Kerlinger's (1986) definitions of attitudes and traits. Following that, the Affective Domain Taxonomy is reviewed. The measurement item formats presented in Chapter 4 are suitable for measuring any attitude on the taxonomy.

A. Attitudes and Traits

1. Measuring an attitude or attitudes is a difficult task, as any attitude is a construct which may be unidimensional or multidimensional. The existence of an attitude is inferred by a person's words and behaviors.
 - a. Kerlinger (1986, p. 453) defines an attitude as, "an organized predisposition to think, feel, perceive, and behave toward a referent or cognitive object. It is an enduring structure of beliefs that predisposes the individual to behave selectively toward attitude referents. A referent is a category [political party], class [social or academic], or set of phenomena: physical objects [tall buildings], events [Martin Luther King Day], behaviors [smoking], or even constructs [patriotism]."
 - b. Kerlinger (1986, p. 453) differentiates between an attitude and a trait. He defines a trait to be, "a relatively enduring characteristic of the individual to respond in a certain manner in all situations." A person who dislikes math is likely to only dislike math (an attitude). A person who dislikes learning, most probably dislikes all academic subjects (a trait). Personality measurement is almost always trait measurement and is very clinical in nature. Trait measurement is beyond the scope of the present discussion.
2. By measuring attitudes, which are indicators of motives and intentions, we gain insight (e.g., explanation and prediction) as to possible or probable individual or group behavior.
3. Krathwohl, Bloom, and Masia (1964, pp. 35-37) have advanced a five level taxonomy for framing attitudinal objectives, which is also useful in guiding educational, social, managerial and marketing attitudinal research measurement decisions.
 - a. We'll first examine an overview of the taxonomy and then approach it from an alternative (and probably simpler) perspective.
 - b. A detailed explanation will then follow with recommendations for measurement.

B. The Affective (i.e., Attitude) Domain Taxonomy: An Overview**1. The Attitudinal Taxonomy**

- a. Receiving (or attending) is the first level of the taxonomy which relates to the examinee's sensitivity to affective stimuli or phenomena, i.e., his or her willingness to attend to each. There are three gradients to receiving or attending: awareness, willingness to receive, and controlled or selected attention.
 - b. Responding involves a response which passes merely attending to a stimuli or phenomena so that the examinee or respondent is at least in a very small way is committing himself or herself to the stimuli or phenomena. Such commitment must not be confused with a value or attitude. There are three levels of responding (acquiescence in responding, willingness to respond and satisfaction in response) each of which is characterized by increasing degree of internalization and voluntary action.
 - c. Valuing refers to the fact that an object, behavior, or phenomenon has worth. Each of the three levels of valuing within the taxonomy represents a deeper degree of internalization. Behavior is consistent and stable so as to indicate the possession of an attitude or value. At the first level (acceptance of a value), the subject merely accepts a belief whereas the highest level of valuing (commitment or conviction) may be described as belief with little if any doubt. Valuing is not a compliance behavior, but the result of an underlying commitment which guides behavior. Most attitudinal assessment starts at this level.
 - d. Organization refers to the building of a value system. Intended research, educational, training, or management, outcomes which require the formation of a value system are classified here. Assessment at this level measures the two dimensions of organization: conceptualization of value and organization of a value system.
 - e. Characterization means that at this "level of internalization the values already have a place in the individual's value hierarchy, are organized into some kind of internally consistent system, have controlled the behavior of the individual for a sufficient time that he [or she] has adapted to behaving this way" Krathwohl, Bloom, and Masia (1964, p. 165). The two components are "generalized set" and "characterization."
2. An alternative View of the Taxonomy
 - a. Interest comprises taxonomy levels receiving, responding, and two levels of valuing: "acceptance of a value" and "preference for a value".
 - b. Appreciation extends from receiving's "controlled or selected attention", across all three levels of responding to two levels of valuing: "acceptance of a value" and "preference for a value."
 - c. Attitudes encompass "willingness to respond", all levels of valuing, and the first level, "conceptualization of a value" within organization. Attitudes reflect a person's preference or feelings about a particular issue or phenomena; whereas opinions are verbally articulated attitudes.

- d. Value encompasses the same elements of the taxonomy as attitudes.
- e. Adjustment ranges from responding's "willingness to respond", across valuing, organization, and characterization by a value complex."

C. The Affective Domain: A Detailed Description

1. Receiving: This first level of the taxonomy relates to the examinee's sensitivity to affective stimuli or phenomena, i.e., his or her willingness to attend to each. There are three gradients to receiving or attending.
 - a. Awareness involves being conscious of the existence of stimuli (e.g., statement) or phenomena (e.g., behavior).
 - (1) Awareness is almost a cognitive behavior. It is difficult to write educational, training, research, or marketing objectives for this level.
 - (2) The purpose in measuring at this level is to determine whether or not the examinee or respondent is aware or conscious of a stimulus, phenomena, person, or event. The chief assessment problem is that awareness emerges without prompting from the examiner or researcher; so, the testing or research environment does not direct the examinee or respondent to the stimulus or phenomena.
 - (3) Measurement Strategies employed are:
 - (a) Items where the respondent (e.g., young children) sorts or matches generally based on awareness criteria, such as color, shape, design, etc. to test for awareness of color, shape, design, form, etc.
 - (b) Ranking response options by degree of desirability when given a description of student or employee behavior tests awareness.
 - (c) Matching and true false items designed at the knowledge level can be used to test awareness, provided such items are written towards that purpose.
 - b. Willingness to receive shows a willingness to tolerate a given stimulus or phenomena and not avoid either the stimulus or phenomena.
 - (1) The examinee suspends judgment regarding the stimulus or phenomena.
 - (2) In measuring willingness to receive, we seek to determine whether or not there is an absence of rejection. Three alternative item response formats are recommended; see immediately below. Strong positive affective options, such as those found in the Likert scale, e.g., strongly agree or strongly disagree are avoided. Item stems should be very tentative and present a rather general disposition towards a preference, intention, or behavior. In measuring attitudes at this level, we are only looking for a "favorably disposed" response.
 - (3) Sample response options
 - (a) Like—Indifferent—Dislike
 - (b) Agree—No Opinion—Disagree
 - (c) Yes—Uncertain—No
 - (d) Interesting—No Opinion—Uninteresting
 - (e) Certain—Not Sure—Uncertain
 - (f) Like—Neither Like or Dislike—Dislike
 - (g) Usually—Occasionally—Rarely

- (4) Measurement Strategies typically employed are:
- (a) Asking students whether they are willing to consider a curriculum topic or if they are indifferent to said consideration.
 - (b) Providing descriptions of preferences, intentions, or behaviors worded so that each can be answered in either a positive or neutral manner. If “no” is a response option, word the statement so that “no” is at least a neutral response. It is better that “no” be a positive response.
- c. Controlled or selected attention is effected when an examinee or respondent moves beyond simple awareness and willingness to attend to a stimulus or phenomena.
- (1) At this level, he or she pays attention to a selected stimulus or phenomena despite distractions or other competing stimuli.
 - (2) In measuring controlled or selected attention, our focus is assessing the strength of the awareness of the attitude.
 - (3) Measurement Strategies
 - a. The interest inventory is a widely used strategy. Stimuli in the form of statements are prepared. Respondents typically select from one of three response options. A respondent could choose either “Yes—Uncertain—No” in responding to the statement, “I have a strong preference for writing.”
 - b. The “forced choice” item format is also employed. The examiner prepares several activities or tasks in pairs. The examinee then selects his or her preference from among the pairs presented. Over time a pattern of preferences emerges. This pattern is evidence of controlled or selected attention. See the paired comparison scale example below.
2. Responding involves a response which passes merely attending to a stimuli or phenomena so that the examinee or respondent is, at least in a very small way committing himself or herself to the stimuli or phenomena. Such commitment must not be confused with a value or attitude. There are three levels of responding each of which are characterized by increasing degree of internalization and voluntary action.
- a. Acquiescence in responding occurs when the subject has agreed to comply with the stimulus or phenomena.
 - (1) It is quite possible that if there were other alternatives and no compliance pressure, the subject might elect an alternative response. Compliance with health and safety requests or regulations is the primary example. There are few educational or training objectives targeted to this response level. Management objectives at this level are common.
 - (2) The purpose of measurement at this level is to assess the subject’s degree of acquiescence and compliance to a stimulus and exhibition of expected behavior such as completing required homework or obeying traffic laws. The key question is the subject responding, i.e. is he or she turning in required homework as prescribed or actually obeying traffic laws.

- (3) Measurement Strategies
- (a) Direct observation is the preferred assessment method; but a survey of those who would know whether or not the subject is responding at this level is acceptable.
 - (b) Activities, checklists, or inventories are employed. Be sure to frame item stems within the experience of the examinee, respondent, or subject. Response options might be “I perform the activity, without being told or reminded.”, “I perform the activity only when told or reminded.”, or “I do not perform the activity.” Only the second response option demonstrates “acquiescence in responding.”
- b. Willingness to respond is when the subject, examinee, or respondent voluntarily consents, agrees to respond to a stimulus (e.g., item) or phenomena. Willingness to respond is demonstrated by the response behavior, i.e., the act of responding.
- (1) The element of resistance or compulsion that is found in “acquiescence in responding” is absent here. There is an absence of overt or subtle compliance pressure. Many educational, training, and management objectives are written to this level.
 - (2) The examiner’s principle interest is to determine the subject’s willingness to respond to a stimulus or phenomena. The reasons for a willingness to respond are not routinely assessed at this level.
- (3) Measurement Strategies
- (a) Direct observation of behavior is preferred. Behaviors might include a hobby, display of interest, or co-operative deportment. Inferences can be drawn from behaviors such as work product turned in on deadline, well constructed and presented, and which appears to be “above and beyond” usual expectations. Don’t rely on one cue when making this or any inference. A “package” of behaviors is needed.
 - (b) Activities, checklists, or inventories are employed. Be sure to frame item stems within the experience of the examinee, respondent, or subject. Response options might be “I perform the activity, without being told or reminded.”, “I perform the activity only when told or reminded.”, or “I do not perform the activity.” Only the first response option demonstrates “willingness to respond.”
- c. Satisfaction in response accompanies the responding behavior as demonstrated by the willingness to respond. Satisfaction in response is an emotional response characterized by a feeling of satisfaction. The feeling of satisfaction is reinforcement which tends to produce other responses.
- (1) At this level, the measurement interest is the emotional state which accompanies the response behavior. The display of emotion may be overt or covert.
 - (2) Measurement Strategies
 - (a) The testing for overt emotional display involves the determination as to which behaviors indicate satisfaction and then the development of a

measurement strategy, e.g., direct observation of behavior which indicates satisfaction, verbalizations, etc.

- (1) Patrons applauded loudly at the opera.
 - (2) Customer service representatives (CSR) expressed appreciation for the new employee recognition program.
- (b) To test for covert or private displays of satisfaction a scenario must be created and the respondent's reactions are documented in a systematic manner which might include an objective technique (e.g., Likert Scale, check list, etc.), a free response (e.g., an open-ended item where the subject is asked to write his or her response), or an adjective check list which contains both positive and negative selection options.
- (1) Indicate your degree of satisfaction with these specific topics within the CSR training program by circling the letters which represent your level of satisfaction which include SD, D, N, A, or SA.
 - (2) In the space provided, please describe your feelings about the field trip to the museum.
 - (3) Please circle those adjectives which best describe your reaction to the video just presented.
3. Valuing refers to the fact that an object, behavior, or phenomenon has worth. Each of the three levels of valuing within the taxonomy represents a deeper degree of internalization. Behavior is consistent and stable so as to indicate the possession of an attitude or value. At the first level, the subject merely accepts a belief whereas the highest level of valuing may be described as belief with little if any doubt. Valuing is not a compliance behavior, but the result of an underlying commitment which guides behavior. Most attitudinal assessment starts at this level.
- a. Acceptance of a value is characterized by behavior that is so consistently displayed that others say he or she must hold that value. The value is sufficiently deeply rooted within an individual so as to exert a controlling influence on behavior. However, at this level one is more likely to change his or her mind with respect to the value under observation than at the other higher valuing levels.
- (1) Examples include:
 - (a) He is said to appreciate diversity in entertainment, food, and friends as he is seen at the theater and local restaurants with different people periodically.
 - (b) She seems to desire to further refine her writing and speaking skills as she visits the writing and public speaking resource centers at least once a week.

Central to both examples is that the subject exhibits behavior which indicates the holding of an underlying value which informs behavior.
 - (2) Measurement Strategies
 - (a) Measurement strategies include direct observation and/or questioning, standard attitudinal indexes or scales, or verbalizations by the subject indicating that the value is held.
 - (b) At this level in the taxonomy, measurement is concerned with whether or not the value is accepted and the degree, to which it is valued, not rejected.

- b. Preference for a value indicates that the individual not only accepts the value, but is motivated to pursue it or wants to be perceived as holding that value, but falls short of a “full commitment” to the value. For example, a person might believe very strongly in a religious faith such that he or she practices that faith devotedly, but not so strongly as to become a member of the faith’s clergy.
- (1) Examples include:
- (a) He appreciates diversity in entertainment, food, and friends as he goes to plays, movies and operas with several different friends and visits a wide variety of ethnic restaurants.
 - (b) She desires to further refine her writing and speaking skills as she visits the writing and public speaking resource centers three times each week. Central to both examples is that the subject exhibits behavior which indicates the holding of an underlying value which informs behavior with increasing investments in time, energy, and resources.
- (2) Measurement Strategies
- (a) The preferred measurement strategy is the situational method where the respondent is given a variety of response choices. There are several situations presented, with the same response options. The pattern of choices is then analyzed. Consistency in choice is the criterion to determine whether or not the respondent has a preference for the value under investigation.
 - (b) As above, measurement is concerned with whether or not the value is accepted and the degree, to which it is valued, not rejected.
- c. Commitment is characterized by a degree of certainty which is beyond the “shadow of a doubt.” Commitment may be seen as religious faith, dedication to a political cause, or loyalty to a group. There is significant motivation to engage in behavior which displays the underlying value; the person holding the value seeks to deepen his or her understanding, involvement with, and further display his or her commitment to the value either by convincing others to share the belief or converting others to the belief.
- (1) Commitment Characteristics
- a. The commitment to the value or valuing of an object or phenomena encompasses a significant time period so that any measurement strategy will consider how long the value has been held and how likely it is to continue to be held, i.e., its stability.
 - b. Considerable energy and other investment in the value must be evidenced.
 - c. There should be sustained behaviors which by their very nature convey commitment to the value.
 - d. There is an evidenced strong emotional attachment to the value and a demonstrated willingness to display that attachment.
- (2) Measurement Strategies
- (a) Where possible direct observation is preferred, but self-report is a common method for gathering commitment data. For those subjects lacking a venue to display their commitment to a value or set of values, a

scenario may need to be constructed. Evidence of commitment is the degree of emotion displayed along with the intellectual quality of position statements respecting the scenario.

- (b) High scores on an attitudinal scale or index are generally considered as a preference for a value. Commitment is assessed via very detailed questionnaires or interviews which typically explore the value in much greater depth and breadth than a scale or index.
4. Organization refers to the building of a value system. Intended research, educational, training, or management, and outcomes which require the formation of a value system are classified here. Assessment at this level measures the two dimensions of value system organization: conceptualization of a value and organization of a value system. The measurement of value systems is beyond the scope of the present work. The interested reader is invited to consult Krathwohl, Bloom, and Masia (1964) or any standard text in educational psychology, marketing research, or psychology.
 5. Characterization at this “level of internalization the values already have a place in the individual’s value hierarchy, are organized into some kind of internally consistent system, have controlled the behavior of the individual for a sufficient time that he [or she] has adapted to behaving this way” Krathwohl, Bloom, and Masia (1964, p. 165). The two components are “generalized set” and “characterization”. The values must be generally set in place before the value system can be characterized. The measurement of value systems is beyond the scope of the present work. The interested reader is invited to consult Krathwohl, Bloom, and Masia (1964) or any standard text in educational psychology, marketing research, or psychology.

References

- Cooper, D. R. & Schindler, P.S. (2001). *Business research methods* (8th ed.) Boston, MA: Irwin McGraw-Hill.
- Bloom, B. S., Engelhart, M. D., Frost, E. J., & Krathwohl, D. (1956). *Taxonomy of educational objectives. Book 1 Cognitive domain*. New York, NY: Longman.
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.) New York, NY: Holt, Rinehart, and Winston.
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Book 2: Affective domain*. New York, NY: Longman.

Appendix 4.9: Scoring & Reporting Quantitative Items

Items for quantitative data collection tools (DCTs) or measures must be written so that a numerical score or data are produced. This may require that nominal or ordinal data be “transformed” into interval or ratio data as is done below. For information on the types of data, means, and standard deviations, see Chapter 9.

We examine scoring and reporting quantitative items at three levels: the item level, subtest level, and composite (also known as the total scale or test score). The use of scores is necessary for data analysis, reporting, interpretation, and decision-making. The use of scores permits more detailed statistical analysis of data.

A. Scoring & Reporting Item Level Data

1. Appendix 4.6 recommends several Likert Scale item response options; we’ll focus first on these. In these examples (Appendix 4.6, A1 to A4), an item mean, standard deviation, and percentage of endorsements (selection) for each response option is the usual reporting practice. Analyzing item level responses, gives the evaluator insight into an item-specific knowledge, skill, or attitude (depending on the item). This detailed insight informs subsequent data analysis, the drawing of conclusions, framing recommendations, and eventually decision-making. Four examples are presented below.
2. **Statement: A blue sky is beautiful.**
 - a. Let’s start with Response String D2 from Appendix 4.6 which measures agreement: Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree.
 - (1) Its common practice to shorten these response option labels from “Strongly Disagree” to “SD,” “Agree” to “A,” “Neutral” or “No Opinion” to “N,” “Disagree” to “D” and “Strongly Disagree” to “SD.”
 - (2) Each Likert Scale response option is numerically weighted, 1 = Strongly Disagree to 5 = Strongly Agree. This “converts” the ordinal categories to interval level data from which a mean and standard deviation can be computed.

SD	D	N	A	A
1	2	3	4	5

- b. An item mean (e.g., 4.1) and standard deviation (e.g., 0.96) can be computed; additionally, the percentage of each respondent selecting one of the five item response choices can be calculated. A high item mean score indicates more agreement with the statement. The percentage endorsing each response option would yield the same data.
- c. Let’s say this statement was presented to 500 people in a mall; it would be difficult to make sense out of 500 separate responses. We computed these descriptive statistics: Mean = 4.6; standard deviation 0.3 and the percentages below.

SD	D	N	A	A
5%	5%	5%	15%	75%

- (1) The Likert Scale ranges from 1 to 5; so, a mean of 4.6 indicates widespread agreement with the statement, given the measurement scale. A very low standard deviation suggests little variance or differences in opinion. The measurement scale guides and bounds the interpretation; we must interpret data within the confines of the measurement scale, in this case 1 to 5 or SD to SA.
- (2) The percentage spread across the 5 response options or measurement scale confirms the mean and standard deviation.

3. **Statement: I eat healthy food.**

- a. Next, let's examine Appendix 4.6 Response String E1 which measures frequency of behavior: Never, Rarely, Occasionally, Often, Frequently, Very Frequently.
 - (1) Each Likert Scale response option is numerically weighted, 1 = Never to 6 = Very Frequently.

Never	Rarely	Occasionally	Often	Frequently	Very Frequently
1	2	3	4	5	6

- b. As above, an item mean and standard deviation can be calculated along with the percentage endorsing (selecting) each response option. A higher item mean indicates that the behavior is engaged more frequently, but never more than "Very Frequently."
- c. Let's say this statement was presented to 500 people at an obesity clinic; it would be difficult to make sense out of 500 separate responses. We computed these descriptive statistics: Mean = 3.1; standard deviation 1.7 and the percentages below.

Never	Rarely	Occasionally	Often	Frequently	Very Frequently
5%	10%	50%	10%	15%	10%

- (1) The Likert Scale ranges from 1 to 6; so, a mean of 3.1 points indicates that respondents occasionally eat healthy food. A standard deviation of 1.7 points suggests some variance or differences in behavior among respondents.
- (2) The percentage spread across the 6 response options confirms the mean and standard deviation; the mutual confirmation between the percentage spread and the mean and standard deviation is a useful self-check.

4. **Statement: The manager is fair in making job assignments.**

- a. Next, let's examine Appendix 4.6 Response String B1 which measures performance: Exceeded Expectations (EE), Met Expectations (ME), or Didn't Meet Expectations (DME).
 - (1) Each Likert Scale response option is numerically weighted, 1 = DME to 3 = EE.

DME	ME	EE
1	2	3

- b. As above, an item mean and standard deviation can be calculated along with the percentage endorsing each response option. A higher item mean indicates higher

performance. Due to the length of the response options or measurement scale, we shortened them.

- c. Let's say this statement was presented in a manager's 360° performance appraisal to her 17 colleagues, direct reports, and supervisor; it would be difficult to make sense of 17 separate responses. We computed these descriptive statistics: Mean = 2.8; standard deviation 0.5 and the percentages below.

DME	ME	EE
6%	59%	35%

- (1) About 60% of the respondents reported thinking the manager was fair in making job assignments. Thirty-five percent agreed that the manger exceeded expectations, with one (1) person disagreeing. The mean of 2.8 on a three point scale, confirms the percentage spread. The small standard deviation suggests little variance in opinion.
- (2) When applying labels (e.g., "small," "really small," "medium," or "large" to a standard deviation, the researcher must draw on his or her professional experience, knowledge of the research literature, and the measurement scale being used. Each qualitative label must be reasonable and defensible.

5. Statement: I went fishing last Saturday.

- a. Finally, let's examine Appendix 4.6 Response String A3 which measures whether or not something happened or were agreed to: No or Yes. These are nominal data.

Statement A	No	Yes
Respondent 1		x
Respondent 2		x
Respondent 3		x
Respondent 4	x	
Respondent 5	x	
Respondent 6		x
Respondent 7	x	
Total	3	4

- b. Three (3) respondents answered Statement A "No" while four (4) answered "Yes." We can't compute an item mean or standard deviation; computing percentages isn't of much value, as the addition or deletion of a single respondent would change the percentage value substantially, which is common with small sample or study group sizes.
- c. When we have two response options with a small number of respondents, we usually report the raw number of endorsements (selections) for each response option (in this case "No" or "Yes"). Three respondents reported "No," whereas 4 said, "Yes."

6. Scoring a Checklist

- a. A checklist is simple data collection tool using binary responses, e.g., “No or Yes,” “Did or Didn’t,” “Included or Not Included,” or “Performed or Not Performed.”
- b. Checklists are usually used to measure or assess an examinee’s or respondent’s skill in following a procedure or to ensure everything is included in an application, emergency rescue kit, preparation for a parachute jump,
- c. Items are usually scored “0” or “1” or “1 or 2” or even with a simple check “√.” When numerical values (“0”/”1”) are used, a total score can be calculated. Item level scores are reported using percentages or just numbers.

Procedure A	Correct (1)	Incorrect (0)
Step 1	X	
Step 2	X	
Step 3	X	
Step 4	X	
Step 5	X	
Step 6		X
Step 7	X	
Total	6	0

The examinee performed 5 of the seven steps in Procedure A correctly, but didn’t perform Step 6 correctly. Depending on the procedure, the examinee would likely not “pass” as it’s usually necessary for all 7 steps in a procedure to be done correctly.

B. Scoring & Reporting Subtest Data

1. Look at Appendix 4.7; notice the Academic Credit Participation Index (ACPI). The ACPI is a multidimensional measure as it measures the seven (7) dimensions (A-G) of the Chain of Response Model (Cross, 1981). Each subtest measures one dimension or part of the theory (see Table 4.1). Let’s look at Subtest A which is composed of 10 items with 6 response options (Appendix 4.6 Response String E2) in more detail.

A. I would rate myself in the top quarter of my classes (currently or when enrolled) for academic credit in:

	VSD	SD	D	A	SA	VSA
1. Completing reading assignments	1	2	3	4	5	6
2. Completing writing assignments	1	2	3	4	5	6
3. Participating in discussions	1	2	3	4	5	6
4. Earning good grades	1	2	3	4	5	6
5. Working jointly on projects	1	2	3	4	5	6
6. Conducting library research	1	2	3	4	5	6
7. Making a class presentation	1	2	3	4	5	6
8. Participating in a group presentation	1	2	3	4	5	6
9. Taking essay tests	1	2	3	4	5	6
10. Taking multiple choice or similar tests	1	2	3	4	5	6

2. To obtain the whole-group perspective on “individual skills,” we’d compute, for each item, a mean, standard deviation, and percentage spread for each response option, including zeros (response options not selected by anybody). We want to know how the individual examinees (adult undergraduate degree completing students) viewed their collective strength on each specific academic skill.

3. Next, we want the whole-group perspective on “academic skills.” So, the individual item responses are summed to produce a single subtest score. Since, there are ten items with six response options each, there is a maximum of 60 points (10 x 6). Also, a subtest mean, standard deviation, and percentage spread across response options are computed.
 - (a) The closer an individual or group subtest score is to 60; the more confidence an individual has in his or her academic skill set; the same interpretation applies to the group. An adult student (or any learner) who is academically skilled, is more likely to persist to graduation. Weaker students need academic support or may opt out.
 - (b) An individual’s subtest test score can let an advisor, teacher, or program manager, identify a potentially weak student. A review of individual item means will identify exact weaknesses and guide strategies to strengthen specific academic skills.
 - (c) At the group level, a program manager will be alerted to the possible need for intervention, if a group of respondents scored 30 of 60 possible points. He or she would know that the group is generally weak on academic skills and/or that a significant segment of the group is weak. So, the program manager will investigate and intervene to strengthen specific academic skills by analyzing each separate item.

4. Often what we want to measure is one dimension; if we were only interested in measuring academic skill confidence, we’d just measure Subtest A and not be concerned about the other six (6) dimensions. What we want to know, determines what we measure.
 - a. Appendix 3.1 is a one dimensional measure. The scoring rules are more elaborate than Appendix 4.2 because its purpose was to inform graduate students as to how other students perceived their individual contribution to major group projects. Score interpretation guidance is provided as well.
 - b. Appendix 4.2 was designed to stimulate discussion in a supervisor skills training course. So, it’s scoring and interpretation is simpler and less formal.

C. Scoring & Reporting Composite or Total Scale (Test) Scores

1. Let’s return again to Appendix 4.7, recall that Academic Credit Participation Index (ACPI) is composed of seven (7) subtests (A-G). Each subtest measures a dimension of the Chain of Response Model (Cross, 1981). See Table 4.9.1.

Table 4.9.1

ACPI Model Dimension	Subtest	Total Points
(A) Self-evaluation	A	60
(B) Education Attitudes	B	60
(C) Goals & Expectations	C	54
(D) Life Transitions	D	66
(E) Opportunities & Barriers	E	78
(F) Information	F	60
(G) Participation	G	72
AAPI Total Score	A-G	450

2. To produce a total test score, we’d sum each of the seven subtest scores. Total points are 450. The closer an individual score is to 450, the more likely the adult student is to successfully participate in academic credit learning experiences to achieve his or her

academic goals according to the theory. For each subtest, a composite score mean, standard deviation, and percentage spread across the response option string is usually computed.

3. By way of practical application, let's suppose the program manager decided to administer the ACPI to each prospective adult student. This risk profile is created based on current research, Cross' theory, his/her professional judgment, and the program's unique history. We've chosen to gauge the degree of risk on the University's course grading system. In doing so, the following risk profile emerges (Table 4.9.2).

Table 4.9.2

Points	Risk for Non-Persistence	Computation
405-450	Very Low Risk	$450 * 0.90$
360-404	Low Risk	$450 * 0.80$
315-359	Average Risk	$450 * 0.70$
270-314	High Risk	$450 * 0.60$
< 270	Very High Risk	

- (a) The ACPI can be given to incoming students; first, an advisor or the program manager can review individual composite scores to identify individuals for further investigation, given their risk score. A student scores less than 270 would receive more intense academic support than a student scoring between 360-404 points. A student scoring more than 405 points will likely receive little or no academic support.
- (b) Second, the manager can focus on individual subtests or subtest items to identify specific areas of weakness. Third, a support plan can be written for each learner tailored to his or her specific needs.
- (b) Also, the ACPI scores of those who drop out or who are not otherwise successful can be reviewed to adjust risk category point definitions to be more accurate given the University's experience with its students.

Reference

Cross, K. P. (1981). *Adults as learners: Increasing participation and facilitating learning*. San Francisco, CA: Jossey-Bass.