

## CHAPTER 4: OBSERVING AND MEASURING COMMUNICATION VARIABLES

### I. Introduction

- A. If you meet an old friend after a long separation and ask, “how ya’ doing?” you’re likely to get the bland and disappointing answer, “Fine.”
  - 1. Vague questions most often yield uninformative responses.
  - 2. So how might you learn more about your friend’s life?
    - a. You would ask pointed questions about specific issues.
    - b. In so doing, you’re taking the overall goal, “learning about your friend’s life,” and dividing it into some of its main components.
- B. The process of variable identification and measurement in the work of communication researchers has a similar goal—learning about particular elements of communication in people’s lives.

### II. Conceptual Versus Operational Definitions

- A. When researchers start to operationalize an abstract term, they take word definition a step beyond our commonsense view of it.
  - 1. A **conceptual definition** describes what a concept means by relating it to other abstract concepts.
  - 2. An **operational definition** describes a concept in terms of its observable and measurable characteristics or behaviors, by specifying how the concept can be observed in actual practice.
    - a. As an example, notice that we’ve focused on communication-related indicators of love such as saying “I love you” and “staring into another’s eyes.”
    - b. Had we been psychologists, we might have operationalized love in terms of cognitive or emotional states, such as how much a person reports feeling close, physically attracted, and committed to another person.
    - c. It is theoretically possible to operationalize any abstract concept.
    - d. While simple concepts, like the volume or pitch of a voice, may be operationally defined by a single observable characteristic, no single characteristic stands very well for complex concepts.
    - e. An operational definition represents the most readily available way of specifying the observable characteristics of an abstract concept that we currently know, but operational definitions rarely capture completely all the dimensions of sophisticated concepts.
    - f. When many observable characteristics are included in an operational definition, researchers must decide which ones are more essential than the others for defining the abstract concept.
  - 3. The first step in research, then, is moving from the abstract, conceptual level to the concrete, operational level; and only after defining a concept operationally can researchers examine it.
- B. Evaluating Operational Definitions: Researchers use a conceptual definition as the basis for devising a good operational definition.
  - 1. Good conceptual definitions describe the primary elements of the research topic being investigated and researchers often refer back to them to assure that the behaviors they observe and measure in their studies actually reflect the conceptual components of their research topic.
  - 2. Researchers try to retain in the operational definition the essential meaning of a conceptual definition.
    - a. This strong linkage between a conceptual and an operational definition is referred to as **conceptual fit**.
    - b. The closer the conceptual fit, the more likely it is that researchers are observing the phenomenon they intend to study.
    - c. The looser the conceptual fit, the greater the danger that researchers are observing a phenomenon different from the one they intended to study.
    - d. Researchers may differ about how a concept should be operationally defined and measured.
  - 3. Barker (1989) suggests keeping the following questions in mind when evaluating researchers’ operational definitions:
    - a. Is the definition or operationalization *adequate*?
    - b. Is the definition or operationalization *accurate*?

c. Is the definition or operationalization *clear*?

### III. Measurement Theory

- A. After researchers specify the observable characteristics of the concepts being investigated, they must determine the ways to record and order in a systematic way observations of those behavioral characteristics.
1. **Measurement** is the process of determining the existence, characteristics, size, and/or quantity of changes in a variable through systematic recording and organization of the researcher's observations.
    - a. Increasingly precise measurement and statistical analysis of behavior are relatively recent developments, pioneered in the nineteenth century and refined in the twentieth century.
- B. Quantitative and Qualitative Measurements: One way measurements are distinguished is with respect to whether they employ meaningful numerical symbols.
1. **Quantitative measurements** employ meaningful numerical indicators to ascertain the relative amount of something.
  2. **Qualitative measurements** employ symbols (words, diagrams, and nonmeaningful numbers to indicate the meanings (other than relative amounts) people have of something.
  3. Quantitative and qualitative measurements provide researchers with different but potentially complementary ways of measuring operationally defined concepts.
    - a. *Quantitative measurements* provide numerical precision about such properties as amount and size.
    - b. Qualitative measurements provide useful information about people's perceptions.
- C. There has been much debate in communication research about whether quantitative or qualitative measurement is more persuasive.
1. Baeslar and Burgoon (1994) found that all forms of evidence were initially persuasive when compared to no evidence.
    - a. Statistical evidence was more persuasive than story evidence, and this held constant at 48 hours and even 1 week for the vivid statistical evidence.
    - b. The persuasiveness of statistical evidence is also supported by M. Allen and Priess' (1997) meta-analysis, which shows that across 15 investigation, statistical evidence was more persuasive than narrative evidence.
    - c. There are some studies that support the persuasiveness of other forms of qualitative evidence besides narratives.
    - d. One example is reported by Kazoleas (1993) who found that while both types of evidence were equally effective in changing attitudes, the qualitative evidence was significantly more persuasive over time, with people also recalling the qualitative evidence better than the quantitative evidence.
    - e. It could be that quantitative and qualitative measurements affect people in different ways.
- D. The debate between quantitative and qualitative measurements is by no means settled.
1. Many researchers and practitioners actually use both types of measurements to enhance both the precision of the data gathered and an understanding of contextual influences on those data.
  2. Studying something in multiple ways within a single study is called **triangulation**, which means calculating the distance to a point by looking at it from two other points.
  3. There are four types of triangulation.
    - a. *Methodological triangulation* involves the use of and comparisons made among multiple methods to study the same phenomenon.
    - b. *Data triangulation* is where a number of data sources are used.
    - c. *Research triangulation* is when multiple researchers are used to collect and analyze data.
    - d. *Theoretical triangulation* involves using multiple theories and/or perspectives to interpret the same data.
    - e. Communication researchers often use triangulation to increase their understanding of the phenomenon of interest, and one way to methodologically triangulate research findings is to

combine quantitative and qualitative measurements.

#### E. Levels of Measurement

1. In Chapter 2, we described a variable as an observable concept that can take on different values.
  - a. These different values are measured by using a **measurement scale**, “a specific scheme for assigning numbers or symbols to designate characteristics of a variable” (F. Williams, 1986, p. 14).
2. Stevens (1958) identified four levels of measurement scales used to describe the type, range, and relationships among the values a variable can take: nominal, ordinal, interval, and ratio.
  - a. These levels of measurement are arranged hierarchically—each level has all the characteristics of the preceding level, and each provides increasing measurement precision and information about a variable.
  - b. **Nominal** variables are differentiated on the basis of type or category: hence, **nominal measurement** scales classify a variable into different categories.
    - i. The term “nominal” is derived from the Latin *nomen*, meaning “name”, and the categories of a nominal scale may be named by words or by numbers.
  - c. A variable measured at the nominal level must be classifiable into at least two different categories.
    - i. The categories must be mutually exclusive; otherwise, comparisons between them are misleading.
    - ii. With regard to communication research, the problem of constructing mutually exclusive categories is illustrated in Christenson and Peterson’s (1988) study of college students’ perceptions of “mapping” music genres (see Figure 4.1).
    - iii. The categories must be equivalent; otherwise, we will be comparing apples and oranges.
    - iv. The categories must be exhaustive; otherwise, they will not represent the variable fully.
  - d. Many communication researchers measure variables, especially independent variables using nominal measurements (see Chapter 7).
    - i. Many **background variables** (also called **classification, individual-difference, organismic, or subject variables**), “aspects of subjects’ [research participants] ‘backgrounds’ that may influence other variables but will not be influenced by them” (Vogt, 1993, p. 16), such as one’s nationality and ethnicity are measured using nominal scales.
  - e. Researchers also sometimes measure dependent variables nominally.
    - i. The most basic way is asking people to answer “yes” or “no” to questions.
    - ii. Another nominal form of measurement is asking people to choose between two or more responses from a checklist.
    - iii. Many communication researchers measure variables, especially independent variables using nominal measurements.
    - iv. Researchers also sometimes measure a dependent variable nominally by first asking people to respond to open-ended questions and then classifying their responses into categories.
    - v. Counting the frequency of use of nominal categories may reveal important findings, but it also limits the quantitative data analyses that can be performed and the conclusions that can be drawn.
  - f. **Ordinal measurement scales** not only classify a variable into nominal categories but also rank order those categories along some dimension.
    - i. The categories can then be compared because they are measured along some “greater than” and “less than” scale.
    - ii. An ordinal scale in which a particular rank can only be used once is referred to as an **ipsative scale**; a scale that allows ranked ties is called a **normative scale**.
    - iii. Ordinal measurements provide more information about a variable than nominal measurements because they transform discrete classifications into *ordered* classifications.
    - iv. Ordinal measurements are still quite limited, for they only rank order a variable along a dimension without telling researchers *how much* more or less of the variable has been

measured.

- v. The points on an ordinal scale are, thus, arranged in a meaningful numerical order, but there is no assumption that the distances between the adjacent points on that scale are equivalent; a variable measured by an interval or ratio scale is sometimes called a *continuous, metric, or numerical variable*.
- g. **Interval measurement scales** not only categorize a variable and rank order it along some dimension but also establish *equal distances* between each of the adjacent points along the measurement scale.
  - i. Interval measurements also include an arbitrary zero point on the scale; however, a zero rating does not mean the variable doesn't exist.
  - ii. Interval measurements provide more information than nominal or ordinal measurements.
  - iii. There are a number of scales used in the social sciences for measuring variables at the interval level.
  - iv. **Likert scales** identify the extent of a person's beliefs, attitudes, or feelings toward some object.
    - (a) The traditional Likert scale asks people the extent to which they agree or disagree with a statement by choosing one category on a 5-point scale that ranges from "strongly agree" to "strongly disagree" (See Figure 4.2).
    - (b) Any adaptations that resemble, even superficially, the Likert scale are loosely referred to as **Likert-type** (or **Likert-like**) scales (See Figure 4.3).
    - (c) Most Likert and Likert-like scales include a middle neutral point because sometimes people legitimately neither agree nor disagree or don't know how they feel.
    - (d) Some researchers use visual Likert-type scales to measure people's attitudes or feelings (See Figure 4.4).
    - (e) Researchers also use Likert and Likert-type scales to rate/evaluate people's behavior.
    - (f) Likert and Likert-type scales typically are scored by assigning the number 1 to the category at one end of the scale and consecutively higher numbers to the next categories, up to the number 5 (or 7, etc.) at the other end of the scale, and these numbers can be sometimes summed.
  - v. **Semantic differential** scales were developed to measure the *meanings* people ascribe to a specific stimulus.
    - (a) Semantic differential scales present a stimulus item at the top of a list of (usually) 7-point scales representing polar opposites.
    - (b) Respondents choose a single point on each scale that expresses their perception of the stimulus object (See Figure 4.5).
    - (c) Semantic differential scales are used frequently in communication research.
    - (d) The selection of the bipolar adjectives for semantic differential scales is not arbitrary or random.
    - (e) Most measure three dimensions: evaluation, potency, and activity.
    - (f) The research population must also see the paired adjectives as opposites, which may be problematic.
    - (g) The construction of effective semantic differential scales, therefore, requires careful pretesting.
  - h. Like Likert and Likert-type scales, semantic differential scales are scored with numbers 1 through 7 assigned to the points on a scale.
    - i. Many scholars question, however, whether the distance between the points on these scales actually are equivalent (e.g., whether the distance between "strongly agree" and "agree" is actually equal to the distance between "agree" and "neither agree or disagree").
    - ii. One type of scale that attempts to ensure that the distances are equal is the **Thurstone scale**.
      - (a) To construct a Thurstone scale, a researcher first generates many statements, usually several hundred, related to the referent being investigated.

- (b) A large number of judges, usually 50 to 300, then independently categorize the statements into 11 categories, ranging from “extremely favorable” to “extremely unfavorable.”
  - (c) The researcher selects those statements, usually about 20, that all judges code consistently into a particular category.
  - (d) While the Thurstone scale is superior to Likert and Likert-type, and semantic differential scales, it takes a lot of time and energy to construct; and this is probably why it is seldom used in social-scientific research.
  - (e) In actual practice, researchers use Likert, Likert-type, and semantic differential scales, and assume that the distance between adjacent points on these scales is equivalent.
  - (f) This assumption allows them to use advanced statistical procedures that require interval-level data (See Chapters 13 and 14).
- vii. **Ratio measurement scales** not only categorize and rank order a variable along a scale with equal intervals between adjacent points but also establish an *absolute, or true, zero point* where the variable being measured ceases to exist.
- (a) Because of the absolute zero point, ratio measurements cannot have negative values, since the smallest value on ratio scale is zero.
  - (b) Ratio measurements are common in the physical sciences, but rare in the social sciences.

#### F. Measuring Unidimensional and Multidimensional Concepts

1. Many concepts/variables can be measured by asking only one question on a single-item scale.
2. Measuring complex concepts requires multiple items.
  - a. These several indicators are then combined in some manner to yield the desired measurement.
  - b. When all those indicators relate to one concept, that concept is called *unidimensional*.
  - c. When indicators relate to several subconcepts, the concept that unites them is called *multidimensional*.
  - d. The statistical procedure that reveals whether a concept is unidimensional or multidimensional is called **factor analysis** (See Chapter 14).
3. **Unidimensional concepts** are measured by a set of indicators that can be added together equally to derive a single, overall score.
  - a. A scale comprised of such items is called a **summated scale**.
  - b. S. Booth-Butterfield and Booth-Butterfield (1991) took 17 observable characteristics associated with the variable of humor orientation (HO), and asked 275 people to indicate on a Likert scale the extent to which each one applied to themselves.
  - c. A factor analysis revealed that responses to these 17 items were related to each other and, therefore, they were measuring a unidimensional concept (see Figure 4.6).
4. Many communication concepts are composed of a number of different subconcepts, called **factors**.
5. Concepts that incorporate more than one factor and, therefore, must be measured by more than one set of scale items are called **multidimensional concepts**.

#### IV. Measurement Methods

##### A. Researchers use three general measurement methods: self-reports, others’ reports, and observing behavioral acts.

1. Most researchers use **self-reports** to measure their target characteristics/behaviors.
  - a. They ask people to **comment** on themselves.
2. There are both advantages and disadvantages to using self-reports.
  - a. Self-reports are an efficient way to ascertain respondents’ beliefs, attitudes, and values.
    - i. These are psychological characteristics that exist inside of people’s heads, which makes them impossible to observe directly, so asking people what they think, like, or value makes sense.
  - b. Self-reports can also sometimes be a good measure of how people behave.
    - i. The researcher asks them how often they do something.
    - ii. Some social scientists distrust self-reports of behavior and believe that observations of behavior are more accurate.

- iii. It is not the measurement method per se that is better, it is a matter of which method yields the most valid information about people's behavior (see Chapter 5).
- iv. People also may provide inaccurate information when asked to step outside themselves and comment on behaviors they not normally think about or remember.
  - (a) You've probably met people who don't remember what they said two minutes before, let alone what they said last week, last month, or last year.
  - (b) Those people who pay more attention to their verbal and nonverbal behaviors, a practice called *self-monitoring*.
  - © Sometimes people don't report the truth.
  - (d) Self-reports about controversial issues or deviant behavior are questionable, due in part, to a **social desirability bias**, the tendency for people to answer in socially desirable ways.
  - (e) A social desirability bias potentially compromises the validity of many self-reports).

#### B. Others' Reports

- 1. A second general method for measuring how much people demonstrate particular characteristics/behaviors is through **others' reports**, that is, asking people to describe other people.
  - a. In some cases, others' reports may be more accurate than self-reports.
  - b. But others' reports can be as susceptible to inaccuracy as self-reports.
  - c. To compensate for the strengths and weaknesses of self-reports and others' reports, researchers sometimes use triangulated measurement. Thus, when respondents and others agree that they behave in particular ways, researchers have more confidence in the accuracy of their measurements.

#### C. Behavioral Acts

- 1. Researchers can sometimes observe a person's *behavior* to assess an operationally defined concept.
  - a. Most researchers believe that direct measures of a person's behavior are more accurate than self-report or others' reports.
  - b. In two instances, however, behavioral acts aren't as useful to researchers as self-reports or others' reports: first behaviors show what people do, not what they feel or believe, and secondly, researchers must be sure that the behaviors observed accurately represent the concept of interest.

#### V. Measurement Techniques

##### A. Three specific measurement techniques are used in conjunction with self-reports, others' reports, and behavioral acts.

- 1. *Questionnaires*: The presentation of written questions to evoke written responses from people;
- 2. *Interviews*: The presentation of spoken questions to evoke spoken responses from people;
- 3. *Observations*: The systematic inspection and interpretation of behavioral phenomena.
  - a. Researchers have developed many **instruments** for using these measurement techniques; that is specific, formal measurement tools together data about research variables.

##### B. Questionnaires and Interviews

- 1. **Questionnaires** are probably the measurement technique used most frequently in communication research.
- 2. **Interviews** are employed in many of the same situations as questionnaires.
- 3. The use of questionnaires and interviews presents three important issues: closed versus open-ended questions; question strategies and formats; and the relative advantages and disadvantages of both measurement techniques.

##### C. Closed versus Open Questions

- 1. **Closed questions** (also called **closed-ended questions**) provide respondents with preselected answers from which they choose or call for a precise bit of information.
- 2. **Open questions** (also called **open-ended questions** and **unstructured items**) ask respondents to use their own words in answering questions.

3. Because closed questions provide limited options or seek quantitative responses, respondents answer using terms *researchers* consider important, and, thereby, give direct information about researchers' concerns.
4. Open questions are more time consuming for researchers to administer and for respondents to answer.
5. Researchers sometimes use both open and closed questions in the same study.

#### D. Question Strategies and Formats

1. Questionnaires and interviews are structured in a variety of ways, depending on the type and arrangement of questions.
2. **Directive questionnaires and interviews** present respondents with a predetermined sequence of questions.
3. In **nondirective questionnaires and interviews**, respondents' initial responses determine what they will be asked next.
4. The list of questions that guide an interview is referred to as the **interview schedule**, or **protocol**.
  - a. **Structured interviews** list all questions an interviewer is supposed to ask, and interviewers are expected to follow that schedule consistently so that interviews conducted by different interviewers and with different respondents will be the same.
  - b. **Semi-structured interviews** allow interviewers to ask a set of basic questions on the interview schedule, but they are free to ask probing follow-up questions, as well, usually to gather specific details or more complete answers.
  - c. **Unstructured interviews** provide a list of topics for interviewers, but they have maximum freedom to decide the focus, phrasing, and order of questions.
5. The strategic *sequence* of queries on questionnaires and interviews is referred to as the **question format** and there are three common types.
  - a. **Tunnel format:** Provides respondents with a similar set of questions to answer and researchers with a consistent set of responses to code.
  - b. **Funnel format:** Broad, open-ended questions are used to introduce the questionnaires or interview followed by narrower, closed questions that seek more specific information.
  - c. **Inverted funnel format:** Begins with narrow, closed questions and builds to broader, open questions.
6. Researchers try to structure their question format to avoid **question order effects**.
  - a. These can occur when responses to earlier questions influence how people respond to later questions.
  - b. Researchers also try to avoid, **response set (style)**, that is, the tendency for respondents to answer the questions the same way automatically rather than thinking about each individual question.
7. There are relative advantages and disadvantages to questionnaires and interviews.
  - a. For questionnaires, they usually are more economical than other methods. and reach a large audience in a short period of time with a short turnaround time (see Figure 4.7).
  - b. For interviews, those done by phone can reach remote respondents quickly, and they provide an opportunity to gather observational data, verbal and nonverbal (see Figure 4.7).
  - c. In the final analysis, choosing between questionnaires and interviews (or using both) depends on three factors: The research population, the research questions, and the available resources.

#### E. Observations

1. **Observations**, the systematic inspection and interpretation of behavioral phenomena, are also used to gather data about communication variables
2. **Direct observation:** Researchers watch people engaging in communication.
  - a. Sometimes these occur in a *laboratory setting*.
  - b. Other times, researchers go out into the *field* to observe people as they engage in everyday activities.
  - c. In some cases the people know they are being observed.

- d. In other situations, people know a researcher is present, but not that they are being observed.
- e. Finally, people may not know that a researcher is present and that they are being observed.
- 3. **Indirect observation:** Researchers examine communication artifacts, texts produced by people, as opposed to live communication events (See Chapter 9).
  - a. These artifacts can include recordings, conversations, books, magazines, TV programs, Internet messages, etc.
  - b. **Trace measures:** Physical evidence, such as footprints, hair, ticket stubs, or cigarette butts left behind.
  - c. **Measures of erosion:** These physical signs show how objects are worn down over time (e.g., the wear and tear of a book's pages or faded nature of a couch's fabric).
  - d. **Measures of accretion:** These physical signs show how traces build up over time (e.g., the number of different fingerprints appearing on a magazine advertisement).
    - i. These measures can be *natural* or *controlled* based on whether the artifacts build up naturally or are set up by a person.
  - e. Both measures of erosion and accretion, whether natural or controlled, only provide a measure of a physical trace; researchers must interpret what that physical trace means.
  - f. **Unobtrusive or nonreactive measures:** Both indirect observations and covert direct observations in which people don't know they are being observed, and the presumption is that their behavior will be natural or unchanged.

#### F. Methods of Observation

- 1. There are many ways observations can be recorded systematically including using electronic devices; taking notes and writing down what occurs as it happens, and using audio or video recorders, as well as observing messages sent online through the Internet.
- 2. To categorize these multifaceted observations, researchers use **coding schemes**, classification systems that describe the nature or quantify the frequency of particular communication behaviors.
  - a. The process of coding observations, like types of questions on questionnaires and interviews, ranges from closed to open.
  - b. Researchers using closed coding procedures can sometimes enter their observations directly into a computer program (See EVENTLOG).
  - c. Developing effective closed coding schemes is a complex task.
  - d. At the other extreme is open coding or categorization during or after observations have taken place.
    - i. Ethnographic researchers typically use this procedure when observing people's behavior (See Chapter 10).

#### VI. Conclusion

- A. Observing and measuring behavior is a central part of many communication research studies.
- B. Just as people in everyday life hate to be misunderstood and described inaccurately to someone else, it is equally damaging to good communication research for researchers to judge hastily or imprecisely what the people they study say or do.