

Self-Report Measurement

JAMES C. McCROSKEY

For over fifty years, communication avoidance, anxiety, and fear have constituted a major concern of social scientists who study communication. In fact, this area represents the oldest continuing research effort in the field of communication. Throughout this half century of research, continuing attention has been directed toward the issue of measurement.

Three major approaches to measurement initiated in the early days of the research continue to the present. Lomas (1934) and Gilkinson (1942) began the stream of research that employs self-report measures. Henning (1935) initiated research using observer ratings. Redding (1936) launched the research that focuses on measurement of physiological arousal.

Almost exactly midway in this five decades of research, Clevenger (1959) published a major review of the research done to that time. He found that all three approaches to measurement generated highly reliable scores. However, he observed that the scores did not seem to be related to each other meaningfully. As he put it:

Surprisingly, instruments which are so reliable display comparatively poor intercorrelations. Results of comparisons of various indices of stage fright suggest that the emotional disturbance which is recorded on physiological measuring devices is different from both the emotional disturbance which the speaker reports having experienced, and the emotional disturbance which a group of judges report having observed, and that the latter are different from each other. (Clevenger, 1959, p.137)

The overwhelming concern of this early research was anxiety or fear relating to public speaking, or, more commonly, stage fright. In recent years communication avoidance, anxiety, and fear have been examined

in other communication contexts. However, a survey of that literature indicates that Clevenger's conclusion in 1959 could as well be drawn today. Self-reports, physiological arousal indicants, and observer ratings, while often significantly correlated, are not measures of the same thing. Isomorphism, concurrent validity, and interchangeability simply are not present.

While Clevenger's (1959) conclusion, as well as the one I draw above, is based on the empirical findings of research, careful conceptualization prior to that research would have made the findings expected rather than surprising. In most cases, such care in conceptualizing communication avoidance, shyness, communication apprehension, and reticence has not been present prior to instrument development. Of particular importance for conceptualization in this area is the trait-state distinction advanced by Spielberger (1966). While this dichotomy is now discredited by many writers, the underlying concept of the distinction is critical for our understanding of measurement in this area.

Human behavior is the product of at least two interacting factors: characteristic predispositions of the individual (traits) and situational constraints on behavior at a given time (states). Individual traits are relatively enduring over time, whereas states are highly variable. Applying this to the common problem of stage fright, a person may be generally apprehensive about giving speeches and thus will experience considerable anxiety when forced into giving a speech. Another person may generally enjoy and not fear giving speeches, but if asked to give a speech on television with insufficient time to prepare may experience a comparable amount of anxiety. The anxiety of one stems primarily from the trait of the individual, the anxiety of the other is primarily the result of the situation itself. Depending on how we choose to measure stage fright in this instance, two measures may be either highly correlated or totally uncorrelated.

Communication avoidance, shyness, communication apprehension, and reticence can all be measured by self-report, observer rating, or physiological arousal indicant at either a trait or state level. However, the three measurement approaches are not equally useful for all purposes or at all levels. The primary thing we must determine before we select or construct a measure is what we want to measure. The three primary options available are physiological arousal, behavioral disruption, and cognitive discomfort. I will consider each measurement approach below and suggest when each may be most useful.

Choosing Measurement Approach

Indicants of Physiological Arousal

If one wishes to measure physiological arousal, the many instruments that record indicants of physiological arousal are an obvious choice. The case for such measures appears strong, at least on the surface. Physiological responses are hard to fake and the instruments, if handled by competent professionals, are not as subject to such human frailties as demand characteristics and experimenter biases as are other instruments. However, there are many pitfalls in the use of such instruments. Few scientists and even fewer teachers are trained in the appropriate use of the technology involved. This equipment in the hands of the untrained individual is worse than useless and very likely to lead to false knowledge claims that are difficult to identify in research reports. In addition, use of such instruments for screening large numbers of students would seldom be economically or strategically feasible. Finally, as Beatty notes in Chapter 6, many issues are not yet settled among even the experts on how such data should be analyzed and interpreted.

The usefulness of physiological measures, for all intents and purposes, is restricted to state issues. Because of the difficulty of calibrating such equipment to each individual, the possibility of using such measures across a sufficient variety of communication settings to generate an estimate of a trait arousal level is extremely remote and, in the case of some instruments, completely impossible with current technology.

Finally, we must address the question of whether we should attempt to measure arousal at all. Arousal does not equal anxiety; arousal simply equals arousal. Considerable research indicates that people who report experiencing anxiety and people who report feeling exhilaration can have highly similar arousal levels. The measures of arousal, then, have insufficient face validity as indicants of communication avoidance, shyness, communication apprehension, or reticence to deserve research attention from researchers concerned with these constructs. The sole exception would be research that employs previously validated measures of these constructs along with measures of physiological arousal in the attempt to discover the physiological impact of these constructs in various communication settings. Physiological measurement, then, is the least useful approach for measurement in this area.

Behavioral Observation

Most scholars concerned with shyness, communication apprehension, or reticence profess their concern as a result of the presumed relationship between these constructs and potentially observable human behaviors that may have important positive or negative effects in the life of an individual. While I share this concern, I also believe that communication apprehension (in particular as it is conceptualized in Chapter 1) can have important effects on the individual that are not observable in behavior. Nevertheless, much of the research in this general area is amenable to the use of behavioral observation.

As reticence is currently conceptualized (the inability to perform competent communication behaviors), behavioral observation probably is the most valid and useful approach to measurement. Predispositional and physiological arousal factors are only marginally related to this construct. Thus, whether one is interested in a general trait of reticence or reticence in a specific state setting, observed behavior is the only indicant with strong face validity.

Communication avoidance and shyness (which I see as virtually isomorphic constructs) are also amenable to measurement by behavioral observation. Both envision the person who is shy or avoidant as engaging in less communication, either generally (trait) or in a given situation (state). Thus observation of behavior should provide a measure with strong face validity if carefully administered.

Behavioral observation is least useful as a measure of communication apprehension because there are no behaviors that are specifically predicted from the communication apprehension conceptualization. Rather, this conceptualization envisions a variety of options available to the individual, many of which are not observable. In short, measurement by behavioral observation has little or no face validity for communication apprehension.

To say that behavioral observation is potentially very useful for measurement of reticence, communication avoidance, and shyness is not to say that it should always be the measure of choice. This approach has several limitations that may force its rejection for some purposes.

Observing behavior is often, although not always, an intrusive approach to measurement. The observer may, and often does, alter the behavior being observed. This is particularly true in dyadic communication settings, but can apply to any setting. One method advocated for overcoming this problem is use of video recording. However, as one who has carefully debriefed subjects who knew they were being

videotaped, this is only a "better than nothing" improvement. Intrusion is still there.

The most difficult problem for the researcher considering the use of behavioral observation is the determination of what behavior is to be observed. For example, if reticence is the reverse of communication competence, what behaviors shall be taken as evidence of reticence? As a field we are in far from complete agreement about the nature of what we shall call "competent" communication. We simply have not yet come to agreement on the set of behaviors that will operationally define this construct. Some believe we never will. In the absence of such general agreement, we have operationalism at its worst — reticence is whatever my measure measures. Such an approach is unlikely to lead to major advance in knowledge.

Behavioral observation probably is most useful for assessing states and least useful for assessing traits. This is not inherent in the technology of the method, but is a function of the limitations on the practical application of the method. Behavioral observation is highly time consuming and expensive in most instances. It usually is difficult to apply appropriately even to assess states, since resources typically are more limited than would be desirable. To assess traits requires extensive observation across many settings to generate valid data. Resources are seldom available for such careful observation. Even the use of this method where it may be most appropriate, the screening of students for reticence/competence for purposes of assignment for communication instruction, is seldom economically feasible. Thus the practical alternatives we often confront are to choose another method or to use this method in an inappropriate and invalid way. Unfortunately, the over-reaction of our society to the need to assess communication competence often leads us as individuals as well as our professional associations to accept the latter alternative.

Self-Report

The most widely employed approach to measurement in the areas of communication avoidance, shyness, and communication apprehension is that of self-report scales. There are several reasons for this clear preference on the part of both researchers and practitioners — some good, some not so good.

Many people argue that the best way to find out something about someone is simply to ask her or him. I cannot argue with that logic, except to point out that it is only correct if the person knows the answer

and if the person is willing to tell you the truth. To illustrate: If you ask me how many ounces of blood there are in my body, I can't give you the correct answer because I don't know. However, if you ask whether I feel apprehensive about interviewing with a university president for a position, I can give you a correct and precise answer. In addition, if you are a good friend of mine and you ask me whether I fudge on my income tax return, I know the correct answer and probably will tell you the truth. But if you are a tax auditor and ask me the same question, it is most likely I will tell you I don't, whether I do or not.

Self-report measures, then, are most appropriate when they are directed toward matters of affect and/or perception in circumstances where the respondent has no reason to fear negative consequences from any answer given. They are least useful when they are directed toward matters of fact that may be unknown or unknowable by the respondent. For research involving the constructs of communication avoidance, shyness, communication apprehension, and reticence, these distinctions are critical to the decision of selection of this type of measure.

Self-report measures are the most useful for measuring communication apprehension. Since this construct is directed toward the cognitions of the individual, it is uniquely suited to self-report measurement, if care is taken to avoid causing the respondent to provide false answers. Such measures have strong face validity (some have other strong indications of validity as well) and are the only measures isomorphic with the communication apprehension construct. Physiological and observation measures lack such isomorphism, and thus must rest their case for validity upon observed correlations with previously validated self-report measures.

Self-report measures are amenable to either trait or state concerns with communication apprehension. Respondents can report their general feelings, their feelings in broad categories of communication situations, and their feelings in specific situations with equal ease.

Self-report measures probably are the least useful for measuring reticence. While subjects can report whether they *feel* competent in general or in specific settings (probably an indirect and imprecise measure of communication apprehension), they are not likely to be in a position to know whether they *are* competent. Most likely such self-reports would be overwhelmingly influenced by the respondent's self-esteem. Respondents with high self-esteem would report being competent, while those with low self-esteem would report being incompetent. Such reports might even be somewhat correlated with observed competence, but they would still lack face validity. It has been my

experience, for example, that many people who consider themselves the most competent in interpersonal communication or public speaking (often experts and teachers in the area) are, in fact, among the least competent.

The use of self-report instruments to measure shyness or communication avoidance is problematic. The question that must be addressed is whether respondents actually know how much they talk compared to others. On the one hand, Bernard and Killworth (1977) present convincing evidence that people cannot accurately report with whom they talk on a given day (the basic data for much network research). On the other hand, at least one measure of shyness has generated a respectable validity quotient when assessed against observer ratings (McCroskey and Richmond, 1982a). With this limited information, we may infer that self-reports of shyness may be more valid for trait measurement than they are for state measurement. However, more research is needed before we can hold that conclusion with much confidence.

While determining advice for or against the use of self-report scales for shyness or communication avoidance is difficult at best, the decision probably should be based on two considerations: (1) Does the self-report measure have a substantial case for validity compared to observer ratings? and (2) Is it practically feasible to use observer ratings as an alternative?

As I noted previously, behavioral observation has strong face validity for the measurement of shyness and communication avoidance. Such face validity is lacking for the other measurement approaches. Thus behavioral observation, other things being equal, should be the measure of choice. However, in many instances it is not feasible to employ this method. When this is the case, the researcher should turn to self-report measurement with extreme care. There are a number of shyness measures available in the literature, many of which really are measures of communication apprehension, and most of them have only modest data supporting validity, if any at all. One of the most widely used measures of shyness (Zimbardo, 1977), for example, has a modest case for predictive validity, but I have been unable to generate acceptable validity quotients against observer ratings for this measure. It is probable that this measure is simply a crude and imprecise measure of communication apprehension, since it generates moderate correlations with established communication apprehension measures, and the Zimbardo conceptualization of shyness is very similar to the conceptualization of communication apprehension outlined earlier in this book.

All in all, self-report measures are potentially very useful for researchers concerned with communication avoidance, shyness, and communication apprehension. They represent an inexpensive and efficient method of assessing large numbers of respondents with minimum effort or imposition. Of course, the choice of self-report instrument must hinge on the instrument's demonstrated validity. The use of unvalidated instruments should be avoided carefully. Validated instruments for most purposes are available and should be chosen over the many others that have appeared in the literature.

Available Measures

A wide variety of self-report measures related to communication avoidance, shyness, and communication apprehension have been reported in the literature. Many of these have been used only once or have negative indications of validity, so these will not be considered here. The ones that are discussed below are instruments that are widely used, have a good case for validity, and/or show some promise for future use. These are grouped in four categories: measures of communication avoidance or shyness, measures of communication apprehension in generalized contexts, measures of traitlike communication apprehension, and measures of state communication apprehension.

Measures of Communication Avoidance or Shyness

As noted previously, I am considering shyness to be the tendency to avoid communication and talk less. Shyness has been defined in other ways by other scale developers. Scales that measure "shyness" conceptualized in a different way are considered in another section.

The Predispositions Toward Verbal Communication (PVB) scale was developed by Mortensen, Arntson, and Lustig (1977). This is a 25-item, 7-step scale with good reliability and substantial indications of validity, even though it has received little use from researchers other than its developers. Although a few of the items on the scale appear to be more appropriate for a measure of communication apprehension, the bulk of the items appear to measure a traitlike orientation toward initiating, maintaining, and dominating communication. In general, the PVB appears to be a very good measure of a trait approach-avoidance orientation. This conclusion is supported by substantial relationships

between PVB scores and both observer ratings and actual coded behavior in controlled settings.

The Unwillingness-to-Communicate Scale (UCS) was developed by Burgoon (1976). The scale has been found to have two independent dimensions, one measuring communication apprehension and the other measuring perceived communication rewards. The apprehension dimension is not relevant here and is measured better by other scales to be discussed in a later section. The reward factor is promising as a measure of the construct of concern. However, there is little support for the validity of the instrument available at this point. It has received little attention from researchers since its publication. One major concern about the scale is the *lack* of correlation between the two factors. Conceptualizations of communication apprehension and communication avoidance, as well as the conceptualization of unwillingness to communicate itself, suggest that there should be a substantial relationship. The PVB does evidence such a relationship, and thus probably should be selected in preference over the UCS.

The Stanford Shyness Survey (SSS) developed by Zimbardo (1977) is the best known measure of shyness. However, its use has been limited primarily to its originator and his students. In any event, it is very difficult to determine what the SSS actually measures. Results reported from use of the measure are a close analogue to results of studies employing communication apprehension measures. At present, there is little evidence to support this measure's use as a measure of communication avoidance. The PVB clearly should be preferred over the SSS for this purpose.

The measure of communication avoidance (SHY) that I developed (McCroskey, Andersen, Richmond, and Wheelless, 1981) was generated as an artifact of attempting to develop a measure of communication apprehension with simplified wording that could be used with preliterate children. Preliminary work indicated the presence of two factors that although substantially correlated (around .60), were clearly distinct. The items on the measure relate specifically to the amount a person believes he or she talks compared to others.

The SHY measure is a 14-item, 5-step scale with good reliability and face validity. Its validity also is suggested by moderately high correlations between SHY scores and observer ratings. Also, in unpublished research I have found PVB and SHY to be correlated above .80, suggesting concurrent validity for the two measures. In general, then, either measure can be chosen with some confidence. However, if length

of scale or simplicity of wording is an important concern, the SHY scale should be preferred because it is shorter and simpler.

Measures of Communication Apprehension in Generalized Contexts

Most of the early measures that fit in this category of instruments were concerned with communication apprehension in the public speaking context, commonly called "stage fright." Only recently have measures appeared that attempt to measure communication apprehension in other generalized contexts. All of these measures attempt to tap a traitlike orientation that applies only to a specific type of communication context (as opposed to a traitlike orientation that cuts across contexts, which will be considered in the next section).

The first widely used measure of traitlike stage fright was the Personal Report of Confidence as a Speaker (PRCS) developed by Gilkinson (1942). Numerous shorter versions of the PRCS have appeared in the literature, the most commonly used being one reported by Paul (1966). Both the longer and the shorter versions have been demonstrated to be highly reliable and numerous research studies point to their validity.

Since the PRCS employs a forced true-false response option, I developed the Personal Report of Public Speaking Anxiety (PRPSA), a 34-item, 5-step, Likert-type scale, to increase precision of measurement (McCroskey, 1970). This scale is highly reliable and was found to maintain that reliability, in subsequent research, with only half as many items (Hensley and Batty, 1974). The concurrent validity of this scale was demonstrated by correlations above .80 with the PRCS. Those same correlations, of course, question the need for the scale, at least in its long, 34-item form. It is essentially equivalent to, not superior to, the PRCS. It might be preferred to the PRCS if the short form (17-item) is used when length of measurement is a problem.

Buss (1980) reports use of two scales that appear to measure communication apprehension in separate categories of communication contexts. One is referred to as a measure of audience anxiety, a 5-item scale with a reliability of .73. Unfortunately, three of the items relate to public speaking and two relate to talking in class, two contexts that have been found to be correlated yet distinct in terms of the level of generalized communication apprehension they generate. Consequently, this scale is not recommended for use.

The second scale is referred to as a shyness scale, a 9-item scale with a reliability of .78. While this scale clearly does not measure shyness as conceptualized here it does appear to be a fairly good measure of communication apprehension in dyadic and small social group contexts. The problem, as with the audience anxiety scale, is that the contexts of dyadic and small group communication are confounded in the scale. Apprehension concerning these two contexts has been found to be correlated, but the levels are distinct from each other and often dramatically different for a given individual. Consequently, this scale is not recommended for use.

Leary (1982) also has developed two instruments that appear to measure communication apprehension in separate categories of communication contexts. Unlike the Buss (1980) scales, these are clear measures of two distinct contexts. The first scale is the Interaction Anxiousness (IA) scale. This is a 15-item, 5-step, Likert-type scale tapping apprehension about interpersonal, primarily dyadic, communication. The second scale is the Audience Anxiousness (AA) scale. This is a 12-item, 5-step, Likert-type scale tapping stage fright, primarily in the public speaking context. Both instruments generate good reliability and there is some evidence for validity.

As can be seen from the above discussions, we have several instruments that appear to be very satisfactory measures of communication apprehension in the public speaking context, but we have no widely used or strongly validated measures of traitlike communication apprehension for the contexts of talking in meetings or classes, talking in small group settings, or talking in dyadic settings. The only measures that seem to hold some promise for these purposes, other than the Leary IA scale, are those I developed with Richmond (McCroskey and Richmond, 1982a). Each scale is a 10-item, 5-step, Likert-type instrument. The reliabilities we have obtained have been very satisfactory and the face validity of the instruments is good. However, at present no other indications of validity have been established. Thus I recommend the use of these scales only on the basis of the absence of other scales and the promising outlook for the validity of these scales, yet to be verified.

Measures of Traitlike Communication Apprehension

The Personal Report of Communication Apprehension (PRCA), in its various forms, has been the self-report measure employed in the

overwhelming majority of studies involving traitlike communication apprehension. There are 20-item (McCroskey, 1970), 10-item (McCroskey, 1978), 25-item (McCroskey, 1978), and 24-item (McCroskey, 1982) versions of this instrument available. All use 5-step, Likert-type response formats. The reliability of all of the forms is very high, usually above .90. The forms are correlated around .90. There is overwhelming evidence for the predictive validity of the measures.

The 10-, 20-, and 25-item versions of the instrument have been criticized for their inclusion of a disproportionate number of items relating to public speaking when the instruments purport to tap traitlike communication apprehension across communication contexts. The most recently developed, 24-item, version of the instrument overcomes this criticism since it includes 6 items for each of 4 contexts: public speaking, talking in meetings or classes, talking in small groups, and talking in dyads. This version also permits the generation of 4 subscores (1 for each context) as well as an overall score. In some as yet unpublished research, the subscores have been found to differ in their predictive power for a variety of dependent variables. Thus the clear choice of which form to use is the 24-item version.

Another measure of traitlike communication apprehension that may be recommended for some uses is the Personal Report of Communication Fear (PRCF; McCroskey et al. 1981). This measure is not completely isomorphic with the various forms of the PRCA, hence its slightly different name. It was designed to employ a much smaller vocabulary than the PRCA so it could be used with preliterate children. With older children and adults the two measures correlate around .80, suggesting substantial concurrent validity but not isomorphism. This measure can be used with small children, but the PRCA-24 should be used with all older groups. One problem with the PRCF when used with young children is that many of the items are worded in a negative way to avoid response bias. Young children have considerable difficulty handling the double-negative response of disagreeing with a negatively worded item. Hence, since this measure must be presented orally to the young children, the administrator of the instrument must take great care to be certain that the correct response from the child is determined and recorded. Carelessness in administration will lead to greatly reduced reliability and validity of the scores.

The Measure of Elementary Communication Apprehension (MECA) was designed to overcome the wording problems in the PRCF (Garrison and Garrison, 1979). The MECA is a 20-item, 5-choice, Likert-type instrument. It is unique in that the response options are

presented in the form of faces, ranging from a broad smile to a broad frown. The problem of response bias is controlled by presenting the faces in reverse order for half of the items. Concurrent validity with the PRCF is very good and fairly good with the PRCA-10. Discriminate validity was also established. This measure should be the instrument of choice when working with small children, but probably should not be used with older children, who may feel the "smiling faces" response options are a bit beneath them.

Measures of State Communication Apprehension

In the past most measurement of traitlike and generalized-context communication apprehension has employed self-report instruments. In contrast, most state measurement of communication apprehension has employed physiological indicants or behavioral observation. The weakness of the latter two approaches for measuring communication apprehension has led recently to the use of self-reports.

The State Anxiety measure developed by Spielberger (1966) has been employed most frequently. This measure can be used to tap any state anxiety, not just state anxiety about a communication situation. It is a 20-item, true-false type of instrument. It has very good reliability and strong face validity. Richmond (1978) employed a modified version of the instrument with good results. She adapted the instructions for the instrument to apply specifically to a communication situation and converted the response options to a 5-step, Likert-type format. She obtained high reliability and a strong indication of validity.

This instrument can be recommended strongly for use as a measure of state communication apprehension. However, the instrument is under copyright (which I do not hold) and fees for use are unusually high. As an alternative the State Communication Apprehension Measure (SCAM) is recommended. This measure was developed by Richmond and I to avoid copyright infringement (McCroskey and Richmond, 1982b). Unlike the Spielberger instrument, the SCAM is a 20-item, 5-step, Likert-type instrument. It has high reliability, around .90, and good face validity.

Conclusion

Self-report instruments are readily available for use as measures of communication avoidance, shyness, and communication apprehension
