Social Approaches to Communication

Edited by

Wendy Leeds-Hurwitz

Foreword by Robert T. Craig

1995

The Guilford Press
New York London

CHAPTER 3

Quantitative versus Qualitative?

JANET BEAVIN BAVELAS

One of the most useful vestiges of my undergraduate education has been an alertness for false dichotomies, that is, for putative opposites such as democracy versus socialism, heredity versus environment, individual versus society. Each of these is, at the very least, likely to be a continuum rather than a dichotomy and is even more likely to be a complex combination of several continua. For example, democracy versus socialism confounds political and economic systems. If we separate the two, we have democracy versus dictatorship and socialism versus capitalism, which create four possible combinations (all of which occur in the modern world). Given especially the prevalence of mixed economies, many more combinations also exist. So much for democracy versus socialism.

The inventor Edwin Land (1962) added to my skepticism about prefabricated choices by suggesting that, when faced with a polarity on which one seems required to choose, there is always the option of thinking orthogonally. That is, rather than being limited to positions along a one-dimensional line, it is possible to cut across it, strike out in another direction, and even create another dimension.

These practical and comfortable old ideas have remained useful, and I admit they shaped my initial reaction to rumblings about quantitative versus qualitative research: Here is another false dichotomy that is best handled by ignoring it and going in a direction completely independent of either side. (Indeed, my primary credentials for writing this chapter are that, over my career, I have been equally criticized from both sides.) It did not take long, however, to realize that it was a mistake to dismiss the issue, because people were taking it very seriously. Soon, I had to admit

that, like everything else since I was an undergraduate, this issue had become much more sophisticated and that there had been massive inflation as well. The qualitative-versus-quantitative debate could now be described, not entirely whimsically, as a series of concatenated false dichotomies, having expanded in scale and complexity to the status of (equally false) isomorphisms.

In mathematical logic, an isomorphism is "a one-to-one relation onto the map between two sets, which preserves the relations existing between elements in its domain."[1] For example, the set of all odd numbers can be mapped onto the set of all even numbers because, within each set, the elements have exactly the same relations to each other (namely, an ascending difference of two). The set of quantitative versus qualitative research is usually constructed as consisting of only two elements in a relationship of opposition. Moreover, debates about quantitative and qualitative research have explanded to co-opt several other sets and mapped them onto the original. An example is experimental versus nonexperimental: Quantitative research is experimental, whereas its polar opposite, qualitative research, is nonexperimental (the polar opposite in the new domain).

Another way to visualize this way of thinking is to imagine a child's stacking toy that consists of short rods with a round cup at one end and a square cup at the other. The starting piece is white, its square cup is called quantitative and its round one is called qualitative. Another piece is red; its square cup is called experimental and its round cup is called nonexperimental. If there is a way to stack the red piece smoothly into the white one (like egg cartons), an isomorphism exists. I propose that a lot of people believe that the experimental end of one piece fits smoothly into the quantitative end of the other, and that the nonexperimental cup also fits smoothly into the qualitative cup.

What surprises me is that there are so many proposed isomorphies in the debate over quantitative versus qualitative research. Table 3.1 summarizes 13 domains that have been stacked neatly into the basic quantitative-versus-qualitative piece. The only disagreement is which way the last piece fits: Is it good to be quantitative/experimental/statistical/deductive (etc.) or is it good to be qualitative/nonexperimental/nonstatistical/inductive (etc.)? That is, researchers seem to agree to let themselves be lined up neatly on opposite sides of these dichotomies; they only disagree about which side it is better to be on.

The purpose of this chapter is to suggest that these are all *false* isomorphies, that diverse aspects of a complex process such as research cannot be simplified into a child's stacking toy. These differences are socially constructed, and to the extent that we insist on maintaining them, we will severely limit the number of approaches we can invent to explore

Table 3.1. False Isomorphies

Quantitative research	Qualitative research
Numbers	No numbers
Parametric	Nonparametric
Statistics	No statistics
Empirical	Not empirical
Objective	Subjective
Deductive	Inductive
Hypothesis testing	Exploratory
Experimental	Nonexperimental
Laboratory	Real world
Artificial	Natural
Not generalizable	Generalizable
Internal validity	External validity
Good guys	Bad guys
or	
Bad guys	Good guys

our common interests. A highly restricted choice of methods inevitably stunts the growth of theory as well. On the other hand, if we reject polarization we may discover new, previously unexplored combinations of both approaches. In the following pages, I am going to defer my deconstruction of the obvious distinctions between quantitative and qualitative research (the use of numbers and statistics) until we have examined the other dichotomies listed in Table 3.1; these may be considered less central and defining, but they are equally important in practice.

EMPIRICAL VERSUS NONEMPIRICAL

One of the most extreme isomorphisms can be dealt with quickly. Sometimes a quantitative approach is equated with empirical research whereas qualitative research is treated as nonempirical. It is usually implicitly bad to be nonempirical, but I have also heard quantitative resarch dismissed as "merely" or "trivially" empirical, so (as usual) the good guys–bad guys mapping is open. However, empirical simply means "derived from or guided by experience," which fits any communication researcher from the most dedicated conversation analyst to an equally hard-core multivariate devotee. The alternatives to empirical conclusions are those based on intuition, authority, faith, or some other means of knowing without recourse to data. Clearly, all researchers do empirical

research (unless they ignore the data and draw their own conclusions, which unfortunately is not unusual).

OBJECTIVE VERSUS SUBJECTIVE

An initially more plausible mapping characterizes quantitative research as objective and qualitative research as subjective. This isomorphism suggests that the two kinds of researchers deal in inherently different kinds of data or at least take dramatically different stances toward their data (e.g., detached vs. interpretive). The question is whether there are ever any inherently objective data.

In a research context, the best working definition of objective is *intersubjective agreement.* All measurement is subjective in the sense of requiring some human inference; even reading the weight on a scale is a human perceptual and cognitive operation. Reading a weight or (more relevant here) which number someone chose on a rating scale is a subjective operation, albeit one that is very likely to yield the same result regardless of who does the reading. Many people would call these objective data because of presumed intersubjective agreement, but to do so creates a misleading dividing line by imposing a dichotomy onto a continuum.

A more accurate approach would be to examine the degree of agreement. Measurements on which people cannot agree well might be called subjective; I would prefer simply to call them unreliable. If we take this approach, there are a couple of surprises waiting. First, objectivity is usually equated with physical, noninferential measures, yet in my experience these are not the most likely to yield high agreement. For example, Ekman and Friesen's (1976, 1978) Facial Action Coding System is purely physical, describing various muscle groups in the face rather than their meaning. Interrater reliability for these "objective" descriptions is often quite low. In contrast, our research group always analyzes the *meaning* of the communicative act, and our reliabilities are consistently high. For example, Chovil (1991/1992) developed a highly reliable system that captured the meaning of each facial display (e.g., as portraying someone else's reaction or as looking quizzical). Coates (1994) assessed the meaning of the entire communicative act at the moment (verbal and nonverbal) and whether this was a mock or serious meaning, again with good intersubjective agreement. I believe that high reliabilities for phenomena that are usually eschewed as subjective are possible because analysts are people who, in their ordinary daily interactions, respond to the meaning of their interlocutors' acts rather than to the physical components. Thus, what is usually considered subjective is quite natural for us; what is

presumed to be objective is often quite unnatural. To capture this knowledge is admittedly more work than creating a 7-point scale but well worth the investment.

A second surprise is the benefit of pursuing high intersubjective agreement. There are many qualitative researchers who would be insulted if someone asked whether others could interpret their data in the same way as they do. This attitude can imply that intersubjective reliability is only for trivial phenomena and that complex and interesting phenomena do not lend themselves to such mechanical approaches. In contrast, we have found that trying to explicate what we are seeing or hearing is a goal worth having. When analyzing communication, we are each isolated individuals. Some may take pride in their apparently unique ability to interpret what others do not notice, but I suspect that they, too, have the occasional nagging fear that what they interpret is not really there. Facing this fear by asking others to analyze the same data can lead to the euphoric discovery that others can in fact see it as well. Every time we achieve high intersubjective agreement in our research, we feel a profound sense of intellectual confirmation. In addition to the personal payoff, there is a more concrete reward for seeking inter-subjective agreement: To do so requires a set of written, highly explicit rules of interpretation, which cover all possible instances and problems (Coates, 1993). These guidelines are themselves an outcome of the research and constitute an important part of the subsequent research report, that is, *exactly* what the authors mean by the phenomenon. In our experience, abandoning our comfortable but vague intuitive judgments in favor of a set of rules often results in better analyses; newcomers applying our rules and unemcumbered by our preconceptions can do a better job than we did.

INDUCTIVE VERSUS DEDUCTIVE

One of the most misunderstood notions in social science research is the aristocratic ideal of formal deductive research, which conjures an image of a highly sophisticated intellectual (obviously a quantitative type) who has such a grand theory of communicative phenomena that he or she can easily spin off predictions about how people will behave. In contrast, the pedestrian, lower-class inductive researcher rummages around in the detritus, trying to come up with something that will never achieve the urbane generality of the deductivist's theory. The radical inductivist takes pride in being a street person, dealing with day-to-day reality far from the ivory tower.

As usual, the dictionary destroys this constructed contrast by showing that neither term fits what we do:

> Deductive and inductive refer to two distinct logical processes. *Deductive reasoning* is a logical process in which a conclusion drawn from a set of premises contains no more information than the premises taken collectively. All dogs are animals; this is a dog; therefore, this is an animal: The truth of the conclusion is dependent only on the method. All men are apes; this is a man; therefore, this is an ape: The conclusion is logically true, even though the premise is absurd.

Who would want to play such silly games? No empirical researcher uses deductive reasoning in this way, much less relies on it entirely, because the data play no role at all.

> *Inductive reasoning* is a logical process in which a conclusion is proposed that contains more information than the observations or experience on which it is based. Every crow ever seen was black; all crows are black: The truth of the conclusion is verifiable only in terms of future experience and certainty is attainable only if all possible instances have been examined. In the example, there is no certainty that a white crow will not be found tomorrow, although past experience would make such an occurrence seem unlikely.

This process is somewhat closer to empirical research in that it includes actual data, some reasoning to and from the data, and especially the inevitable amount of uncertainty. But it is not all that we do.

What the dichotomy misses is that virtually all researchers engage in a *sequential process* that includes aspects of both forms of reasoning at various points. We get an idea inductively, usually from our own observations or from reading the literature (i.e., others' observations), and then apply deduction: If my hunch is correct, I should see more supporting instances in my data, or if my theory is correct, I should be able to predict the outcome of a formal study. If we are wrong or if new possibilities emerge from the data, we construct a new hypothesis (inductively) and go on to deduce how to tell whether this new idea works better, and so forth. This common process is obscured by the fixed deductive format of the research article. Students often think that the way published research results are presented is the way (and the only way) they are obtained (Bavelas, 1987). More experienced researchers know better but may still act as if the way they report the results is also how they got to them.

HYPOTHESIS TESTING VERSUS EXPLORATORY RESEARCH

Closely related to the deductive-versus-inductive dichotomy is another construction that divides (in one view) the elite hypothesis testers from

the mere triflers, who are perhaps too intellectually limited or too cowardly to scale the heights and play for big stakes. Or, seen from the other side, the intellects of hypothesis testers are, if not actually in rigor mortis, at least severely anal retentive: Hypothesis testers are permanently imprisoned inside their own paradigms, whereas the explorers go boldly (and qualitatively) where no one has gone before. Obviously, if we again reject class systems and look at the *process* of any particular research project, these stereotypes quickly stop working.

For example, the moment a quantitative, hypothesis testing researcher adds more variables than he or she has clear and firm hypotheses to cover, the research becomes exploratory. Given the current popularity of multivariate research, this approach is rapidly becoming the norm. The exploration is done by the statistic of choice rather than by the researcher, but it is just as preliminary and subject to confirmation as any other exploratory approach. Classical hypothesis testing research should have few variables and simple statistics, sometimes just a single one-tailed test.

Researchers who identify themselves as qualitative often do not recognize the important role hypothesis testing plays in their analyses. They probably start with hypotheses about a certain kind of data being particularly rich, or about their method being more appropriate than others. And, fortunately, they cannot avoid hypotheses about what they begin to find. Inductive researchers do not simply describe the data; they move from specific instances to some level of generality that subsumes these instances. There is a very exciting point in such research when one begins to see a pattern and looks eagerly to find out whether it will hold up. This is hypothesis testing of a kind that needs to be legitimized as a valuable tool but also honed to remove some of its potential defects.

EXPERIMENTAL VERSUS NONEXPERIMENTAL

No self-respecting qualitative researcher would even fantasize about an experiment, much less consciously commit one. Yet the generic meaning of experiment is very inclusive: "A test, trial, or tentative procedure; an act or operation for the purpose of discovering something unknown or of testing a principle, supposition, etc."

This definition reveals a common root with the simple term "experience" and requires only that we try something out, with or without hypothesis, whether to explore or to test. The degree of formality is not prescribed.

In scientific parlance, the *true experiment* is formally defined by its purpose and method. The purpose of a true experiment is to seek information about causality; the method is random assignment of sub-

jects to conditions that have been created ("manipulated") by the experimenter, which is only possible with certain kinds of data. Nothing in the definition of a true experiment requires that the independent or dependent variable have to be quantitative. Actually, experimental conditions are usually qualitative categories, and the outcome can also be categorical or frequency data. That is, the experimenter is free to ask not only "how much" something happened but, alternatively, "how often," "whether," or "what sort." For example, we (Bavelas, Black, Lemery, & Mullet, 1986) varied experimentally whether eye contact occurred or not and measured the pattern of facial expression that resulted. Thus, both the independent and dependent variables were qualitative.

A great deal of communication research is quasi-experimental in that the key variable is neither manipulated nor manipulable (e.g., friendship, gender, marital status, or any other personal characteristic). Although these designs can be very good and can approximate a true experiment, they are technically not experiments, especially if we insist on an experimental–nonexperimental dichotomy. Moreover, "nonexperimental" includes a lot of methodological territory, such as correlational, observational, archival, or case studies. The answer, obviously, is not to dichotomize but rather to talk about kinds of control or intervention and consequent degrees of causal inference that can be made. If the goal is not causal inference, or if the phenomenon of interest is not amenable to experimentation, nonexperimental work is the better choice. If the goal is to get from point A to point B as fast as possible and a highway exists, a car is the best option. But if the goal is more leisurely, or if the terrain is rough and roadless, hiking may be a better choice.

Even a true experiment does not exist simply because the researcher intended to conduct one. I see a surprising number of supposed experiments where (if one reads the procedure section carefully) subjects were not actually randomly assigned or where the conditions differed in more ways than the level of the independent variable. These are not true experiments but failed experiments; they cannot serve their purpose and can only mislead us about causality.

THE LAB VERSUS THE REAL WORLD; ARTIFICIAL VERSUS NATURAL DATA

After years of hearing people talk about the "real world," I think I have finally located it: It is anywhere off a university campus, as far as possible from a research laboratory (where the world is presumably not real). As a keen science fiction fan, I only wish it were true that the lab is not the real world. It would be wonderful to encounter unreal worlds just by

conducting experiments. I keep peeking hopefully into our lab, but it remains just a pleasant, well-equipped room with no paranormal properties. Our lab is different from (and I behave differently in) my office, my kitchen, a movie theater, or the nearby park. But these places are also different from each other, and I behave differently in each of them as well. So which of them is the *real* real world?

The pastoral variation on this theme is to bless all research outside the lab with the label "natural"—or with the more common but questionably derived term "naturalistic." In contrast, lab/experimental/quantitative research, and indeed any behavior that occurs in the presence of an experimental psychologist, is characterized as "artificial" (though only in the negative senses of that term, i.e., not acknowledging its origin as "created by art"). The radical eco-methodological view is that *any* observation renders data artificial or even not real. It is ironic to hear this position supported by a lofty appeal to experimental (lab) physics via the Heisenberg principle, a demonstrably spurious and false analogy (Bavelas, 1984).

Critics of lab research have raised valid questions about the experimentalists' implicit assumption that the behaviors they study somehow occur in a sterile vacuum, unaffected by the messy contexts of everyday life. But the key point of the criticism surely must be that *all* behavior is situationally grounded, that there is always a context that affects behavior, and that context can be hidden but not eliminated by being held constant. It is therefore completely inconsistent to imply that all nonlaboratory behavior has the special quality of being real and natural by virtue of occurring outside the lab. The logical consequence of these critical questions would be to pursue the specifics of how different contexts affect different behaviors, not to propose that some contexts for behavior are real and others are artificial. I know of few people pursuing these questions.

GENERALIZABLE VERSUS NOT GENERALIZABLE

The concept of generalizability can be a more useful way of posing the above questions than is a prejudgment about the physical location of the research. Generalizable conclusions come from a proper sample of a defined population of people, settings, and events—an ideal virtually no researcher achieves. Our question is usually: to which other people/settings/events are my ad hoc results generalizable? Notice that I ask "to which" rather than "whether" the results will generalize or "how much" they might. We cannot sensibly ask whether or not results will generalize because generalizability is not a single property of research, a sort of one-size-fits-all quality that some studies have and others do not. It is probably not even a quantitative ("how much") continuum, although this

is often implied, ironically, by qualitative researchers who argue that their results are "more generalizable" than other kinds of data. Such statements imply that the researcher has systematically inventoried all people/settings/events, created a dimension that ranges from "like few others" to "like most others," and has a method by which particular configurations of people/settings/events can be validly placed on this dimension. A more modest and realistic approach is to consider closely the features of the data one is committed to studying: In what concrete and abstract ways are these data like data in which other other settings?

INTERNAL VERSUS EXTERNAL VALIDITY

The final variation on this particular theme is the claim that quantitative research maximizes internal validity whereas qualitative research maximizes external validity. The implication that we can pursue quite different goals in different studies is a useful one (and even better if we are open-minded about which goals other people are pursuing).

If a study has, because of the above analysis of its generalizability, some claim to being similar to something else, it has some degree of external validity. However, any researcher is entitled to be indifferent to this goal if he or she has a different passion. Some of us can be quite taken from time to time with the links between events that are usually put under the stuffy rubric of *causality*. The measure of understanding we achieve in these efforts is often called internal validity. In my view, discussions of internal validity are usually far too narrow and puritanical, aimed mainly at showing us how hopeless it is to aspire to demonstrate that X causes Y. This narrow focus neglects all the other fuzzy and interesting relationships between X and Y that we might wish to explore, such as "somehow leads to," "accompanies," "precludes," "mutually influences," and so forth. Some students in my research methods courses are invariably distressed when I explain these two kinds of validity and propose that they cannot pursue both in a single study—unless that single study is going to be the only one they do in their life. Realistically, they must choose a priority and pursue it, conceding weakness in the other area. In any case, whenever one of these two broad goals dominates my research, I find the rewards and demands far too great to permit any time for disparaging those who pursue other goals.

NUMBERS VERSUS NO NUMBERS

Finally, the obvious and accepted distinction between quantitative and qualitative research is that quantitative researchers seek, use, and even worship numbers, whereas qualitative researchers avoid them and may treat them as unholy.

This debate is an old one. Many quantitative researchers agree with the scornful opinion of William Thomson (Lord Kelvin), the English physicist and mathematician:

> When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of *science*. (Thomson, 1891–1894, cited in *Bartlett's Familiar Quotations*, 1968, p. 723)

George Miller (1962), one of the pioneers of mathematical psychology and of applications of information theory to human communication, had his own view of Lord Kelvin's dictum:

> In truth, a good case could be made that if your knowledge is meager and unsatisfactory, the last thing in the world you should do is to make measurements. The chance is negligible that you will measure the right things accidentally. Nevertheless, many social and behavioral scientists, assured that measurement is the touchstone of scientific respectability, have rushed out to seek numbers before they knew what the numbers would mean. (p. 79)

Thus, we might ask whether it is "unquantitative" to use meaningless numbers—unreliable, invalid, questionable, or superficial numbers. If so, a lot of quantitative research is in big trouble.

To give equal time to criticism of both sides, we might also ask whether it matters how many numbers one uses, whether more and more numbers are less and less qualitative. If so, quantitative-qualitative would become a quantitatively defined continuum, which hardly seems fair. The alternative, purist approach is to treat even the slightest hint of quantity as an impurity. This dichotomy would require that all words such as "many," "often," "several," "usually," and so on, be expunged from the word processers of qualitative researchers. One can hardly argue that these terms are acceptable because they are used loosely, with no attempt at precision.

PARAMETRIC VERSUS NOMINAL MEASUREMENT

A more sophisticated distinction than the simple numbers-no numbers dichotomy is between different levels of measurement (i.e., nominal,

ordinal, interval, and ratio) (Stevens, 1946). Note that the term "levels" and the order in which these four kinds of measurement are always introduced make it clear (to quantitative researchers) that this is an evolutionary scale. One emerges from the swamp of nominal measurement and struggles toward the pinnacle of evolution, a ratio scale. In this view, nominal measurers are not yet *Homo sapiens*; they are Lord Kelvin's poor beknighted protoscientists.

Qualitative researchers would probably order the levels in the opposite direction, on a moral dimension. The identification and discovery of new kinds of phenomena are their ultimate goals. Qualitative researchers see nominal description as "natural" for communicative phenomena, and their reaction to Arabic numerals (other than as page numbers) resembles the impassioned prejudice of the Crusaders. To convert their categories into numbers is heresy; ratio and interval scales are a descent into the "black arts."

Seldom in this debate does a simple question arise: Which kind of measurement best captures what the researcher wants to know? I would give absolute primacy to the researcher's immediate goals and to the data he or she is working with, rather than to membership in a particular school of numbers. Two illustrations from our team's research may be useful.

In our work on equivocation (e.g., Bavelas, Black, Chovil, & Mullett, 1990), we began like earlier observers who had described the phenomenon categorically—a message was clear or not; it was evasive or not. However, when we began collecting such messages systematically, nominal measurement alone became unsatisfactory because these fascinating messages clearly varied among themselves in both degree and kind. To draw an arbitrary line would be to continue to construct a two-dimensional world of good and bad messages, ignoring the very subtlety we were interested in. Therefore, we used both quantitative and qualitative approaches to the same messages. The reactions to our hybrid approach were puzzling: One communication journal cut out our qualitative analyses and would publish only the tables of numbers. Another communication journal published a review that was scathing about our use of numbers.

Our measurement of motor mimicry (e.g., wincing at someone else's injury) (Bavelas, et al., 1986) went in another direction. In our experiments, people did not react in degrees of motor mimicry; rather they displayed *patterns* of increase, decrease, or no change. Therefore, in this experimental, hypothesis-testing lab study, our measurement was strictly nominal because trying to make it otherwise would have been inappropriate to the phenomenon.

STATISTICS VERSUS NO STATISTICS

The fastest way to identify an author as quantitative or qualitative is to rifle the pages rapidly under your thumb looking for statistical tables (vs. transcripts). Depending on your preference, you can then decide whether to read the article or not. This is an excellent way to avoid being influenced by new ideas.

To deconstruct this difference, recall that there are the two broad kinds of statistics, *descriptive* and *inferential*. Descriptive statistics are used for data reduction (e.g., I tell my class their average on the midterm rather than reading all the individual marks out to them). One may then go on to inferential statistics, which recognize and assess the possibility of chance outcomes (e.g., whether women really did better than men on the midterm). Let us consider each of these in turn.

If the essence of statistics is data reduction, a great deal of quantitative analysis is nonstatistical because it generates *more* rather than fewer numbers. (One of my students, originally trained in this tradition, was worried that she did not have "enough" tables after her planned, simple analysis.) In contrast, most qualitative research definitely does achieve data reduction: Any transcription is selective and reduces the original conversations to a new form. The author's summary of the process or phenomenon discovered in these transcripts is another bite of the apple.

Clearly, quantitative researchers use inferential statistics, often lots of them. Tables and tables of inferential statistics actually create the paradox of *increasing* the probability of chance conclusions: Without appropriate protection of alpha levels and especially without replication of findings, the more inferential statistical analysis the researcher conducts, the greater the possibility of spurious findings. Equally clearly, one will never find a *p*-value in a qualitative research paper. However, one does find another *p*-word, *patterns* in the data, that is, descriptions of events that are interesting precisely because they seem systematic (nonrandom).

At this point, the reader may feel I am playing a shell game, but I am not sure where the peanut is either. It does appear that, in many instances, quantitative research achieves less data reduction and a less accurate assessment of the role of chance than does qualitative research.

This confusion is an appropriate note on which to end, because my goal has been to shake up all the neatly aligned pieces and let them fall in new combinations. Even better, imagine letting your data find their own best fit. In my view, all of us should focus on empirical data, aim for nontrivial objectivity, and make thoughtful choices about numbers and other possibilities: Is it time more for exploration or for hypothesis

testing? *Which* real world is of interest? Is an experiment desirable or appropriate? How much generalization is possible? All these should be dictated by respect for the phenomenon and the state of our knowledge about it.

NOTE

1. Unless otherwise indicated, all definitions are from the second edition of the *Random House Dictionary* (1993).

REFERENCES

Bartlett, J. (1968). *Bartlett's familiar quotations* (14th ed., E. M. Beck, Ed.). Boston: Little, Brown.

Bavelas, J. B. (1984). On "naturalistic" family research. *Family Process, 23*, 337–341.

Bavelas, J. B. (1987). Permitting creativity in science. In D. N. Jackson & J. P. Rushton (Eds.), *Scientific excellence: Origins and assessment* (pp. 307–327). Beverly Hills, CA: Sage.

Bavelas, J. B., Black, A., Chovil, N., & Mullett, J. (1990). *Equivocal communication.* Newbury Park, CA: Sage.

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I *show* how you feel": Motor mimicry as a communicative act. *Journal of Personality and Social Psychology, 50*, 322–329.

Chovil, N. (1991/1992). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction, 25*, 163–194.

Coates, L. (1993). *Methods for developing scoring systems and calculating reliability in discourse analysis.* Unpublished manuscript, University of Victoria, Victoria, British Columbia, Canada.

Coates, L. (1994). *"Just Kidding." Spontaneous ironic humour in dialogue.* Unpublished manuscript, University of Victoria, Victoria, British Columbia, Canada.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Journal of Environmental Psychology and Nonverbal Behavior, 1*, 56–75.

Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Land, E. H. (1962). Unpublished commencement address, Stanford University.

Miller, G. A. (1962). *Psychology: The study of mental life.* New York: Harper & Row.

Random House unabridged dictionary (2nd ed.). (1993). New York: Random House.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.