STRATEGY ISSUES IN THE DEVELOPMENT OF QUALITY OF EMPLOYMENT INDICATORS

Robert P. Quinn, Survey Research Center, The University of Michigan

Problems confronted in the development and application of systems of social indicators are not qualitatively different from those of any other scientific endeavor. Several characteristics of research on social indicators tend, however, to highlight a number of strategy issues that under more customary circumstances can either be easily resolved or comfortably ignored. Among these characteristics are: the magnitude and continuity of the activity; the contribution to it of multiple investigators; its more-than-usual social relevance; and its multiple audiences.

This paper discusses several general strategy issues that must be confronted in the development of social indicators systems in general and systems for monitoring quality of employment in particular. Each of these issues confronted my colleagues and I in the development of two national surveys of the American workforce (Quinn, et al., 1971; Quinn & Mangione, 1973; Quinn & Shepard, 1974). Hopefully, others engaged in similar activities in the future will be able to profit from our efforts in this area.

Defining the Domain of Investigation

Recent discussions of employment and working conditions have been characterized by a boomlet of new terms that generally begin with the phrase "quality of" and are rounded off by such words as "employment," "work," or "working conditions." None of these terms has really been adequately defined, perhaps for good reason. The phrases have often been used almost as passwords to signify to the listener that something about work is being considered that is somehow new and different from the more traditional concerns of management, labor, and government--a strictly exclusionary definition. Where non-exclusionary definitions have been attempted, they have for the most part been so broad as to include "traditional" areas of concern along with the "newer" ones.

Given this broadness, the selection of measures to be used in an indicator system often becomes highly arbitrary. The system based on such a definition may as a result become a catch-all and vulnerable to the inclusion in perpetuity of measures that happened to be relevant to research problems that were "hot" only at the time when the system was developed.

My colleagues and I faced this problem of selection at a practical level when we developed the second of our national surveys in 1973. The first survey's measures of working conditions had covered a wide variety of content areas as suited the heterogeneity of those contributing to the survey and their particular research needs--the domain had indeed been broadly and loosely defined. The crunch came three years later when a decision to repeat the national survey was made. We were faced at that time with the problem of identifying those materials that would be repeated, those that would be added, and those that would be scrapped.

The "quality-effectiveness strategy" was developed in order to solve this problem and to capitalize on the survey's analytic potentials. This strategy defines three general concepts: working conditions, effectiveness, and quality of employment.

The term "working conditions" refers to descriptions of characteristics of a worker's job obtained from any informed source. These descriptions may focus on any characteristic of the job from the cleanliness of the physical work environment to the degree of time pressure for performance, or from the degree of challenge the job provides to the income a worker receives. Since working conditions characterize a job, they are independent of the individual who does the job. This means that they do not include the worker's evaluation of the conditions measured and that different people doing the same job should describe it similarly.

Effectiveness. The term "effectiveness" refers to states or events that have a positive or negative value from the perspective of some person or set of people. Three such perspectives are distinguished: those of employees, their employers, and society as a whole.

Quality of employment. The term "quality of employment" refers to a judgment about working conditions based on the impact that the working conditions have on effectiveness. Thus, good quality of employment from the perspective of a worker would be some combination of working conditions that produce health (one criterion of effectiveness as judged by the worker). Good quality of employment from the perspective of an employer would be working conditions that lead to a productive, profitable organization.

These definitions provided a standard for selecting measures of quality of employment to be carried forward into the second national survey: select only those measures of working conditions that had a demonstrable association with some criterion of effectiveness.

In the 1969 survey, job satisfaction served as the only criterion measure of effectiveness, a criterion governing the selection of the working conditions to be treated subsequently as quality of employment indicators. The 1973 national survey uses four major criteria of effectiveness other than job satisfaction: physical health; depressed mood; drinking behavior; and life satisfaction. A companion study conducted in five employing establishments focuses upon the criteria of performance, absenteeism, turnover, and participation in activities outside of work. In each of those studies the measures of working conditions (i.e., potential quality of employment indicators) have been expanded to include conditions likely to be associated with the particular criteria of effectiveness under investigation. To the extent that they are in fact able to predict these criteria, they will be included as quality of employment measures in later monitoring efforts. What is unusual about this strategy is that measures of satisfaction, health, and so forth,

although relevant to the system of indicators, are not themselves indicators. Once they have served their purpose, they are shelved until some later time when it may be necessary to use them again to determine whether the quality of employment indicators have retained their predictive powers.

Making Value Assumptions Explicit

Identifying and measuring the "quality" of anything requires standards against which quality can be evaluated. As much as we would like to avoid doing so, the identification of such standards may require that social scientists take normative stands. Unfortunately, social scientists may not be without their biases in their selection of standards.

While it is impossible to rule out entirely the intrusion of arbitrary standards in the measurement of quality of employment, it may nevertheless be possible to circumscribe their effects and to make them more obvious. Two types of normative assertions are often invoked and confused in discussions of quality of employment. The first type prescribes various desirable working conditions, such as having a job that provides a great deal of autonomy. The second type of assertion prescribes only desirable outcomes or criteria of effectiveness, such as good health or satisfaction with one's life. The quality-effectiveness strategy, while permitting normative statements, restricts such statements to outcome criteria. Whether a particular working condition is "good" or "bad" thereby becomes an empirical question rather than a value issue: it is "good" to the extent that it is associated with any or all of the criteria of effectiveness thus specified. This certainly does not avoid value questions. It simply makes them a little more amenable to discussion by limiting the range of variables to which they apply and by confining them to the assignment of priorities among criteria of effectiveness.

To keep reminding ourselves of precisely whose values were being given priority in the indicators of quality of employment we have used in our two surveys, we have found it useful to distinguish three different perspectives for evaluating criteria of effectiveness. The first perspective, that of employers, is a familiar one and includes such matters as productivity, withdrawal from work, counter-productive behavior, and adaptability to changing work procedures. Among the outcomes desired from a second perspective, that of employees, are the equally familiar ones of job satisfaction, mental health, physical health, and so forth.

A third perspective can also be invoked: that of the community or the society. Some of the costs and benefits associated with working do not enter into the personal accounting of either employers or employees. For example:

- --Workers whose jobs undermine their health place an additional demand on the nation's already overburdened system of health-care delivery.
- --Workers whose skills and education are underutilized by their jobs represent an obvious social waste.
- --A worker whose expression of dissatisfaction takes the form of reactions that result in termination may become a candidate for subsequent collection of unemployment compensation, an obvious drain on local resources.
- --The income-deficient worker may burden society with a family prone to illness, future welfare costs, and substandard economic contribution.

The assignment of any criterion to a particular perspective may at times be somewhat arbitrary and perhaps even uncharitable. The assignment does not mean to imply, for example, that from the point of view of employers the physical or mental disorders of their workers are of no importance, only that from the perspectives of most employers there are other more important outcomes. Conversely, the assignment does not imply that employees are necessarily indifferent to productivity. Indeed, the harder it is intellectually to assign a criterion to a particular perspective, the more important that criterion is likely to be. According to this rule of thumb, monitoring and action priorities might profitably be assigned to those working conditions that affect outcomes that are patently relevant to all three perspectives. Work-related illnesses and injuries are one example. They are obviously important to the ill or injured worker, represent a cost to his or her employer (in terms, for example, of sick-pay and filling the worker's position while he or she is away from work), and are costly to society as well (e.g., in terms of the illnesses' or injuries' drain on the nation's already scarce medical resources).

The categorization of criteria of effectiveness in terms of perspective serves principally to clarify the value priorities of the investigator. For example, when we first compared the priorities of our two national surveys against our own list of criteria, we discovered that we had been overwhelmingly concerned with criteria of effectiveness that were valuable from an employee's perspective. Using the quality-effectiveness strategy, our resulting indicators of quality of employment would therefore be similarly biased. Our five-establishment study was designed to compensate for this by attempting to identify those working conditions that were associated with criteria of effectiveness that were important principally from an employer's and a societal perspective. Periodic reviews of the investment of our research energy with reference to this list have generally been very sobering by providing constant reminders of when we have gone overboard in our pursuit of matters that involve only limited perspectives.

Differentiating the Domain of Investigation

Given some agreement upon the domain of investigation and the value perspectives involved, there remains yet a question that must be answered before a monitoring system can be instituted: What are the basic dimensions underlying the domain? Simply talking about "good" and "bad" aspects of jobs is not sufficient. Further differentiation is obviously

necessary. But how is such differentiation to be achieved, and what should constitute the basic vocabulary of the system? There are two obvious ways of answering this question. One of these casts the basic dimensions to be investigated in terms of existing theories of behavior. The second defines such dimensions empirically with only minimal reference to theory.

There are certainly plentiful a priori schemes for categorizing working conditions or workers' motivational orientations toward work, and each of these has its well-reasoned, if not always well-documented, theoretical foundations. At the simplest end are those categorizations that comfortably divide the world into two classes: jobs that are "motivating" versus those that are simply "satisfying." At the more complex extreme are those categorizations that define a variety of "needs" for working and assume that each achieves prominence as lower-level needs are progressively satisfied. Other classification schemes proliferate between these two extremes. All of them, however, present a common problem in that adopting them demands an implicit subscription to their assumptions and their logic. To adopt any one of them in a system of social indicators involves the simultaneous adoption of all its attendent theoretical trappings. While this may provide some short-run theoretical continuity to the monitoring effort, it may reduce the usefulness of the data thus collected for subsequent "secondary" analyses by other investigators who do not subscribe to the same theory. A theoretically "tight" monitoring system with an idiosyncratic theoretical orientation may therefore be self-defeating in that it may provide rich data for those subscribing to the theory involved but may toss only scraps to the non-believers.

On the other hand, the basic dimensions of the monitoring effort can be determined on a wholly empirical basis. This would involve subjecting the system's measures to some kind of dimensional analysis--factor analysis, cluster analysis, small-space analysis, or what-have-you. Many such dimensional analyses of jobs are already available, and they concur in their identification of from about five to eight "basic" dimensions of work. Such agreement is, however, based exclusively upon dimensional analyses of satisfaction ratings of job characteristics. Unfortunately, the same facets do not emerge when dimensional analyses are made of the importance ratings that workers assign to job facets. Nor do they routinely emerge from dimensional analyses of working conditions (i.e., job descriptors).

How, therefore, are facets of jobs to be compartmentalized and differentiated? If this is to be a wholly theoretical procedure, whose theory should be subscribed to and at what cost in terms of making the indicator system unpalatable to others? If it is to be strictly an empirical matter, what are to be the inputs to the relevant dimensional analyses: importance ratings of job characteristics? working conditions with regard to these job facets? satisfaction with regard to these facets? or something else entirely?

Using "Standards" versus "Fit" Models

An important issue in understanding responses to work is whether people are better off by maximizing all the "good" qualities of their jobs or whether they are better off by each person obtaining personally-attuned optimum levels of the particular outcomes that he or she desires most. Measures of the "fit" between the worker and his or her work environment have often been hypothesized to be better predictors of effectiveness than are simply the amounts of the working conditions in question. Indicators of quality of employment that do not accomodate differences in what workers want from their jobs rankle. They are superficially contrary to many theories of human behavior, not to mention common sense. They seem at odds with the advocacy of individually-oriented programs or worker training or job change. And they run counter to the ideology that every person should be treated as an individual.

But indicators of quality of employment that ignore individual differences thrive. Moreover, they constitute the vast majority of those indicators that are publicly available. That such indicators have persisted and remain useful is due in large part to the selectivity of their foci and their assumptions about workers' needs or desires. They focus principally upon those areas of working life that are of concern to most workers and areas where what is desirable to the vast majority of them can be safely assumed (e.g., health and safety, income adequacy).

While "fit" models relate more meaningfully to theories of human behavior, how feasible is their application in a system monitoring quality of employment? A classic example of attempts to translate "fit" models into measuring instruments are those that weight satisfaction ratings of particular job facets (e.g., fringe benefits, security, competent supervision) by the importance that different individuals assign to these facets. While such a weighting procedure is theoretically persuasive and seems easily performable, attempts to improve job satisfaction measures by this procedure have without exception failed to justify it. Whether to include measures of individual preferences in systems of quality of employment indicators seems, therefore, less a theoretical issue than a practical one. It is difficult to argue against their inclusion on theoretical grounds. The only remaining problem is to determine when they constitute a necessity and when they are only a nicety.

r b ms f C mm ca i [1]

Because of its presumed usefulness to society, many results of research involving social indicators are likely to reach a broader audience than social scientists are accustomed to addressing. This, naturally, creates uncommon problems of data presentation and interpretation, only some of which can be overcome by careful use of language. A deeper-rooted problem involves the use of complex measures, particularly psychological measures based on continuous scales. Those who are not social scientists expect social scientific data

to be quantifiable in the same terms that the data of other sciences are. That is, absolute quantities of things can be tallied, or the percentages of people of various types can be specified. Pollsters tend to perpetuate this expectation by declaring, for example, that such-and-such percentage of Americans disapprove of what the President is doing--as if approval and disapproval had been measured as absolutes.

The question we are most frequently asked about our national surveys is a very straightforward one: "How satisfied with work is the average American worker?" Our answer, characterized by the customary lucidity of the social scientist is: "minus two." This answer, while perfectly accurate, is not very helpful to the lay questioner. Our inability to answer otherwise has two sources. First, all the questions in our job satisfaction measure use four- or five-point, fixed-alternative scales, the interval properties of which can certainly be questioned. The second reason involves the complexity of the construction of the satisfaction measure. The five "facet-free" questions using five-point scales are averaged, as are the 34 "facet-specific" questions that use four-point scales. The distributions of each of these means are then converted to standardized z scores. Next, a mean of these two z scores is obtained for each individual, and the resulting mean is finally multiplied by 100 to remove decimals.

The latter arithmetic whoop-de-do, which is not really very complicated as psychological measures go, has certainly not made the initial set of job satisfaction questions lose sight of the psychological reality under investigation. It nevertheless leaves non-scientists completely in the dark, even those accustomed to weather reports that present such equally (or more esoteric) numbers as wind-chill factors and MURC indices. None of this helps the image of social science as a publicly useful discipline. While we can answer some complex questions about job satisfaction based on this measure, superficially simple ones, like "How satisfied with work is the average American worker?" are embarrassingly difficult.

There seems no obvious way around this difficulty. Talk of "educating the public"--or even educating relevant sectors of the public--to the complexities of psychological measurement is largely wishful thinking. This will continue to be true so long as the pollsters persist in leaving the impression that matters are really very simple and so long as posters in the New York subways announce the absolute percentages of Americans with "mental health problems." Occasionally, awkward, but workable compromises can be reached. One such successful, apparent compromise is the widely-used measure of drinking behavior developed by Cahalan, Cisin, and Crossley (1969) for use in population surveys.[2] This measure, far more complex than our own job satisfaction measure, involves very complicated combinations of the frequency of alcohol consumption, the amount consumed at each sitting, and individual variability in amount consumed. These combinations produce a continuous scale of values that the measure's developers break at arbitrary, but reasonable points. To each part of the distribution thus differentiated a publicly meaningful label is assigned: "heavy drinkers," "infrequent drinkers," etc. No one suffers greatly from this assignment. The non-professional users of the measure feel comfortable in having available exact percentages of "heavy drinkers." The professionals use the categories simply as a set of ordered classes, ignore the labels assigned them, and wink knowingly at each other.

But segmenting the continuous distribution of a scale into ordered classes and assigning a label to each can at times result in misunderstandings. To simplify analysis and data presentation, a continuous distribution is often dichotomized at its median and the upper and lower halves of the distribution designated "high" and "low." One of our reports from the 1969 national survey did so. A median split on job satisfaction was performed and a table distinguished "Satisfied" from "Dissatisfied" in order to show the association between job satisfaction and some other variable. Naturally, 50 percent of the workers fell into each category. An early reader of the table, unsophisticated in such "conventional" presentations of data, found the table's marginals very interesting. The marginals "obviously" indicated that half of the work force were dissatisfied with their jobs. The incident is instructive in that it confounded two problems--the problem of communicating data to those unfamiliar with certain conventions and the problem of balancing descriptive and analytic goals in monitoring systems. The user of the data in this example was expecting the survey to provide only descriptive statistics in readily interpretable terms; the marginal distributions were therefore very interesting. Our staff, on the other hand, was concerned with the association between a particular working condition and job satisfaction; the marginal statistics and their verbal designations were only following journal conventions.

Any measure used in a population survey or in a system of social indicators is likely to achieve a premature legitimacy. This is especially likely to be true where the sample size is large and there is an impressive array of institutions or people directing, conducting, or financing the effort. This situation can produce yet another problem of communication. Even a good measure can usually be improved, and its inclusion in a system of social indicators provides an excellent opportunity for such improvement. This raises the question of how much alteration a monitoring system can undergo during its history without losing touch with its users, particularly those who would borrow some of its measures for use in more limited studies. The latter group present an especially difficult problem, since it is particularly tempting to them to appropriate the system's measures. Not only are the measures legitimized, albeit often prematurely, but national norms are available. Given writing and publication lags,

borrowers of a system's measures may find themselves using early versions of measures that have been long since improved upon by the system's designers. For example, some of the difficulty in comparing surveys using the Cahalan, Cisin, and Crossley measure of drinking behavior stems from the survey's uses of various "developing" versions of the final instrument. Clearly, freezing measures for eternity does little for the quality of a monitoring system. Good judgment must be exercised, however, in the timing of changes and improvements, as well as their magnitude. Frequent changes of only minor importance may do a disservice to the system's users.

FOOTNOTES

[1] Problems of Communication

[2] Another example is Belloc, Breslow, and Hochstim's (1971) survey measure of overall physical health.

REFERENCES

Belloc, N., Breslow, L., & Hochstim, J. Measurement of physical health in a general population survey. American Journal of Epidemiology, 1971, 93(5), 328-329.

Cahalan, D., Cisin, I., & Crossley, H. American drinking practices. Rutgers, N.J.: Rutgers Center of Alcohol Studies, 1969.

Quinn, R., Seashore, S., Kahn, R., Mangione, T., Campbell, D. Staines, G., & McCullough, M. Survey of working conditions. Washington, D.C.: U.S. Government Printing Office, 1971. Document 2916-0001.

Quinn, R.P., & Mangione, T.W. The 1969-1970 survey of working conditions: chronicles of an unfinished enterprise. Ann Arbor, Michigan: Survey Research Center, 1974.

Quinn, R.P., & Shepard, L.J. The 1972-73 quality of employment survey: descriptive statistics, with comparison data from the 1969-70 survey of working conditions. Ann Arbor, Michigan: Survey Research Center, 1974.