UJBM, Vol. 6, No. 2, July - December 2007, pp 45-60
ISSN 0975-3311 | https://doi.org/10.12725/ujbm.11.5

METHODOLOGICAL ISSUES IN MANAGEMENT RESEARCH

Sambit Kumar Mishra*

ABSTRACT:

In order to produce information that will be dependably useful to managers, research must be carefully planned, carried out, and interpreted. Good research does not just happen. It is the result of deliberate application of well-tested methods. The aim of this paper is to outline some of those methods and why they are important.

Introduction

Managers want to make good decisions. Any decisions will, by definition, be made on the basis of some presumed information. Even if a decision were to be made by throwing dice, that process would almost certain stem from "information" indicating that no better basis for the decision could be discerned, e.g., that a randomly determined choice would be likely to be better than a decision open to bias. At least to some extent, it is axiomatic that the better the information, the better the decisions.

* *Faculty of Post Graduate Department of Management Studies, Siddaganga Institute of Technology, Tumkur, Karnataka, 572103 E-Mail: sambit_kumar_mishra@rediffmail.com, sambit_mishra_2001@yahoo.co.in*

It is useful to distinguish between data, facts, and information. Data are simply observations, usually in the form of numbers thought to represent some systematic process underlying them, i.e., a process generating the numbers. Data do not mean anything or tell us anything until they are interpreted in some way. Merely to have an observation that on a particular day 43 patients were reported to have received a particular service is not in itself meaningful. Facts are merely data elevated in confidence to a point of suggested certainty. The observation that 43 patients received a service may be registered as a fact if it seems likely that the data are sufficiently trustworthy to justify confidence that the data are correct. Facts come in all varieties: numbers, declarative statements, equations, and so on. But facts are not necessarily information. For example, if one is told that the land area of Bahrain is 231 sq.mi., that might be factual, but it might not constitute information. A useful definition of information is that it is any communication that reduces uncertainty with respect to some decision. That Bahrain is 231 sq. mi. is information only if one is in the position of having to make some decision involving that information, e.g., about whether to invest in a manufacturing enterprise in that country or whether, in a game, to choose Bahrain as the smallest of the Arab states. Even the precise estimate for land area might not be information if one already knows that Bahrain is very small and that is all the precision of knowledge that one requires.

These distinctions are important in thinking about research because not all research findings are necessarily factual, and the facts that are acquired from research efforts are not necessarily information. The reason is that research findings can be—must be—interpreted at different levels. It may be, for example, that in a particular organization more patients were processed by Unit A than by Unit B. That could lead to a conclusion that the management style employed in Unit A results in greater productivity than the management style employed in Unit B. That, in turn, could lead to the conclusion that Management Style A is superior to Management Style B. And so on. Reflection may show, however, that the only undeniable "fact" available is that more patients were <u>recorded</u> as having been processed by Unit A than by Unit B. The apparent better productivity of A could reflect no more than errors in recording patients processed. Or the difference, although real factually, might have happened simply by chance. The difference might be real enough but also so small as to be uninteresting. Or, even if the difference were real and sizable, there is a considerable leap in the conclusion that it should be attributable to differences in management style, and there is even a greater leap in the conclusion that one management style should be considered generally superior to the other.

An Overview

Research, as suggested, may result in specific findings that can be regarded as factual and that may represent information in the sense of reducing uncertainty about decisions. Research may also be supportive of more general theory about the phenomena of interest. Thus, research on a particular management style in a particular context may be supportive of a more general theory of or about management. Even though a particular research project may be rather specific in many of its details, it may still bear on very general principles of organizational management. Conversely, a rather general theory may permit derivation of recommendations concerning specific practices. Although the precision either way, reasoning from specific findings to general theory or from general theory to specific practices, may be limited, the process is likely to result in better decisions that would be made by chance and probably better than would be made overall by intuition.

Theory and research should reflect a continuing interplay of observation and synthesis. It is possibile that theories may sometimes be better than our facts, i.e., the results of specific research. Any one research project is certain to be limited in important ways, at least in terms of its generalizability to other types of persons, to other settings, to different versions of the intervention, and so on. Moreover, any given research project is subject to limitations resulting from a wide range of errors that are inevitable. Measures are not perfect, interventions are not perfectly implemented, individuals differ in their reactions to the same set of conditions, reactions to interventions or other arrangements may vary over the course of a day, or a week, or a month. Thus, if there is a plausible, well-reasoned theory that tells us that a particular intervention ought to decrease the time taken to do some task, but the expected decrease does not happen, it may not necessarily be the case that the theory is wrong. It may well be that the data just were not up to the task of testing the theory.

Causal attribution

The aim of research is, generally, to reveal causal connections between variables, i.e., to demonstrate that one variable is a cause of another variable. When the demonstration is sufficient, then we are, at least potentially, in the position of being in control of the "effect" (caused) variable. We may, if we are able to change the causal variable, e.g., eliminate it, strengthen it, produce a change in the effect variable. Research is sometimes carried out for descriptive purposes, but the results of descriptive research are rarely of great interest in their own right. For example, a manager may wonder whether costs of some activity are higher on some days of the week than on others and may assemble data relevant to that question. That costs are higher, let us say, on Mondays than on other days may be a curious "fact," but it is of value only insofar as it suggests the possibility of another, follow-on study that will explain the higher cost, i.e., that will explain why the cost is higher, what causes the higher cost on Monday. If that can be determined, then the manager may be able to intervene to reduce the excess cost or may understand that the cost, although larger, is warranted because it is attributable to desirable features of the operation of the organization, e.g., more services are provided on Mondays or more serious problems are dealt with on Mondays.

Making correct causal attributions is often far from straightforward. At the heart of the matter is that a causal inference requires dealing with *the counterfactual.* That is, making a causal inference requires at least an implicit answer to the question "What would have happened if the causal event had not occurred?" That may be a relatively simple matter if a causal connection is direct and well understood. If more services are offered, then costs are almost certain to increase. Moreover, it is likely to seem pretty clear that if services are cut, costs will go down. If, however, it is observed that a group of new workers given three days extra training make "only" two errors per shift in the month following training, it may not be so easy to be sure how they would have performed had they not had the training. One cannot both given the training and not give the training. An inference is required.

Descriptive research may be quite useful when it helps to understand some process, particularly if that process is amenable to intervention. It may be important to know, for example, whether supervisors actually read documents that cross their desks for signatures in order to plan for streamlining of the flow of paper through an organization. Such a study might be limited to enquiry by questionnaire, but it might involve interviews or even observations.

Necessity for Comparison Observations

When one thinks about it, no observation is meaningful in and of itself. Observations acquire meaning by the opportunity to compare them to some expected or observed value. Stars are dim only because some other stars are bright. An object is likely to be seen as blue only if it might well have been another color. An observation (or set of observations) made in an organization is interpretable only in relation to some prior expectations. Those expectations might have been derived in any number of ways: logic, common sense, theory, or other observations.

We attribute causality to some intervention in relation to some event because we are capable of imagining that the outcome might have been otherwise, i.e., we invoke the counterfactual possibility. If an assistant tells a manager, "We are having fewer staff complaints from the long-term care unit," and the manager replies, "That's because of the new procedures we instituted last quarter," the manager is implicitly accepting the proposition that had *the procedures not been instituted* the complaints would not have decreased. What might be the justification for that proposition?

Logic? Common sense? It may seem to the manager and the assistant simply logical if things are not going well and a new set of procedures to improve matters is introduced, improvement should occur and that any that does occur should be attributed to the intervention. Things are rarely so simple. Almost always some plausible rival explanations exist, and, if one of them is correct, it may change the decision about what to do in an important way. For example, some studies had shown that the testing of prospective employees for honesty resulted in reductions in episodes of theft or inventory loss. More definitive studies, however, seemed to indicate that the primary effect was the result simply of the demonstration to employees that management cared about theft and was resolved to do something about it. Testing employees for honesty can be expensive and hard on employee morale; by comparison communication of managers' concerns for theft may be inexpensive and accepted as a legitimate managerial function. Complaints from a unit might be reduced for similar reasons, i.e., simply because a manager expressed concern for them and a determination to reduce them.

Theory? Similar to logic but more explicitly reasoned is a theoretical justification for a conclusion. A manager might rely at least to some extent on a theory of management that would predict that a particular type of intervention would have the effect intended and result in an outcome reflected in a decrease in complaints. The legitimacy of the conclusion would depend heavily on the resemblances between the specific intervention and the type represented by the theory and between the outcome supposed by the theory and that realized in the empirical study. A characteristic of a theoretical explanation is that the mechanism for producing the change is explicit. For example, a particular intervention might operate to reduce complaints by improving conditions leading to complaints, by helping staff to be more accepting of deficient conditions, or by improving staff morale so as to decrease the likelihood of complaints.

Other observations? By contrast, a manager might have access to other observations that would support the conclusion that the change could be attributed to the intervention. The manager might note, for example, that complaints had been at a consistently high level for quite some period of time, decreasing only when the intervention was implemented. Or the manager might know that on one or more other units in which no intervention was tried, no decrease in complaints was found. The manager might even be satisfied to know that after the intervention, complaints were no more frequent than reported in other similar organizations in his or her geographic area.

Comparison data

Interpretation of any observations is, obviously, much enhanced by having available data from other sources with which to compare observations at hand. Essentially, the reason is that the comparison data make it possible to make a judgment in relation to the counterfactual notion of what the data would have looked like under other circumstances, e.g., if the intervention had not occurred. If complaints seem high and an intervention is tried, followed by a decrease in complaints, the counterfactual asks what the complain level would have been without the intervention. Answering that question is not always easy, and the answer is not only obvious. Complaints might have been going down anyway; a change in patient mix might have made fewer complaints more likely; new personnel in the unit, independent of the intervention, might have elicited fewer complaints.

Comparison data could be helpful in varying degrees, depending on just what kinds of comparisons were available. Inspection of rate-of-complaint data prior to the intervention might show that the idea that complaints were going down anyway was unlikely.

Those data would not necessarily be helpful in determining whether the change might have been due to changes in personnel or patient mix.

True (randomized) experiments

The best comparison data would result from circumstances in which everything was identical to the conditions in the group exposed to the intervention *except* the occurrence of the intervention itself. The problem is that so often it is difficult, sometimes impossible, to be sure that everything was the same except for the intervention. Two units might, however apparently similar, differ in subtle ways in patient mix, or in personnel, or in some other variables that might be related to the occurrence of complaints. No matter how carefully units were selected and how similar they seemed to be, the possibility exists that the patients in them would differ in some unknown, but important way. By "important" is meant that the difference would be related in some way to an outcome of interest, say the tendency to complain.

The best way to maximize the likelihood that patients on two units will be the same even on unobserved variables is to assign them *randomly* to the two units. By such an assignment process, if a patient with a low tendency toward complaining were assigned to the unit to receive the intervention, probability would guarantee that a similarly low patient would be likely (but not certainly) to be assigned to the other unit. Over a series of observations, chance generally evens things out. The logic of scientific inquiry then is that if one can assume that two groups (units, in this case) were equivalent to begin with, then any later difference between them can be attributed to differences in the way they were treated, i.e., the intervention. The equivalent comparison group answers the question posed by the counterfactual, what things would have been like in the absence of the intervention.

That, in a nutshell, is the essence of the randomized clinical trial, almost universally regarded as the strongest basis for an inference about a causal relationship between two variables. Any other comparison data are considered inherently weaker as a basis for causal inference and, therefore, more likely to be misleading.

Threats to validity of the inference

It is common practice to refer to problems with various comparison data series in terms of "threats to validity," meaning threats to the validity of a causal inference. It will help to review some of the more likely threats and to indicate how they might crop up in specific comparisons.

Selection. If groups, say intervention and comparison groups are assembled in any way other than by random selection, there is at least a possibility, often a near certainty, that differential selection into the groups will have made them different from the beginning, thus making any final difference difficult to interpret. For example, if groups are assembled by asking for volunteers for the intervention, that group may be more enthusiastic, higher in risk-taking, more eager to please, of more of almost anything else than the left-over group of non-volunteers. Thus, if the intervention group turns out better than the comparison group, that difference cannot unequivocally be attributed to the intervention itself. Even if one takes intact groups, e.g., patients in two general medical units of a hospital, one cannot be certain that they do not differ in important ways. Medical studies, for example, have run afoul of the problem that patients selected for one treatment over another tend to be different from patients no so selected. In one well-known study, patients who received surgical treatment had better outcomes than patients not subjected to surgery, but that ultimately was shown to be because patients not selected for surgery were too sick to endure it.

Maturation. Naturally occurring processes may be mistaken for intervention effects under some circumstances. If one unit has very new leadership and another is well established, changes associated with the maturing of the leadership could be mistaken for effects of an intervention. If costs of some process are increasing over time, an intervention might appear to result in increased costs even though it were cost neutral.

History. Sometimes in the middle of a research study, external events will occur that have an impact on the phenomenon being studied. A researcher unaware of a change in accounting procedures in an organization might mistakenly conclude that an intervention decreased costs when, in fact, the change was illusory. Events outside the system may, similarly, have effects that could be taken for the effects of an intervention. It is entirely likely, for example, that the events of Sept. 11 and subsequently may have affected all sorts of responses made within work and treatment facilities. Productivity probably dropped for a while, various complaints probably decreased, and absenteeism increased. Because of the enormous salience of Sept. 11, few investigators would be likely to miss its effects on most data being collected during that time. Similar effects might well be missed, however, if they were less obvious. History is a special threat to studies involving comparisons of data for the same group over time, but if history has more impact on one group than another, it can affect any comparison.

Regression artifact. Observed values of any variable must be considered to be in some part in error. If an observation is toward one extreme or another of a distribution, the probability is that error is in some degree involved in the location of the observation. For example, if on some occasion it is observed that complaints from a treatment unit are much higher than average, the probability is great that the number is high in part because of the odd confluence of factors not likely to persist for long, e.g., a temporary staff shortage, a malcontented patient, an equipment breakdown. Then, with no intervention at all, the number of complaints would fall at the next occasion of measurement (called regression toward the mean). Because interventions are likely to be initiated exactly when complaints are high, the natural effects of regression can easily be mistaken for an intervention effect. Incidentally, and conversely, an intervention that happens to come along when things are noted to be especially good may appear to have a bad effect because regression works in both directions.

Research Designs

As noted earlier, the randomized experiment is generally regarded as the gold standard for research, particularly for supporting causal inferences. So why are randomized experiments not used exclusively? In the first place, they simply cannot be done for many problems because the variable of interest cannot be experimentally controlled, whether for practical or for ethical reasons. We are interested in the effects of education on people's lives, but we cannot "give" education to people who do not want it. Similarly, we are interested in the effects of sanctions on criminal behavior, but we cannot, ethically—or legally—punish some people and not punish others on an experimental basis. Second, randomization may be unacceptable to some persons or groups, i.e., they may be unwilling to be randomized to experimental conditions. If randomization is possible, it is usually preferable as a research option, but often alternatives must be sought.

A wide range of options for research is available, each with specific advantages and deficiencies dependent on the nature of the problem. These options are frequently referred to a quasi-experimental designs. All research designs are, however, simply ways of systematizing observations in such a way as to maximize information concerning the counterfactual proposition. Those research designs likely to be of greatest usefulness are the following:

Nonequivalent comparison group design

The most widely used quasi-experimental design is, undoubtedly, the nonequivalent comparison group design, by which is meant a research scheme that calls for collecting data from an intervention group and a comparison group that do not involve random assignments. Consequently, the groups must be *regarded* as nonequivalent, as potentially having been *selected* in such a way that they are unequal to begin with. Two hospitals differing in management types may be compared, and they may actually be quite equivalent save for management, but there is no way of knowing that, so they must be regarded as nonequivalent. Obviously, the more reason with which one can argue that the groups or organizations are equivalent, the better the case for interpreting any outcome differences between them. That is why when, as is inevitable, plans must be made to compare interventions between two groups or organizations, it is important to select the two (or more) with as much care as possible. One would not opt deliberately to compare two organizations of widely disparate size or from very different communities unless those variables, size or geography, were the exact conditions of interest.

Initial differences between groups or organizations may be dealt with statistically or by developing other, supportive comparison data sets. Statistical allowances for differences is often, but not always, straightforward, depending on the nature and size of the initial differences. Statistical corrections for initial differences also depend for their persuasiveness on the quality and relevance of the data available. If initial measures are available on the variables to be measured as outcomes, allowance for differences is not likely to be terribly controversial, at least if differences are not large and other assumptions are tenable. When, however, it is necessary to use proxy variables, controversy is inevitable. For example, if one has to use a symptom-count as a proxy for severity of illness, the correction for potential initial differences in severity may be questionable.

Interrupted time series

Often data are assembled systematically over time, usually for administrative or clinical rather than for research purposes. If, in the middle of a data series, some event occurs relevant to the processes involved in producing the series, it may be possible to estimate the effect of the event by comparing the values in the series prior to the interruption to those obtained afterwards; hence the label "interrupted time series." Consider, for example, that on a particular date, say the first day of July, the procedures involved in providing aftercare for patients is changed and later that year a manager wonders whether the change may have affected subsequent follow-up visits to an outpatient department. The manager discovers that records of such visits are available by week for the six-month period prior to the change and later. The manager might then plan to continue assembling the data until six months of records are available subsequent to the change. It might then be possible to compare the data series before the change with the data afterwards and reach some conclusions about the effect of the change. Basically, what the manager wants to do is estimate what the data would have looked like without the change (the counterfactual) and compare that estimate with the actual data. The manager's statistician would look for a change in level (intercept) of the two series at the point of the change and for differences in slopes of the series between the two time periods, e.g., whether a general trend in an upward direction prior to the change might reverse to a downward trend after.

The interrupted time series can be a very useful and persuasive research strategy under the right circumstances. Those circumstances begin with the availability of records if the time frame is long (managers will not want to wait for very long to assemble the necessary data) and will include the stability of the series (widely fluctuating values may frustrate attempts to detect changes) and the abruptness of the change (effects of phased-in interventions are difficult to detect). The research strategy does lend itself very readily to many administrative interventions if quick answers are not required or if a long-term perspective is needed.

Observational data with statistical corrections

A very common research strategy, but scarcely a design in the usual sense, is to assemble observational data and attempt to "rule out" rival explanations by logic and statistical means. A recent example is a study based on a large data set including extensive questionnaire responses of 78,000 nurses. The investigators were interested in factors determining the occurrence of breast cancer and discovered that nurses who had worked on night shifts for an extended period of time had an increased risk of breast cancer. That is, they compared breast cancer rates in nurses who did and did not work night shifts and related breast cancer rates to number of years working night shifts. Since many other variables could have been related both to night shift work and to breast cancer risk, they "adjusted" the data for those variables by statistical means. The data for the women who did not work night shifts, when adjusted for "confounding" variables, presumably showed what the data for the other nurses would have looked like if they had not worked night shifts.

An advantage of observational data is that they are usually fairly easy to obtain and are often already available in existing data files. On the other hand, the investigator must often make do with whatever data are available in those files. In the nurses data file, night shift work was defined only as at least three nights per month; a more refined measure would have been desirable. Even if observational data are being collected de novo, the investigator must rely on whatever values for variables happen to show up in the data. If a particular condition of interest is rare, then it simply will not show up very often, although when it does, the investigator can try to make sure that the case gets included in the data set. If a manager is interested in the effectiveness of bilingual supervisors, he or she will be limited to observing the effectiveness of those supervisors actually on the job.

Moreover, observational data are often not likely to be very persuasive when sample sizes are small. Statistical corrections for confounding groups usually require a fairly large number of cases.

Observational data are particularly subject to biases resulting from "data dredging," looking through large quantities of data until something interesting appears to turn up.

Measurement

Good measurement is of absolutely critical importance to good research. It is unfortunate that problems in measurement so often go unrecognized and, if recognized, are treated so lightly. It is impossible to demonstrate effectiveness of any intervention without reliable measurement. Think, for example, of trying to "influence" the a variable with values created by throwing dice. That is exactly what is involved when measures of outcomes are unreliable. Measures that have only modest reliability can be expected to show at best only modest effects of interventions. Poor (unreliable) measures will almost always result in an underestimate of the effects of any intervention.

Prescriptions for good measurement are easily made but not so easily followed. Good measurement begins with good definitions of just what it is to be measured, good definitions of constructs. Multiple measures are highly desirable, especially when measures differ substantially in their structure and likely sources of bias. Measures should, ideally, be nonreactive, i.e., they should not be readily susceptible to self-serving biases and other sources of distortion. They should also be closely related to the phenomena of interest so that values on the measure map clearly onto the underlying variables of true concern. Such prescriptions are obvious; realizing them is difficult.

One measurement problem that is very often overlooked is the assessment of the intervention or, the *independent variable*, as it is often called in the jargon of research methodology. Just as we cannot expect to show much effect on a variable that is poorly measured, so we cannot show much effect from a variable that is only weakly implemented. Researchers too often take the independent variable for granted, and only infrequently do they attempt to quantify it. If an accounting system is to be evaluated, then it is important to know to what extent and how well the system is implemented. If those using it are poorly trained, if they do not like it and, hence, do not use it, or if it involves technical difficulties that result in frequent down-time, a fair evaluation of the system cannot be obtained. Just as it is important in evaluating medications to know how much of a dose was actually received by patients, so it is important to know how much of an intervention was actually achieved in a managerial setting.

Methodological consultation

The design of a good, persuasive research project nearly always requires technical expertise for the many decisions to be made. Managers who wish to become involved in research should be quick to seek methodological—and statistical—consultation. Consideration may also need to be given to the need for consultation on measurement.

Statistical analysis

Data must be subjected to statistical analyses before they are interpretable. Summary statistics such as means, standard deviations, and correlations are easily calculated; in fact, they can be done on simple spread-sheet programs. Their interpretations are not always straightforward, but they usually pose no great problems. Analyses aimed at inferring causes, however, are usually more complex and often are highly complex, requiring specialized knowledge and software. Research outside laboratories, and most of that in laboratories today, is necessarily a multidisciplinary enterprise. A managerial team must include specialists in administration, purchasing, accounting, maintenance, and so on, and a managerial research team must, similarly, include an appropriate mix of specialists.

Research will in every way be improved if statistical expertise is enlisted from the very beginning of a project. Few projects can be brought to a fully satisfactory conclusion by turning things over to a statistician to make sense of after all the other work is done.

Generalizability of research

An issue of persistent concern in the interpretation and use of information derived from research is the extent to which the results of the research may be generalized to other settings, i.e. to settings similar, but not identical, to that in which the research was done. Concerns about generalization often relate to fairly obvious characteristics of settings, e.g., the identity of the organizations or the people associated with them, but legitimate bases for concerns may lie at a deeper level. Superficially, a manufacturing facility and a health care facility seem quite dissimilar, but with respect to the fundamental organizational processes by which they operate, they may be quite similar: leadership requirements, information needs, accountability provisions, quality controls, etc. Researchers can facilitate generalization greatly by being explicit in their assumptions about the settings in which they are working and the nature of the variables underlying the processes they are studying.

To some extent every research study has to be regarded as constituting a "special case of real life." That is, real life is complicated, messy, dynamic; research studies, of necessity, either impose order on or derive order from the situation and the resulting data. Research requires simplification along many dimensions, and that simplification may restrict the generalizability of the interpretations of findings. A great strength of experiments is that they are designed in such a way as to keep everything as simple as possible and as nearly identical as possible between groups. Those restrictions may mean, however, that conditions in the experiment are considerably different from those that obtain in "real life," thus making direct applicability of the findings questionable. For example, in an experiment to test a new accounting system, everyone involved would be likely to be very carefully trained and supervised, well beyond what would achievable in any ordinary, functioning organization. Research requirements often cause investigators to focus on specific outcome measures that may not capture all the consequences of interest. The important issues for generalizability often have to do with specific research arrangements related to the intervention, the circumstances of its implementation, and the measurement of outcomes. With respect to generalizability, those issues may be more important than more obvious things such as the identity of the organization of the types of people working in it or served by it.

In biomedical research a distinction is often made between *efficacy* and *effectiveness* of interventions, the former referring generally to effects of interventions implemented under carefully controlled and sometimes artificial conditions, and the latter to effects obtainable under more-or-less real life conditions. In biomedical research it is regularly necessary to carry out studies related to the transition from the "ideal" conditions to the laboratory to the conditions of the every day world. Undoubtedly, management research may sometimes require the same kinds of transitional studies.

Observational studies offer the advantage that they usually occur in real world settings, and they may not require as many constraints as are necessary to carry out experiments. (It is the case, however, that observational studies may at least require simplifications involved in quantifying inputs and outcomes.) A major shortcoming of observational studies, however, aside from their general messiness (because nothing is really controlled) is that the legitimacy of any causal inferences may be very much in doubt.

In the first place, the direction of causality may be more apparent than real. In a study of the effects of shift work on family relationships, it appeared that shift-workers had poorer family relationships than workers with only daytime jobs. Further investigation suggested, however, the probability that a good many workers choose shift work in order to get away from unpleasant family relationships. When that possibility was taken into account, the effect disappeared. Causal inferences must be developed with great care in observational studies.

Secondly, however, observational studies may be questionable as a basis for causal inferences because of the necessity for assuming that naturally occurring, observed correlations would hold for deliberate interventions. For example, an investigator might determine that a relationship existed between amount of training and productivity of workers. That finding might lead to the supposition that giving more training to workers would increase productivity. That inference would be a big leap, however. It is not only likely, but probable, that workers who have, on their own, acquired more training differ systematically from workers without such training. It cannot at all be assumed that "giving" workers training would have the same effect as their having gotten it.

Conclusions

Thus, observational studies, too, have limitations on their generalizability. All the foregoing leads to the generalization that it is unusual that any important uncertainties can ever be resolved by single studies, whatever their nature, and even multiple studies should have only modest effects on our confidence that we know what we are doing.

Finally, it is important to remember that generalizations can be invoked at different levels: findings, principles, and theory. A particular research finding, let us say that implementing a particular form of feedback to workers about their performance, may generalize to other settings so that those that implement the same feedback get pretty much the same results. At a more general level, it might be that the principal that feedback improves performance is supported, so that a range of feedback arrangements could be counted on. And at a still more general level, the results of a feedback study could be regarded as generally supportive of a "theory of participative management," strengthening it in some small way as a basis from which to derive a wide range of ideas about improving management and performance.

Bibliography

1. Geoff Lancaster, **Research Methods in Management: A concise introduction to research in management and business consultancy,** Elsevier, 2005

2. Bill Taylor, Gautam Sinha, Taposh Ghoshal, **Research Methodology : A Guide for Researchers in Management and Social Sciences,** Prentice-Hall of India Private Limited, New Delhi, 2006

3. Donald R. Cooper, Pamela S. Schindler, **Business Research Methods,** The McGraw-Hill Companies, 2005

4. Alan Bryman, Emma Bell, Business Research Methods, Oxford University Press, 2007

5. Mark N.K. Saunders, Adrian Thornhill, Philip Lewis, **Research Methods for Business Students,** Pearson Education, 2003

6. Cauvery R., **Research Methodology,** S. CHAND, 2005

7. Uma Sekaran, **Research Methods for Business: A Skill Building Approach,** Wiley Text Books, 2005

8. John Adams, Hafiz T A Khan, Robert Raeside, David I White, **Research Methods For Graduate Business And Social Science Students,** Sage Publications, 2007

9. Hair, Joseph F. , Babin, Barry, Money, Arthur H., Samouel, Phillip, **Essentials of Business Research Methods,** John Wiley & Sons, 2005

10. Pervez Ghauri, Kjell Gronhaug, **Research Methods in Business Studies A Practical Guide,** Pearson Education, 2005

11. George A. Marcoulides, **Modern Methods for Business Research,** Lawrence Erlbaum Associates, 1998