EBMgt Using Organizational Facts

Lex Donaldson, University of New South Wales

Abstract

Managers using evidence-based management use data that exist in their organization to draw inferences and make decisions. This process is prone to a number of errors, including the small numbers problem, measurement unreliability, range restriction and confounding. This chapter makes recommendations about how managers can minimize these errors and make better-informed decision. Recommended procedures include statistical corrections and computation methods that make data more interpretable as well as changes in the way data are reported and aggregated within organizations.

Keywords: **Organizational facts**
Organizational data
Meta-analysis
Small numbers problem
Aggregation
Reliability

A long time ago (when I was young!) there was a popular television show called DRAGNET. This was a drama about police in the US. The detective hero, a hard-boiled type, would interrupt weeping witnesses to curtly say: "Just give me the facts, Ma'am". This became a catch phrase: "Just give me the FACTS". Evidence-based management seeks to have managers make judgments based on facts by analyzing facts appropriately to make these judgments better. These goals are difficult to achieve. What seem like facts in many organizations may be misleading or easy to misinterpret. However, methods for reporting and analysing data in academic research can help managers overcome these problems in using organizational data. By identifying critical errors in use of data and ways to avoid them, this chapter seeks to assist management in organizations to become more soundly evidence-based.

This chapter is informed by a theory of organizations I developed: statistico-organizational theory (Donaldson, 2007, 2008, 2010). Other organizational theories draw upon economics (Williamson, 1975) or biology (Hannan & Freeman, 1989) as their foundations. Statistico-organizational theory draws upon statistics and other methodological principles as ideas upon which to build a new theory of management. Statistics and methodological principles provide insights about what errors can occur in academic research. Many are errors inherent in making inferences from numerical data. These same errors occur when managers look at organizational data. Statistics and methodology offer academic researchers guidance about when the errors will be most egregious and how to avoid them. Similarly, statistico-organizational theory uses statistics and methodology to foresee what errors managers may make in drawing inferences from their data.

In particular, statistico-organizational theory uses the principles that small samples lead to random errors, range restrictions and measurement errors lead to under-stated correlations, and confounds introduce spurious correlations. While statistics and methodology identify factors that make these errors larger or smaller, statistico-organizational theory identifies the corresponding situational variables that lead managers to make larger or smaller errors in making inferences from their organization's data. For instance, a smaller organization tends to have smaller numbers of observations and so figures derived from its data tend to contain more random error.

And, importantly, but more tentatively, statistico-organizational theory offers managers advice on how to avoid or minimize making errors in drawing inferences from numerical data. Of course, not all organizational decision-making is, or should be, based solely on numerical data. Some of the evidence used in organizations will be qualitative, but it is the quantitative data that are the focus of statistico-organizational theory and of this chapter.

Modern organizations generate numerical data on many different variables, such as sales, quality, customer satisfaction and so on. These data are held inside the organization, often across units of varying sizes, hierarchical levels, tasks, and metrics. They can be about the organization, e.g., the costs of each of its departments. Or they can be about the organization's environment, e.g., the percentage share the company holds of the market. The data could be generated internally, e.g., the costs of its departments. Data could come from outside, e.g., data from a market research firm about customer perceptions of the firm's products. Once gaining access to these numbers, managers seeking to make inferences from them are prey to certain problems, as predicted by statistico-organizational theory. Statistico-organizational theory focuses on four sources of error in organizational data: small numbers, measurement error, range restriction and confounding. We will discuss each error in turn and identify how managers may reduce them.

Four Sources of Error

Small Numbers

Definition: Smaller samples have more sampling error than do larger samples. When organizational data are based on a small number of observations then any statistic calculated from them can contain random error (Moore et al., 2009). Thus the true value of that statistic varies above or below the value calculated from the data. This error arises whenever a sample is taken from a population or universe. The smaller the sample, the larger the sampling error will tend to be in the sample figure (Moore et al., 2009). Often managers look at a figure not as a mere description but rather as an estimate of some larger population or universe.

In statistics, sampling error has known characteristics. For example, sampling error decreases as the number of observations increases (*N*) (Moore et al., 2009). This knowledge may be used to identify where in the organization management will face a sampling error problem. A figure based on only a small number of observations can contain considerable sampling error, causing the observed figure to vary randomly about the true figure. Given the prevalence of situations in which managers face data based on small numbers of observations, sampling error is likely to exist frequently, rendering figures misleading.

Examples

For instance, an automobile dealership may calculate the average number of sales of automobiles by its salespersons over the last year. But this is in order to forecasts what next year's average will be. Thus the average is being used to convey information not just about this year's sales and this year's customers, but about the future, such as next year's sales and next year's customers, few of whom will be this year's customers. Hence a figure about this year is being used as information about the future (Ehrenberg, 1975). Last year's sales were by last year's salespersons, but next year's may be by a different set of salespersons. Yet the average from last year may be used to try to predict how many extra sales would be obtained by adding, say, three additional salespersons. Thus last year's figure is being used to generalize about the future and to different customers and salespersons.

This is often the case in management because managers look at figures (e.g., averages) from the past to predict the future. The reason is that managers are trying to decide what actions to take in the future. The action-orientation of managers means that they are using the data to estimate a population or universe that is wider than those cases actually studied. Hence managers are often involved in sampling from a population or universe (e.g., all the present and future customers and salespersons) and so their data are samples and so are subject to sampling error.

Problem: The small numbers problem is commonplace. The potential problem of small numbers of observations (*N*) occurs in any organizational characteristic that draws on only a small number of cases. Even a large company may make a small number of products (e.g., heavy industrial machinery) or have just (say) five different information technology systems, so any analysis of these variables will tend to be beset by the problem of small numbers. A small company, if defined as having a small number of employees, will face the small numbers problem in any analysis of its employees, e.g., mean job satisfaction, absenteeism, quitting rates etc. Countries with small populations (e.g., New Zealand) tend to have smaller organizations and so be prone to the small numbers problem in their analyses. Even a large organization in a large country, with many employees, can fall into the small numbers problem if it conducts analyses of, say, employees in one small branch.

Recommendation: The best solution to the problem of small numbers is to avoid it by having larger numbers of observations. Larger organizations have this large number of observations for numbers of employees and related variables. However, for this to hold, the large organization has to aggregate its data rather than leaving them disaggregated in the parts of the organization, such as its branches, departments or divisions. Large organizations that do aggregate their data can calculate statistics fairly free of sampling error. Also, organizations have to aggregate their data over time, such as accumulating sales for the year, or monthly, rather than daily. However, f they have sufficient sample sizes, such as from being a large organization, they can use data that are somewhat temporally disaggregated, such as by month, to spot trends over time. This ability of large organizations to make superior inferences from data is an "inference advantage", giving them a comparative advantage over smaller, rival firms.

A small or medium-sized organization could gain this inference advantage by merging with another organization to form a larger organization. The data from the various business units can be aggregated together in the head office. In this way a parent company could bestow a "parenting advantage" (Goold et al., 1994) of superior inference on its "children" (i.e., subsidiaries). While firms may not merge for this reason, when they do so the inference advantage can be part of the economies of scale frequently sought in mergers.

There are other creative tactics for data combination in smaller settings. A small organization that remains independent can gain access to data based on larger numbers of observations by pooling data with other firms, in a strategic collaboration or though an industry association, government bureau, consulting company or other means. Similarly, organizations in a small country may gain access to data based on larger numbers of observations through international organizations, e.g., UNESCO, or by being subsidiaries of large multinational organizations.

Historically, US organizations have had an inference advantage in that their size could be large by world standards because the US was larger than most other countries. For instance, it could be easier for a US company to spot trends because of its larger size. However, the integration of countries into the European Union and the economic rise of more populous countries such as China and India are reducing the US inference advantage.

If a large number of observations is not available to managers through these various stratagems their data will be infected with random error. Managers and staff analysts need to recognize this explicitly. This may be done by calculating the confidence intervals around the statistic, e.g., mean or correlation coefficient. Hence, instead of the single figure that emerges from the calculation, there is a range of figures, so that the true statistic is acknowledged as lying within a possible range. For instance, sales figures being looked at by managers in an organization could have their confidence intervals routinely supplied by the organization's IT system. Alternatively, managers or their staff analysts could use personal computers to calculate the sampling error around figures of interest to them.

Since managers at higher levels will tend to look at aggregate data, they will tend to make superior data-based inferences than those at lower levels. Looking at such aggregate data require a centralized collection of organizational information. It also requires use of standard definitions for the particular type of data throughout the organization. Furthermore, it implies that there are routine operations of recording and collating data. Moreover, there will need to be organization-wide rules, e.g., that all branches report their week's total sales to head office by 4 pm on Saturday, so that some increased formalization is also required. In these ways, for an organization to benefit from better inferences based on its data, some degree of centralization, standardization and formalization of data collection is needed. Therefore, for upper-level management to have the benefit of aggregate data, incremental increases are required in the levels of the structural variables of centralization, standardization and formalization (Pugh et al., 1968).

In theory, data can be brought together in databanks used by lower-level managers, in order to gain the benefits of aggregate data without centralizing decision-making. However, lower-level managers will tend to focus on “my figures”, that is information from their own unit. Such feelings are prompted by a sense of responsibility and accountability, as well as by personal identification with one’s immediate work setting.

Aggregating data at higher organization levels reduces sampling error, thus eliminating the spurious variation in figures endemic to the data analysed at lower levels in the organization. However, not all the variation is necessarily spurious. There may be some true difference in figures about lower-level sub-units, such as geographic regions. Aggregate data better display these true differences. For instance, the mean sales figures differ by region and these differences are true, rather than being spurious. Such real differences may then be a basis on which the organization delegates decentralized authority down to those subunits. For example, in a national retail chain, winter clothing could differ between California and the Mid-west regions so that those regions are given autonomy to select their own winter clothing. But such decentralization is only be valid if derived from aggregate statistical analyses that show true differences between California and the Mid-west regions. Therefore decentralization of decision-making that reflects real differences between parts of the organization is only valid when made after centralized data analysis, because that aggregate data will have larger numbers of observations and so avoid much of the random error from small numbers of observations. Thus, such centralization should precede decentralization. This is a sounder managerial strategy than assuming *a priori* the existence of differences between parts of an organization that are based on commonsense or myths that may merely appear to be borne out by the spurious differences from sampling error.

Measurement Error

Definition: Whenever something is observed and measured, from profit to intelligence, its score is likely to deviate somewhat from its true score. This deviation is measurement error. This is also known as unreliability (Cohen et al., 2003). Nothing is measured perfectly whether it be accident rates or customer sales. Having some measurement error is unavoidable. Research methodology offers guidance as to where measurement error will be most egregious and how to reduce it.

Measurement error tends to be greater where the variable is a difference score, meaning that it is one variable minus another. In psychology this is well understood to occur when measuring, say, the difference between desired pay and actual pay. Psychometrically, the difference score has lower reliability than the two variables that compose it (e.g., desired pay and actual pay) when those two variables are positively correlated (Johns, 1981). The higher the positive correlation between them, the lower the reliability of the difference score (Johns, 1981). It is possible for the two variables to be highly reliably measured and yet for their difference score to be low in reliability. In other words, two variables that are measured with very little error can be the origins of a variable with great measurement error that is formed by simply taking their difference.

Examples: Profit is widely used to measure the financial performance of companies and often their constituent business units and subsidiaries. Yet profit is a difference score: sales minus costs. This allows profit to be unreliable. There can be much measurement error in profit even if sales and costs are both measured with very little error.

For instance, the Walt Disney Company reports financial data for 2002 for its four business segments. Sales and costs are highly, positively correlated, .981, which implies that sales and costs must be measured quite reliably (otherwise their correlation would be attenuated to be much less than 1.0). The most conservative estimate of the reliability of sales and costs is that they are both .981 (for details of calculations see Donaldson, 2010, pp. 109-110). Despite this very high reliability of sales and costs, the reliability of profit can be as low as .22 (see Donaldson, 2010, pp. 107-110). Thus, although little measurement error may exist in sales and cost measures, there can be much measurement error in profit. The average error in the profits of the business segments could be 91 per cent. For instance, the reported profit of the Media Networks business segment of $986 million could truly be only $229 million, an over-statement error of $757 million (Donaldson, 2010, p. 90).

Problem: Measurement error, that is, unreliability, of a variable reduces its potential correlation with another, a condition referred to as attenuation (Hunter and Schmidt, 2004). When two variables are correlated, the lower their reliability, the more that the observed correlation understates their true correlation. For instance, if training of business unit employees truly correlates .3 with business unit profit, but profit has reliability of only .22, then the observed correlation is only .14. Such a small correlation could easily be dismissed by managers as "too small to bother with," or "non-significant." Thus the policy implication drawn could be that training is ineffectual and training budgets should be cut or not increased, whereas the sounder policy would be to maintain or increase training because it is effective.

Recommendation: Measurement error can be reduced by using multiple indicators and combining their scores. There could be multiple measures of financial performance such as profit, sales growth etc. Performance could be broadened, so that to financial performance variables are added measures of consumer satisfaction, quality, innovativeness and other aspects of organizational performance, i.e., a balanced scorecard (Kaplan & Norton, 1996).

Consider the correlation between an independent variable (that is, the presumed cause), say business unit employee training, and a dependent variable (the presumed effect), say, business unit profit. The problem of the unreliability of profit could be avoided in the following way. Rather than conducting a single regression with profit as the dependent variable, two regressions would be conducted: a regression with sales as the dependent variable and a regression with costs as the dependent variable. If training correlates positively with profit then training will have a positive relationship with sales in the first regression and training will have a negative relationship with sales in the second regression. Further, if the beneficial effects of training on sales and costs come from the beneficial effects of training on profit, then the sizes of the regression coefficients of sales and costs will be equal. In this case, the standardized regression coefficients of sales and costs will be the correlation between training and profit. For instance, if the regression of sales on training is +.4 and the regression of costs on training is –.4 then the correlation between training and profit is +.4.

This correlation is attenuated far less than if only profit were used in the regression, because sales and costs are usually measured far more reliably than profit because of its being a difference score. To return to the example above, if the true training-profit correlation is +.3, if sales is measured with reliability of .981 then the observed correlation is .297. Hence, the attenuation (i.e., reduction) in correlation due to the error in measuring sales is only .003, which is trivial. The same applies to costs. Thus using sales and costs to estimate effects on profits allows use of variables that avoid most of the measurement error of profit and the attenuation of correlation that comes from it. The results more accurately answer the question of the true correlation between an independent variable and profit, providing a higher correlation. Thereby, managers can appreciate the full merit of a practice, such as employee training, and make more optimal decisions about allocating resources and support to those practices.

Where, nevertheless, the manager's analyses include a correlation involving profit (or other variable with considerable measurement error) a correction formula can be applied (Hunter and Schmidt, 2004, p. 97):

$$r_t = r_o / \sqrt{r_{xx}}$$

Where r_t is the true correlation, r_o is the observed correlation and r_{xx} is the reliability of profit. For instance, returning to the example above, if the observed correlation between the training of business unit employees and business unit profit is .14, and the reliability of profit is only .22, then the true correlation is .14 divided by .47 (the square root of .22), which is .3. Thus the correlation can be corrected to give an estimate of the true correlation without the measurement error. A method to estimate the reliability of profit in a data set is given in Donaldson (2010, p. 107). The reliability of the variable that is being correlated with profit is also needed. Then the correction formula from Hunter and Schmidt (2004, p. 97) can be applied to turn the observed into the estimated true correlation. A staff analyst could readily perform these steps. The resulting higher correlation would give the manager a truer sense of the amount of effect training has on profit. The correction formula for providing a truer estimate of a correlation could be utilized on any correlation for which the reliability of one or both of the variables is not high. The method for obtaining the reliability of a difference score can be applied to any difference score variable (Donaldson, 2010, p. 219), e.g., pay disparity defined as desired pay less actual pay.

We have focused here on the measurement error in profit that arises from it being a difference score. Often financial performance is a ratio, but the ratio involves profit, e.g., the ratio of profit-to-sales, so that the same problem of potentially high measurement error exists, because profit is still a difference score (sales minus costs). Also, many other financial performance variables are difference scores and so they are prey to the same high degree of measurement error. For instance, sales growth is the difference between sales in one time period and sales in the preceding time period. More generally, any growth rate is the difference between a variable in one time period and that variable in the preceding time period (Donaldson, 2010, p. 97), e.g., growth in capacity, so the problem of difference scores producing unreliability can occur in more than just financial performance variables.

Range Restriction

Definition: Range restriction occurs when a variable in the data has less than the range it possesses in the population or universe (Hunter & Schmidt, 2004). Such range restriction reduces any potential correlation between that variable and another. Range restriction can occur unwittingly through limiting the range in the data that are gathered. For example, if data are gathered in a factory that has all low-skill employees, then any positive effect of job skill on job satisfaction will be severely attenuated and thus understated. Managers interested in such a relationship would do well to ensure that they draw data from across their organization to obtain the true range of skill among its employees, which will yield a higher, truer correlation.

Example: Range restriction can occur in an organization though organizational learning and adaptation. Suppose that the Acme Corporation has poor safety. It starts training employees in safety, conducting courses in some of its branches and not others. At this time safety training is found to correlate considerably with safety (lack of accidents). Therefore, Acme mandates that all its employees receive safety training. Now the correlation between training and safety decreases to about zero. Managers draw the lesson that safety training no longer works and discontinue it. In reality, the safety training still works very well. All employees score highly on amount of safety training leading to range restriction in that variable. This in turn attenuates the correlation between training and safety causing it to decrease to near zero. Managers need to ensure that any study using correlations or other methods of association has enough variance in the variables to reveal their true correlation.

Lack of variation can occur within an organization through its sub-units following a standard approach through, say, training or mimicry (DiMaggio & Powell, 1983). The managers of the sub-units may wish to vary from each other but lack the power and autonomous decision-making authority to do so. Again, organizations may be required to follow a standard template by powerful outside organizations, such as funding agencies or governments (DiMaggio & Powell, 1983). Or organizations may be required to follow a standard template by professional service firms, such as auditors (DiMaggio & Powell, 1983). Any of these mechanisms will reduce variation and restrict range leading to attenuation of correlations. Managers may need to look at the scientific evidence to better observe effects of these practices or to aggregated data from industries or larger firms besides their own.

Range restriction can occur in an organization though differential survival. Suppose that a government department for small business conducts a study of success factors in small retail shops in a shopping mall. It studies ten shops and finds that over-stocking has very little correlation with success measured as profit. However, over the past ten years fifty such shops have opened and forty have failed (Aldrich, 1979) so that they are not in the study. Across the fifty shops there is a strong negative correlation between over-stocking and profit, so that over-stocking is a major cause of failure in these shops. Yet to see that a manager needs to look at data that captures the variation that exists over time but is missing from cross-sectional data. Hence, the governmental officials and their staff need to conduct longitudinal studies that include those businesses that exit from the population.

Problem: The above essentially explains that a true correlation can be reduced towards zero. Denrell (2005) has cogently argued that a true negative correlation can falsely appear positive through what is here termed range restriction. Projects vary in risk, with low risk projects having outcomes that vary only from slightly positive to slightly negative, whereas high risk projects have outcomes that vary considerably from highly positive to highly negative. But the projects with negative outcomes are discontinued, so that a cross-sectional study finds only the survivors, which have positive outcomes. Among the survivors, the average outcome is higher for the more risky than for the less risky projects. Hence the study concludes that risk leads to higher outcomes. But this is false, because if the discontinued projects are included the relationship is much less positive and can actually be negative. Once again, managers need to conduct inquiries that are not limited to survivors.

There is another possible problem about range and that is range extension. Range extension is the opposite of range restriction. Range extension occurs when a variable in a data-set has more than the range it possesses in the population or universe. Such range restriction over-states any correlation between that variable and another. Range restriction can occur unwittingly through taking just extreme cases. For example, suppose management want to investigate the relationship between personality and performance of its salespersons. Out of fifty salespersons, they took the top three performances and the bottom three and correlated their extraversion scores with their performances. The correlation is +.6 leading the managers to conclude that "success as a salesperson is really mostly due to having the right personality", so they select on extraversion.

Taking only the highest and lowest performers, however, produces range extension, so that the correlation is exaggeratedly high. In a correlation coefficient, cases distant from the mean have more leverage and so produce higher correlations than cases nearer the mean. Hence, excluded the middling cases produces an exaggeratedly high correlation. Returning to the example, the forty-four middling performing salespersons would have had a lower correlation, e.g., +.2. If all the fifty salespersons had been used in the correlation, the true, lower, true correlation of (say) + .3 would have been found. If the management had seen this true correlation they might have correctly concluded that extraversion was a factor in performance, but only *one* of a number. This might have led them to seek other factors that affect the performance of salespersons, such as mental ability -- so that it is not just “having the right personality”.

Recommendations: Range extension can be avoided by managers by not just taking extreme cases for analysis but rather studying all the cases or a representative sample. In practice managers need to avoid just focusing solely on the “winners and losers” when they are seeking to find associations that will inform them about the true effects of drivers of outcomes.

The best way to avoid both range restriction and range extension is to use a set of observations that captures the actual range in the relevant population. Capturing the range in one variable will usually also capture the range in any other variable with which it is being correlated, so that range problems are avoided simultaneously on both the independent and dependent variables. Hence, getting the right data to use in the analysis is the preferred approach.

Nevertheless, if the right data aren’t available and yet a correlation has been obtained from data that are afflicted by either range restriction or range extension, a correction formula may be applied to obtain an estimate of the truer correlation (Hunter & Schmidt, 2004, pp. 107-8):

$$r_p = Ur_d / \sqrt{(U^2 - 1)r_d^2 + 1}$$

Where r_p is the correlation in the population (that is, the organization), U is the ratio of the standard deviation of the population divided by the standard deviation of the data, and r_d is the correlation in the data. For instance, returning to the Acme Corporation example, suppose that standard deviation in the data was half that in Acme overall and that the correlation between training and safety in Acme's data was only .2, then the population correlation would be .38. Thus the considerable restriction in range of the data relative to the population leads to the correlation in the data under-stating the population correlation by about half. Yet by using the formula the population correlation can readily be found. The correlation for Acme overall gives a truer picture to the manager of the effect that increasing training will have on safety in Acme. This formula will increase the value of a correlation that has range restriction and decrease a correlation that has range extension.

The correction formula just given could readily be used by a staff analyst. The key would be to determine the range of the relevant population. Acme conducted a study of the relationship between training and safety in only one plant in the organization. This plant had less range of training than Acme as a whole, so that the observed training-safety correlation under-states the true relationship. The range of training may be known in organizational records and so can used in the formula above to correct the correlation in the data so that it is increased up to its truer value. Alternately, the range of the safety variable, if it is known, can be used to make the correction. This corrected correlation gives the managers a truer picture of the strength of the association between training and safety in their organization.

Confounds

Definitions: A confound occurs when the true relationship between two variables is obscured because of some contaminating effect. Usually in social science this is considered to be due to the influence of a third variable or some unmeasured factor common to both variables. We will discuss this case before turning to two other ways in which confounding can occur in numerical analyses: confounding due to definition and confounding due to reverse causality.

Problem: Suppose training all employees has a beneficial effect on organizational effectiveness of correlation +.4. This could be obscured by confounding by spurious correlations such that the observed correlation between employee training and organizational effectiveness was zero, so that the training budget was not increased, or cut. This confounding could occur in three ways.

The first way is a confound due to a third variable correlated with both employee training and organizational effectiveness. Employee mental ability also positively affects organizational effectiveness, but those divisions recruiting persons high on mental ability do not feel it necessary to train them, whereas those divisions recruiting persons low on mental ability do feel it necessary to train them. Therefore there is a negative correlation between employee mental ability and employee training. Given that employee mental ability is positively correlated with organizational effectiveness, there is a spurious negative correlation between employee training and organizational effectiveness that is due to employee mental ability. This masks the relationship between employee training and organizational effectiveness and could reduce it from +.4 to zero.

The second way to have a confound is due to definition. If organizational effectiveness is measured by profit, then by its definition, profit is positively correlated with sales and negatively correlated with costs. Suppose that sales is negatively correlated with employee training because managers of high-sales divisions judge that training is not needed, then there is a spurious negative correlation between employee training and profit. Or suppose that costs are positively correlated with employee training because managers of high-cost divisions judge that employee training *is* needed, then, again, there is a spurious negative correlation between employee training and profit. These spurious negative correlations between employee training and profit mask the relationship between employee training and organizational effectiveness and could reduce it from +.4 to zero.

The third way to have a confound is due to reverse causality. For a firm, having a structure that fits its strategy positively affects firm performance. But firms tend not to adopt a fitting structure until their performance is poor, so performance negatively affects fit. Therefore, in a cross-sectional correlation the positive correlation of fit with performance is masked by the negative correlation of performance with fit. The result could be an observed correlation of zero.

Confounding by a Third Variable

The true effect of *X* on *Y* can be obscured by a third variable, *Z*, that is correlated with both *X* and *Y*. *Z* leads the observed correlation between *X* and *Y* to be greater or smaller than its true value. In social science this is often dealt with by including *Z* into the analysis and conducting a multivariate analysis which gives the relation between *X* and *Y*, controlling for *Z*. Managers can also do this in their analyses inside their organizations.

Example: To continue the earlier example, having discovered that across all fifty salespersons extraversion correlates +.3 with performance, when they go on to study the effect of mental ability on performance they could control for extraversion by including it in a multivariate analysis. If extraversion happens to correlate with mental ability, given its correlation with performance, extraversion will introduce a spurious correlation into the observed correlation between mental ability and performance. Including extraversion in the multivariate analysis essentially removes this spurious effect rendering the observed correlation the same as the true correlation (ignoring sampling error etc.).

Recommendation: There could be numerous control variables included into the multivariate analysis to control for them. However, to control a confounding variable by inclusion into a multivariate analysis, the analyst has to know that a variable is a confound. Thus, the analyst has to identify all the variables that confound a relationship. This can be a tall order. In practice many multivariate analyses explain much less than all the variance in the dependent variable (e.g., the performance of salespersons). This leaves open the possibility that there are some other variables that are correlated with both the dependent and independent variables and so confound the focal relationship.

Some social scientists who are worried about the possibility of such unidentified confounds use experimental methods with a control group. In the experiment the subjects are isolated from the effects of all other variables except the independent variable. The corresponding changes in the dependent variable are the true effects of the independent variable on the dependent variable.

However, in organizational research it is hard to prevent all other variables from affecting the experimental subjects, in part because, once again, not all the other variables may be known. Therefore, experimentalists often make use of having a control group, which registers the effects on the experimental subjects of all other variables that affect the dependent variable. The effect in the control group can be measured and subtracted from the effect in the experimental group, yielding the true effect of the independent variable on the dependent variable. This is very attractive approach because the control group registers the effects of *all* the other variables that affect the dependent variable without the analyst needing to identify *what those other variables are*.

Ideally a control group needs to be identical to the experimental group in every respect, including type of person. This may be facilitated in university laboratories by random assignment of subjects to the experimental and control groups. Experiments in organizations may adopt the somewhat weaker approach of trying to have the control group be as similar as possible. To do so, they may use formal groupings of the organization as control groups, such as members of two plants that are manufacturing the same product, so that they have similar technology, routines and employee skills etc. The treatment is then introduced into one plant, the experimental group, but not the other, the control group.

However, having two near-identical organizational sub-units would be unlikely in a small organization (i.e., with few employees) and is more likely in a large organization. The reason is that size positively affects structural differentiation (Blau and Schoneherr, 1971). Similarly, formal structuring of activity through rules that are in common between the two plants provides standardization between the two plants, but is more likely in large than small organizations, because size positively affects bureaucratic standardization (Pugh, et al. 1969). Again, measurement of dependent variables in organizational experiments often relies upon the formal control systems to measure productivity, absenteeism and costs etc., but these are more likely in large organization (Pugh, et al. 1969). Therefore, while organizational experiments are an attractive idea they may be infeasible in small and medium sized organizations, and feasible only in large organizations (see Donaldson, 2010, pp. 171-3). Thus, managers in many smaller organizations may find that they cannot use experiments in their organizations to control for confounds.

Although this seems to leave organizational experiments as a viable option for managers in large organizations, this could be beguiling. In a large organization having a sub-unit use a new approach (the experimental group) while another sub-unit uses the old approach (the control group) is facilitated where opinion is divided among its managers about the efficacy of the new approach. This leads to the philosophy of "trying a little as an 'experiment'". The new approach may be championed only by staff personnel (e.g., Human Resources), or by a local manager who allows his or her plant to be used as the site of the experiment (Rodgers & Hunter, 1991). Top managers may be skeptical or hostile about the new approach.

Yet, in research on the effectiveness of Management-by-Objectives, top management support has been shown to be a strong moderator interacting with the management technique to impact its effectiveness. High top management support produced high performance outcomes whereas low top management support produced only low performance outcomes (Rodgers & Hunter, 1991). Moreover, it was in organizations with low top management support that experiments could be run because there were simultaneously organizational sub-units using the new technique (the experimental groups) alongside of organizational sub-units not using the new technique (the control groups). In contrast, if there is high top management support for a technique it is likely to have been adopted organization-wide so that there are no organizational sub-units that are not using it and hence no control group, so that an organizational experiment with control groups cannot be run in that organization.

Thus, organizational experiments that utilize control groups can unwittingly introduce conservative bias (Donaldson, 2010, pp. 173-6). Managers looking at the only modest benefits shown by the experiment may wrongly conclude that the technique is ineffective in their organization. Initial skepticism may be reinforced. Initial skepticism leads to "let's try an experiment to see if it works in our organization", which concludes the initial skepticism was justified. Thus, there can be a self-fulfilling prophecy, in that the initial skepticism leads to running an organizational experiment, which inherently will tend to show only modest benefits, confirming the initial skepticism – so that the innovation is never adopted organization-wide Thus, organizational experiments unwittingly lead to dysfunctional conservatism in that innovations that would be highly beneficial if adopted organization-wide and implemented in a full-bodied way, are not.

Aggregation of data sets can eliminate confounding. The confounding variable may introduce a spurious correlation that is positive in one data-set, but negative in another data-set. If the two data-sets are aggregated together then the two spurious correlations will off-set each other so that the net confounding effect may move towards zero. If the confounding variable is genuinely independent of the independent variable then there will be no correlation between them. Then, any correlation between the confounding variable and independent variable in a data-set will be wholly spurious and specific to that data-set. Combining such data-sets will tend towards a zero correlation between the confounding and independent variables, thereby eliminating the confounding element from the combined data-set (Donaldson, 2010, pp. 177-197).

Moreover, random variation from data-set to data-set in the confounding spurious correlation is eliminated by the aggregation of data-sets, so that false inferences are avoided about the focal relationship being stronger in some situations than others. Aggregation of data-sets is essentially similar to averaging the results of data-sets. The aggregate figure provides a better estimate of the general effect across the multiple situations than do the disaggregate figures that display a lot of variation, some of which is actually spurious due to variations in idiosyncratic confounding from situation to situation. Where there is a confound existing in a population, aggregation will not eliminate it, but will still eliminate the specious variations across situations.

Similarly, in an organization, the data from each sub-unit are prey to spurious effects from confounding that are idiosyncratic to that data-set. As data go up the organizational hierarchy they become aggregated. For example, the sales of salespersons become sales of the branch. In turn branch sales are aggregated into area sales, then territory sales, regional sales and company sales. At every level in the hierarchy the combination of data can lead to the offsetting of spurious confounds in one direction (e.g., positive) by spurious confounds in the opposite direction (e.g., negative). Therefore, the more aggregate data typically found at higher levels in organization will tend to be less infected by confounding. Hence organizational data at higher levels in the hierarchy will give a truer picture than data at lower levels.

If the confound is real, so that it exists even in aggregate data, it may still be small (e.g., a correlation of +.1) so that the aggregate figure is substantially valid. Only if the confound is high relative to the relationship between the independent and the dependent variable will the confounding still be substantial after data aggregation. For the true confound to be high relative to that relationship, there must be correlations between both those variables and the confound that are both greater than the relationship between the independent and the dependent variable.

Example: For example, if X and Y correlate +.5, then, for the confound Z to completely obscure the XY relationship, the correlations between X and Z, and Y and Z must both be +.7, because the confound is the product of their correlations, .49 (= .7 x .7). Only then will the confounded correlation between X and Y appear, wrongly, to be near-zero, +.01 (= .50 - .49). This kind of strong confounding is unlikely, because the confounding variable (Z) rarely will be strongly correlated with *both* the dependent and the independent variables. Partial confounding is more feasible.

Continuing the example, if correlations between *X* and *Z*, and *Y* and *Z* were both +.4, then the confounding of *XY* by *Z* would be .16 (= .4 x .4), so that a true *XY* correlation of +.5 would appear to be lesser at +.34 (.5 - .16). Managerially, partial confounding only matters if it is strong enough to cause managers to make the wrong decision, e.g., not implement a new technique because it wrongly appears not to be beneficial.

Aggregation of data-sets reduces confounding more, the greater the number of data-sets aggregated, but at a decreasing rate (see Donaldson, 2010, p. 190). Thus, the big decreases in confounding come from aggregating the first few data-sets, so that even modest aggregation of data can help considerably to reduce the confounding problem. Therefore, managers would be well advised to use aggregation of data in their organization to reduce confounding.

This benefit from reducing confounding is independent from the benefit of reducing sampling error (Donaldson, 2010, p. 195). Thus, data aggregation benefits inference making in two ways: reducing confounding and reducing random error from sampling. This gives managers two reasons why they should aggregate their organizational data.

Confounding Due to Definition of a Variable

Definition of this type of confounding: Some variables are difference scores, meaning that the variable is defined as being the difference between the level of one variable and the level of some other variable. For example, profit is sales *less* costs. These definitional connections between the difference score variable and the variables from which it is composed lead to associations among them.

Example: For instance, by definition, profit is positively correlated with sales and negatively correlated with costs. Therefore, if sales happen to be positively correlated with some other variable, there will be a spurious positive correlation between that variable and profit. This would confound any true relationship between that variable and profit. Similarly, a confound can arise from costs, and it could reinforce the confound from sales.

Recommendation: The way to control for these confounds is to enter both sales and costs into any analysis of the relationship between some variable and profit. Thus, a multivariate analysis is required in which sales and costs are present as control variables. A manager interested in the causes of profit could have a staff analyst use a personal computer to conduct such a multivariate analysis.

Confounding Due to Reverse Causality

Definition: When the presumed cause is in actuality the result of the presumed effect, reverse causality exists. In this case, observed relationships can be misinterpreted.

Example: Managers are perennially interested in their organization's performance and that of their rivals. They conduct analyses to pinpoint the causes of organizational performance. However, performance levels often feed back to produce changes in the organization. In particular low performance can be the crisis that triggers changes in the organization's structure, strategy or leadership

(Chandler, 1962). Thus although some organizational characteristic, such as organizational structure, may positively affect organizational performance, organizational performance can also negatively affect that same organizational characteristic (e.g., organizational structure). Hence a true positive effect on organizational performance can be masked by a negative effect of organizational performance. If the positive effect on performance was of the same degree of correlation as the negative effect from performance, then the observed correlation would appear to be zero. This would give the false message that the organizational characteristic (e.g., structure) has no effect on performance when it actually has a beneficial effect. A manager might conclude that it is best not to increase the level of that organizational characteristic when actually such an increase would be beneficial for the organization. Similarly, successful firms may invest more in R and D, confounding the effect of R and D on firm performance.

Recommendation: The solution is to conduct a study over time so that organizational performance is measured after the organizational characteristic (e.g., structure). In some academic studies an appropriate time lag between the organizational characteristic and organizational performance is two years (Donaldson, 1987; Rogers, 2006). This means waiting two years to see the performance effects of some organizational characteristic. Some managers may resist such delay, being impatient to know the results of the analysis so that they can take action. Delay may seem irresponsible to the manager. By the time the lagged effect of the organizational characteristic on organizational performance is known, it may appear to the manager to be "merely academic", because it addresses the past rather than the present. However, managers would be well advised to curb any such natural impatience and use studies with substantial time lags to avoid confounding by reverse causality, and the false lessons it teaches.

Conclusions

Evidence-based approach entails using data to inform decisions. However, data contain errors that can mislead anyone looking at them. Numerical data -- which have been the focus of this chapter -- risk particular kinds of errors. Small numbers of observations leads to random error around the true value. This is best avoided by aggregating the observations in an organization, which involves some centralization, standardization and formalization of the structure that collates and analyses the data. For a small organization, its managers can seek to obtain data that aggregate across numerous organizations such as from a parent company, market research firms, or a governmental bureau.

Measurement error occurs in organizational variables, especially such as profit, even if there is little measurement error in the sales and costs figures from which profit is derived. A solution is for staff analysts interested in profit to conduct separate analyses for sales and costs, and then combine their results. Also, the under-stated profit correlations can be corrected upwards by use of the formula given in this chapter.

Range restriction leads to observed correlations that under-state the true correlations. The solution is for staff analysts to obtain the full range of the variables of interest, or to correct the correlations upwards in value using the formula given in this chapter.

How best to deal with confoundings depends on the kind of confounding in the data. Confounding by a third variable is best dealt with by aggregating the organizational data. (Thus, data aggregation has the twin benefits of reducing the error from this confounding and from small numbers of observations.) Confounding due to the definition of a variable can be avoided by including in the analysis the constituent variables, e.g., sales and costs in an analysis of profit. Confounding due to reverse causality can be reduced by measuring the effects *after* the causes.

In these ways, these major sources of error in numerical data can be avoided by managers and their staffs. Table 13.1 gives a summary of the errors and the recommendations for reducing them. Other errors made by managers who are seeking to make evidence-based decisions may be reduced by the use of the techniques of qualitative academic social research, and scholars other than the present author would be better able to identify these.

Hence, managers seeking to best use their organization's data to make evidence-based decisions can adopt ways to help attain this goal through structural arrangements or use of analytic approaches. Aggregation of data will minimise random errors in the figures inferred from the data and such aggregation is facilitated by centralizing, standardizing and formalizing the data handling procedures in the organization, to produce a big N from which to compute the statistics of interest to managers, e.g., sales trends. This data aggregation will also reduce confounds and their misleading effects. Problems introduced by the use of profit can be ameliorated by using sales and costs in regressions conducted to identify profit drivers. Under-stated correlations involving profit can also be corrected upwards to their true value by use of a formula. And under-stated correlations due to range restriction may be corrected upwards by another formula -- though the more enlightened analysts will avoid the problem by capturing the full range of a variable in the organization in their studies.

Through this combination of structural and analytic approaches managers can extract the most information value from the data their organizations possess. In this way, managers may make sound inferences and our organizations will better attain what may be termed "inference security." Such organizations will realize their potential inference advantage, so that the benefits of evidence-based management will come to be more fully realized.

References

Aldrich, H. E. (1979). *Organizations and environments*. Englewood Cliffs, NJ: PrenticeHall.

Blau, P. M., & Schoenherr, R. A. (1971). *The structure of organizations*. New York, NY: Basic Books.

Chandler, A. D., Jr. (1962). *Strategy and structure: Chapters in the history of the industrial enterprise*. Cambridge, MA: MIT Press.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum.

Denrell, J. (2005). Selection bias and the perils of benchmarking. *Harvard Business Review,* April, 114–119.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited:Institutional isomorphism and collective rationality in organization fields. *American Sociological Review 48(2)* (April), 147–160.

Donaldson, L. (1987). Strategy and structural adjustment to regain fit and performance: In defence of contingency theory. *Journal of Management Studies 24(1)* (January), 1–24.

Donaldson, L. (2007). Statistico-organizational theory: A new theory of organizations, in G. M. Schwarz, S. Clegg, T. G. Cummings, L. Donaldson & J. B. Miner: We see dead people? The State of Organization Science, *Journal of Management Inquiry*, *16(4)* pp. 300–317.

Donaldson, L. (2008). Statistico-organizational theory: Creating organizational management theory from methodological principles. In D. Barry & H. Hensen (Eds.), *The SAGE handbook of new approaches to management and organization* (pp. 135–145). London, England: SAGE Publications.

Donaldson, L. (2010) *The meta-analytic organization: Introducing statistico-organizational theory*. Armonk, NY: M.E. Sharpe, Inc.

Ehrenberg, A. S.C. (1975). *Data reduction: Analysing and interpreting statistical data*. London, England: John Wiley.

Goold, M., Campbell, A. & Alexander, M. (1994), *Corporate-level strategy: Creating value in the multibusiness company*. New York, NY: John Wiley and Sons.

Hannan, M. T., & Freeman, J. (1989). *Organizational ecology*. Cambridge, MA: Harvard University Press.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. 2nd ed. Thousand Oaks, CA: Sage.

Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance 27(3)* (June), 443–463.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: Translating strategy into action*. Boston, MA: Harvard Business School Press.

Moore, D. S., McCabe, G. P., Duckworth, W. M., & Alwan, L. C. (2009). *The practice of business statistics*. 2nd ed. New York, NY: W.H. Freeman.

Pugh, D. S., Hickson, D. J., Hinings, C. Robert & Turner, C. (1968). Dimensions of organization structure. *Administrative Science Quarterly 13(1)* (June), 65–105.

Pugh, D. S., Hickson, D. J., Hinings, C. R. & Turner, C. (1969). The context of organization structures. *Administrative Science Quarterly 14(1)* (March), 91–114.

Rodgers, R. C., & Hunter, J. E. (1991). Impact of management by objectives on organizational productivity. *Journal of Applied Psychology 76(2)* (April), 322–336.

Rogers, M. (2006). Contingent corporate governance: A challenge to universal theories of board structure. PhD thesis, University of New South Wales, Sydney, Australia.

Williamson, O. E. (1975). *Markets and hierarchies: Analysis and antitrust implications*. New York, NY: Free Press.