THE RIGHT DECISION

Evidence-based Decision Making for Fire Service Professionals

Paul S. Maxim, Len Garis and Darryl Plecas

THE RIGHT DECISION

Evidence-based Decision Making for Fire Service Professionals

Paul S. Maxim, Len Garis and Darryl Plecas

Cover photo: Brian Biggs
Creative Commons Licence
The clock tower at 484 Yonge Street, Toronto
(formerly the tower of Fire Station #3)

About the Authors

Paul Maxim, Professor, Wilfrid Laurier University

Paul Maxim obtained his MA in criminology at the University of Ottawa and his PhD in sociology at the University of Pennsylvania where he specialized in criminology and research methods. He is currently a professor in the Department of Economics and the Balsillie School of International Affairs at Wilfrid Laurier University in Waterloo, Ontario. His primary areas of research interest are population and labour economics.

Len Garis, Adjunct Professor, University of the Fraser Valley

Fire Chief for the City of Surrey, British Columbia, Len Garis and is an Adjunct Professor in the School of Criminology and Criminal Justice at the University of the Fraser Valley, a member of the Affiliated Research Faculty at John Jay College of Criminal Justice in New York, and a faculty member of the Institute of Canadian Urban Research Studies at Simon Fraser University. His focus is on addressing public safety challenges through evidence-based decision making and innovation.

Darryl Plecas, Professor Emeritus, University of the Fraser Valley

Darryl Plecas is Professor Emeritus in the School of Criminology and Criminal Justice at the University of the Fraser Valley. Prior to his retirement, he held the RCMP Senior University Research Chair at UFV. He received the University's Teaching Excellence Award, and in 2003 received an Innovative Excellence in Teaching, Learning and Technology Award at the Fourteenth International Conference on College Teaching and Learning. He has published numerous articles relating to Canadian criminal justice. He also has extensive experience in applied policy and program evaluation, and effective decision-making.

For more information about this book or its authors, contact Dr. Irwin M. Cohen at the Centre for Public Safety and Criminal Justice Research in the School of Criminology and Criminal Justice at the University of the Fraser Valley, at irwin.cohen@ufv.ca or 604-854-4553.

For additional fire and public safety publications, please visit www.cjr.ufv.ca

Table of Contents

Canadian Safety and Security Program

This project was supported by the Canadian Safety and Security Program (CSSP) [CSSP-2013-CD-1109, Establishing Evidence-based Policy (Fire Community of Practice)] which is led by Defence Research and Development Canada's Centre for Security Science, in partnership with Public Safety Canada.

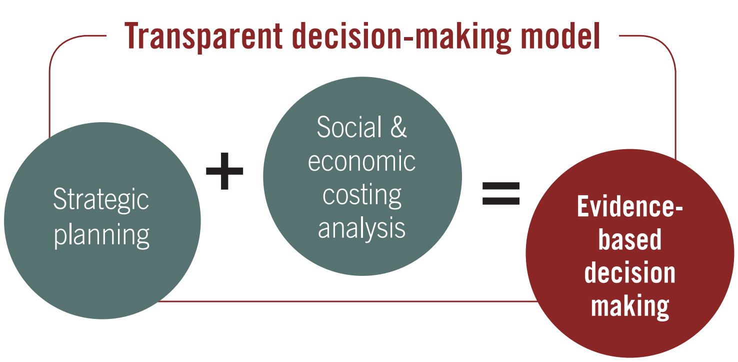
Partners in the project include the CSSP, the City of Surrey (Surrey Fire Department) and the University of the Fraser Valley. The CSSP is a federally-funded program to strengthen Canada's ability to anticipate, prevent/mitigate, prepare for, respond to, and recover from natural disasters, serious accidents, crime and terrorism through the convergence of science and technology with policy, operations and intelligence.

Forward

Canadian fire services are key to protecting lives while maintaining our civil infrastructure, but their high capital and labour costs often pose significant financial challenges. They are one of the major budget items for most cities and municipalities. Fire service professionals make crucial decisions regarding the level of service they can provide their communities and the demands they are going to place on those communities.

While understanding the invaluable role of fire services, both the public and municipal leaders are asking that significant decisions be based on hard evidence. Questions such as what is the issue's underlying strategic value and what are the associated costs and benefits are commonly raised.

This manual has been created to help decision makers address those concerns.



Evidence-based decision making is one of the more effective tools you can use to rationalize why a particular approach or program option was chosen. Evidence-based decision making is not new. Rather, it is a framework that brings together strategic planning with social and economic costing analysis within a transparent decision-making model.

This manual provides an overview of some of the more crucial components of evidence-based decision making. Some of the material may appear a little daunting at first sight. But we would ask that you read it in small chunks and go over it more than once. As with any volume of this type, the material often makes more sense when you try to link it to a problem or issue your own organization is facing. Overall, we hope you will find this manual useful in improving your decision making and justifying your choices.

We would like to thank the following people for their assistance in producing this manual. First, we thank Collette Squires, who contributed both content and research assistance. We also express our gratitude to Larry Thomas for providing valuable fire service context in the creation of the examples expressed throughout.

Introduction

Effective Decision Making in a Changing World

While the primary function of fire departments has remained consistent over time—to protect human lives and property—how departments deliver their services is becoming more complex. The rate of fire incidents has generally declined over the past few decades, but the public is asking departments to respond to a broader range of calls. Those calls often require more sophisticated equipment and better trained personnel. Furthermore, we are asking fire services increasingly to integrate their functions with other first responder agencies, such as Emergency Medical Services.

As a result, leaders and managers continually face this question: How can we provide quality service in light of more complex demands while being sensitive to resource and economic restraints? Choices and trade-offs need to be made, and consequences need to be considered. The pressure increases on decision makers when politicians, municipal staff, and ultimately, the public scrutinize these decisions. The days are gone—if, indeed, they ever existed—where government and taxpayers take a request for more equipment and more personnel at face value.

The days are gone—if, indeed, they ever existed—where government and taxpayers take a request for more equipment and more personnel at face value.

Politicians, city managers and higher executives are increasingly forced to make choices within tight resource constraints.

More than ever, leaders in fire departments need to make decisions in ways that are transparent and justifiable. Good decision making, we will argue, needs to be informed as much as possible by evidence, research, and sound information. We term this approach evidence-based decision making. We make and justify evidence-based decisions by referencing independently supported and verifiable facts. This approach helps ensure that the decisions we make are sound and defensible. Used effectively, evidence-based approaches can help you produce the results for which you are searching.

So why is evidence-based research important? Why does this approach to problem solving matter for fire services? Among some key reasons are the following.

- Policies and programs not guided by sound evidence frequently cost too much, waste resources, or simply yield poor or unknown results.
- External decision-makers who approve departmental budgets may not view departmental requests as justified if they lack compelling evidence.
- Policies and strategies that are evidence-based often produce better results, which can increase your credibility and support for the department as a whole.

This manual will help you understand how to find and use the information and research needed to make evidence-based decisions. It will also help you to put your decisions within a compelling framework to convince others of their merit.

Of course, not all decisions are or can be based on facts. Both professionally and in our personal lives, we refer to values, preferences and political choices. To believe or do otherwise would be to deny the complexities of social life.

This manual will help you:

1. Find and use information and research to make evidence-based decisions.
2. Put your decisions into a compelling framework to convince others of their merit.

Yet, even in those circumstances, evidence-based decision making can help you link the values, principles, and ideologies that guide your department to independent evidence and supportive research.

The evidence that we will learn to use comes from a variety of sources. Some is available as administrative data that government and other formal organizations routinely collect. Some is generated in the course of formal policy and program evaluations, and some will come from the work of government and academic scientists. Other sources of information will include your own organization and, often, your own unit or department.

Learning to Navigate the World of Evidence

This manual will help you to navigate the world of evidence without feeling intimidated by it. As we will discover, not all evidence or data is of equal value. Even good information needs to be placed in a context where we can evaluate its accuracy and meaning. In other words, this manual will help you figure out what you need to know about data generation without having to be a scientist or scholar.

Besides learning how to assess evidence, we will also discuss how to use evidence to formulate a persuasive argument. Data alone is not sufficient to inform and support your decisions. We need to frame public justifications for our policy or program decisions logically and coherently. Requests not grounded in a sound strategic or business plan will have very little chance of success. As we will discover, many arguments or justifications that are put forward simply do not make sense. We will examine some major logical fallacies that are to be avoided at all costs.

This manual will also explain how to conduct an environmental scan and a SWOT analysis (an assessment of an organization's Strengths, Weaknesses, external Opportunities, and Threats). You will learn why those frequently form part of information collection before a new policy or program is developed, or before strategic priorities are determined. You will learn about cost-benefit analyses and costing studies, which are critical components of strategic planning when resources are tight.

Using examples from fire services around the world, this manual will show you how to define a problem. It will help you to think critically and creatively about it, and find the evidence you need to inform your decision. Additionally, it will provide simple explanations of various forms of research so you will know how and when to use them to support your case.

Before we begin, though, it is helpful to think more deeply about the reasons for doing all of this. How and why has evidence-based decision making become so important? Why should you, or anyone else, care about the process?

Medicine and Health Care Professionals That Have Led the Way

We can trace the origins of evidence-based approaches back to the 1980s. Faced with significant financial challenges, the government of the United Kingdom started to emphasize the need for policies and best practices supported by compelling evidence and empirically-sound research. Decision makers had wasted too many resources, they believed, on choices that had no evidence to back them up. They too often decide on the basis of personal preference, traditional practices, and ideas that had little more to support them than they were popular at the time.

As anyone who has been in their field for any length of time knows, the world is full of scam artists selling the latest managerial elixir or practice. Within the UK, it was obvious to the government that investments were needed, but those investments needed to be effective and efficient.[1]

This approach influenced many other fields, most particularly health sciences, where researchers could directly link poor practices to increased levels of harm for patients. Evidence-based medicine evolved as a way to reduce the gap between academic research and clinical practice. Ideally, this would ensure the best possible outcomes and the most appropriate care for patients. Researchers and health care professionals scrutinized policies and procedures to see how they could run medical facilities in more efficient and effective ways.[2]

The need to change existing ways of doing things in the world of medicine was becoming increasingly apparent. For example, one major study suggested that it took approximately 15 years to incorporate the results of research into recommended policy. As a dramatic example, let us consider that the research basis underlying a cure for a particular form of cancer might already exist. However, the lag between that discovery and even partially implementing it in a clinical setting takes about a decade and a half. Even after that extended period, only about 40 per cent of practitioners are using that information.[3]

Meanwhile, people who could benefit from the results of that research continue to suffer or die because information had not influenced the practices of the medical profession in a timely way. Worse still, implementing the answer might be intentionally delayed if other groups saw greater benefit and financial profit in "managing" the disease rather than in actually curing it.

An evidence-based approach tries to use the best available information generated through research, experiments, observation, and other factual sources to influence the creation of the best decisions and policies possible. Sometimes, this can directly conflict with other forces, values and interests, as the previous hypothetical example illustrates.

Case Study

A real example from the medical field shows another way this can play out. In the United States, the Institute of Medicine (IOM) released a report in 2000 revealing that hospital errors across the country resulted in a loss of nearly 100,000 lives each year. The report not only explained *what* was happening, but *why*. These errors were not resulting from individual incompetence or ineptitude, but were the result of system failures.

The report recognized that people in the health care system were among the best educated and dedicated of any workforce in the country. The problem was not inadequate people; the problem was that the system needed to be made safer.[4]

This IOM report raised alarms throughout the medical profession, especially since healthcare professionals considered themselves to be devoted to *saving* lives, not *costing* lives. Their ethical beliefs and commitment to competence and applying scientific knowledge were undermined by actions—and sometimes inactions—that actually increased the death rate among their patients.

Although this report was shocking, its release created a unique opportunity. The Institute for Healthcare Improvement (IHI) had already spent a decade working to reduce the loss of life due to human error, and this report was a goldmine of helpful data. Subsequently, the IHI launched the 100,000 Lives Campaign with the highly ambitious goal of reducing the embarrassing and devastating accidental hospital and clinical deaths to the only acceptable level possible: none at all. Within two years an estimated 122,342 deaths had been prevented.[5]

While recognizing the differences between healthcare and firefighting, the success of this initiative proves important ideas that are still very relevant:

- The policy decision to launch the 100,000 Lives Campaign was driven by the ideological commitment to save lives and reduce the death rate, but these values were grounded in sound research. The take-away here is that core values may guide our policy development, but we need to anchor and partner them with credible research and information.
- The Campaign had a very specific goal: saving 100,000 lives. This number came from what the research suggested might be possible. Quantifying their goal enabled the IHI to inspire others to get involved and to aim for a specific target.

The healthcare practitioners faced the same challenges that other professionals face in similar situations. There was the difficulty of trying to shift attitudes and behaviours in a large work force. There was the pressure to stay current with the latest research. There was the challenge of creating change in the face of evolving technology and budget restraints. Furthermore, overcoming skill and knowledge deficits was necessary, along with the large challenge of inspiring system-wide support for a new vision of how we should deliver service. Despite these obstacles, positive change occurred.

Lessons Learned

Jeffrey Pfeffer, from the Stanford Graduate School of Business, examined the results of the IHI initiative and noted that the IHI had acted based on the following premises, all of which can apply to other fields.

- Things can be improved.
- Improvement will come over time, through a succession of actions, each of which will provide the opportunity for learning.
- We should not wait to solve everything before beginning to improve some things.
- We should be modest and realistic about our insights and abilities.
- We need to do something, because in the absence of informed action, nothing will change. And we can learn as we proceed.[6]

The decision in health services to make things better resulted in significant improvements in saving lives. Similarly, there is no reason why fire services cannot improve their results and increase their effectiveness by adopting an evidence-based approach.

Effective Decision Making: The Task of Good Leaders and Managers

Decision making is possibly one of the most important roles of leaders and managers. Their decisions influence the direction of their units and affect the morale and well-being of personnel who work for them. Poorly made decisions increase conflict and diminish morale. Well-made decisions that lead to tangible, positive results can increase departmental success and improve morale.

Nevertheless, even when leaders and managers see the value in an evidence-based approach, several factors can get in the way. Some administrators feel pressured to make decisions quickly and with incomplete information; others might use outdated information. Additionally, most people rely on personal experience, observation, or gut instinct when having to make a choice. As trained professionals, our personal experiences and judgments are often valid, but they comprise *part* of the picture only. Using evidence-based research helps to ground our experiences and opinions in a broader context of information that is ultimately more convincing. Besides, practices evolve. The fire service of the early twentieth century is not that of the new millennium.

Poorly made decisions increase conflict and diminish morale. Well-made decisions that lead to tangible, positive results can increase departmental success and improve morale.

When developing a new strategy or policy it is best to assess what you know, what others around you know, and what the wider field of research tells you about it. It is also prudent to commit to evaluating that new policy or plan after you have started it so you can generate your own evidence to show its effectiveness. That helps to advance the field as a whole, and your department's research can then inform other departments on what works, what does not, and why. Often we are reluctant to assess a program or practice because we might find that it does not work. That is not a problem. Both as individuals and as a society, we typically learn more from our failures than from our successes.

The Nature of Empirical Research

What are we really talking about when we use the term *evidence*? Unlike the evidence that might come out of a police investigation of a crime scene, evidence in this context has a specific meaning. It refers to the results of empirical research coming from systematic data collection grounded in formal assessments, experiments, or other research models. It is a systematic approach to answering a research question that generates information or facts that are replicable, observable, credible, verifiable, and supportable.

When assessing the research available to you, some of it will be:

- *Quantitative*, generating numbers and statistics, or
- *Qualitative*, generating subjective information that is helpful in determining preferences, values, or perspectives of those responding to the questions.

Either of those sources can generate valid data. The key is in knowing when and where to use what kind of evidence, and to be able to find out whether it is adequate for the purposes at hand.

While there are many good sources of supporting evidence, academic research has the added benefit of being peer-reviewed. This means that other independent scholars and researchers examined the research to see if it was credible and well designed. This does not mean to say that the work is either perfect or infallible. Nevertheless, it does increase your ability to trust in the results. Research must be peer-reviewed before it is published in most academic journals. Some academic journals can be highly technical and very intimidating to those outside the field. Fortunately, many sources summarize significant academic findings or translate the results into everyday language.

Common Research Methods

In the medical field, the gold standard for research has been the randomized controlled trial. Here researchers randomly assign individuals to receive various preventive, therapeutic or diagnostic interventions, and then follow up to see the effect of the intervention. One possible intervention might be no intervention at all. This enables researchers to compare the control group (which received no intervention) to the test groups, which received the various interventions in question. Drug testing is frequently done this way. In a later chapter we will examine different frameworks for collecting evidence and discuss why researchers hold the randomized controlled trial in such high esteem.

In the social sciences, having randomized tests involving a control group is also possible. For example, we could randomly assign people with smoke alarms in their homes as a test group for comparison with another random group without alarms, which would be the control group. This is one way of answering the research question, "Do people with smoke alarms have less costly fire incidents than people without smoke alarms?" Researchers will set up such experiments to 'control' the factors that might skew the results. This increases the *validity* of the research, so that you can have greater confidence or trust in the measurements and results. Researchers are also concerned about the *reliability* of their result—meaning, if we continued to replicate this study repeatedly, would we get the same results? Would we get the same results if we ran this test in a different community? Or, is it unique to this community only and, if so, why is that? Research needs to be both *valid* and *reliable* so you know the results are legitimate and trustworthy, and not a fluke or coincidence.

Engineers conduct a great deal of fire research in laboratories under controlled conditions. From those studies we learn about how fire behaves under varying conditions. Similarly, we test most safety equipment in lab settings. There is also a substantial amount of research conducted on logistics and dispatching. The evidence from those studies has done much to improve response times. Currently, research continues on the use of sophisticated GPS techniques to locate resources, match events to existing databases and keep track of equipment and personnel.

Making Better Decisions

By now, you probably can see that there are benefits in making decisions influenced by sound, credible research. Quite simply, if you have done your homework, it is likely you will have a better-informed decision. Defending your decision is also easier since the process is more transparent and is based on something other than your hunch, best guess, or personal opinion.

It is important to recognize, though, that evidence-based decision making is best suited for objective questions. As we noted at the outset of this chapter, other decisions are influenced primarily by our preferences, values, or beliefs, and are less likely linked to research.

However, the two merge when we want to find the most effective ways to address issues that ultimately connect with our values. In the previous example of the 100,000 Lives Campaign, the underlying value for saving human life provided the motivation to do things differently. Evidence-based research helped them know *what to do* and *how to do it*. That way, they could achieve their desire to save more lives.

Notes

1. Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review," *British Journal of Management*, Vol. 14, No.3, pp. 207-222.
2. Vishwanath V. Baba, Farimah HakemZadeh, (2012) "Toward a theory of evidence based decision making," *Management Decision,* Vol. 50 Issue 5, pp. 832-867.
3. Antman, E.M., Lau, J., Kupelnick, B., Mosteller, F., Chalmers, T.C. (1992) "A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction." *Journal of the American Medical Association*, Vol. 268, No. 2, pp. 240-248.
4. Kohn L, Corrigan J, Donaldson M, (Eds.) (1999) *To Err is Human: Building a Safer Health System*. Washington, DC: Committee on Quality of Health Care in America, Institute of Medicine. National Academies Press.
5. The Commonwealth Fund. (2006). Saving lives, raising hopes. Retrieved from http://www.commonwealthfund.org/Content/From-the-President/2006/Saving-Lives--Raising-Hopes.aspx
6. Cited in: National Institute of Corrections (2010) *A Framework for Evidence-Based Decision Making in Local Criminal Justice Systems*. 3rd Ed. Center for Effective Public Policy. P. 11.

Defining the Problem

Not all Decisions are Alike

We make hundreds of decisions daily, ranging from what to have for breakfast, to deciding in which room to hold a meeting, to whether or not we should buy a new car. Many of those decisions are informal, relatively insignificant and have few consequences. Working through a formal process to address those issues would cause our lives to grind to a halt.

On the other hand, there are significant decisions we face in our personal and our professional lives where the consequences are not small. Examining the issues in detail and working through a formal process is worth our time and effort. Generally, that formal process involves creating a clear definition of the problem, outlining the alternatives, and weighing the costs and benefits associated with selecting any of those alternatives.

Evidence-based decision making can help us in those circumstances where we need to make an economically, socially or politically significant decision. An advantage of evidence-based decision making is that it allows us to use known results to estimate a measurable outcome.

Evidence-based decision-making can help us in those circumstances where we need to make an economically, socially or politically significant decision.

One can never know the actual consequences of a decision before the event. However, by drawing on experience and the available evidence, generating a reasonable and defensible expectation of a specific outcome is possible.

All of us will make decisions that lead to undesired outcomes at times. That is a reality of life. The fact that we made the wrong choice is different from making a bad decision. There is a difference between not making the correct decision and bad decision making. As we will outline, bad decisions are avoidable. Bad outcomes from good decisions, however, are fortuitous events over which we might have little control. So what then, distinguishes a good decision from a bad decision? Simply, good decisions are ones where the problem is clearly articulated.

They are ones where we bring as much of the appropriate and available evidence to bear as possible. A good decision is one where you can look back and with a clear conscience assert that under the same circumstances, and with the same evidence, you would come to the same conclusion.

While getting a less than ideal outcome from a good decision is unfortunate, one other advantage of having made a good decision is that we can draw lessons from it. If the decision making process is transparent, it is possible to consider why it resulted in a negative outcome. Did we make some incorrect assumptions? Were we missing some important information? Did we encounter new or different circumstances? An open and formally structured process allows us to accumulate knowledge so that we are less likely to make the same mistake in the future.

In this chapter, we will consider the following:

- What is the issue and how do we problematize it?
- How can we identify the options and alternatives?
- How can we think creatively to generate new ideas?
- How do we generate alternatives?

What is the Issue?

Typically, even rational, systematic decision makers will start by making a list of alternatives. Lists are good and they definitely have their place. Nevertheless, as John D. Rockefeller once said in a different context, “A list is not a plan.”

Before we start to generate options, we need to ask: What is the purpose of the decision? What is our intended goal? Those questions are embedded in an analysis of the problem. The framework of that analysis is generally a strategic or a business plan. Making a decision without planning is common.

Typically, however, unplanned decisions do not end well. Planning allows us to make decisions logically and systematically. Proper planning makes decision making simpler and it makes it transparent. That is, we can show critics that the choice we made was rational and reasonable under the circumstances. Proper planning makes decision making defendable even when the results are not as expected.

When we ask the question, “What is the issue?” we are essentially asking, “How does the decision we are facing fit into and advance the mandate of our organization?”

Before Doing Anything, Ask "Why?"

Too often, we find ourselves backed into a corner when confronted with the seemingly simple request about whether we should choose Option A over Option B. This is a popular strategic move by someone who wishes to force an issue. For example, an employee may ask for a meeting to discuss performance and salary. As an opening gambit, the employee might ask, "Are going to give me the same raise as last year or will I also get the promotion I have coming in recognition of my service to the company?"

Clearly, the employee is attempting to force a false choice. In this instance, we call it a false dichotomy because the question assumes that only the two options A or B are possible. In fact, many options may exist. Before considering the many possibilities, assessing the employee's contributions to the organization is necessary. Ideally, there should be a performance assessment policy in place. Lacking that, however, you might ask some of the five Ws. *Why* should you be rewarded based on your performance? *What* have you contributed to enhancing the effectiveness of your unit? *Where* can we see evidence of your contributions? *Who* in your unit have you helped or supported this year? *When* can we expect to see the returns on your performance?

Perhaps these are not the most appropriate questions to ask in the circumstances, but you get the idea. The notion is to tie the request back to the goals of the unit or organization and to ensure that the choices that we are considering are consistent with those goals. Typically, we are trying to ensure the bases for the choices are not irrelevant. Decisions to reward employees simply because they are friendly, consistently show up for work on time, or always dress neatly are difficult to defend.

When all else fails, ask yourself, "Can I defend my decision to others in the organization, my boss, or the public?" As a former colleague once said, "I make every major decision assuming it will appear on the front page of tomorrow's newspaper. If I can accept that, then I have likely made a reasonable choice on reasonable grounds."

That mandate is normally part and parcel of our strategic plan. Sometimes it is embedded in our operational plan or standard operating procedures (SOPs).

As an example, let us assume that a Fire Chief has just come back from a conference where statistics show that typical residential fires are on the decline. However, the incidence of incendiary events where hazardous chemicals are involved is on the rise. The issue has been reinforced in the press by the widespread coverage of several horrendous explosions involving the rail transport of crude oil through residential areas. While the local department has the "basic" equipment, training and personnel available for a response, the Chief now thinks this level of preparation is inadequate.

The issue the Chief faces is whether he should initially increase the amount of HazMat (hazardous materials) equipment in his department, or whether giving his staff more advanced HazMat training would be better.

At face value, it is hard to argue that one should not pursue extra preparation for such events. The reality is, however, that in developing its strategic plan, the department needed to balance the requirements of the local community. Few chemical processing facilities exist in the area and there are few, if any, shipments of dangerous chemicals on the region's roads or rail services.

By embedding the decision within the framework of a pre-existing plan—such as a strategic plan—the choices made are defensible on strategically assessed grounds.

The goals outlined in the department's strategic plan say that the primary needs of the community are to provide better service to the more remote regions, and to reduce overall response times.

Consequently, the real needs of the department might be another pumper truck or an upgrade in the department's dispatch system. By referring to a planning framework, we can see that a greater emphasis on HazMat is not a priority. Furthermore, in all likelihood, the incremental investment in that area would be wasted and provide no return on investment.

Investing further in HazMat response fits with the overall goal of the department to improve the safety of the community. The real issue, however, is whether the proposed investment fits with the real and immediate needs of the community.

The issue is not one of improving the overall safety of the community, the issue is really how best to address the most likely threats the community faces.

Undoubtedly, the Chief could have listed the options available to the department regarding HazMat acquisitions and the best alternative among those options could be selected. The point, however, is that decision was not one to be considered. The key was to refer to the department's operational focus or, ideally, its strategic plan.

Obviously, if the community's circumstances changed—if, for example, truckloads of chemicals were being shipped on the local roads—the strategic plan would need to be reassessed and, perhaps, the priorities changed.

Again, by embedding the decision within the framework of a pre-existing plan or operational framework, the choices made are defensible on strategically assessed grounds. In that case, a delay in implementing a new dispatch system to provide more HazMat training is justifiable.

Generating Ideas

Often, choices appear obvious. Do we spend more on equipment or personnel? Is our communications equipment at the end of its working life expectancy or not? In other instances, the alternatives are not always self-evident. It is not an A or not-A choice. In later chapters, we will examine how to conduct environmental scans and SWOT analyses. These are relatively formal procedures that systematically review what others have done or might do in similar circumstances.

Before resorting to those approaches, however, several more modest ways exist to generate alternatives. You might want to consider the following options.

Talk to people outside your normal circles

Too often we limit our social and professional circles to those we already know or with whom we work. Often, this generates a group-think mentality where we reinforce the belief in a limited number of options. Outsiders, however, may face similar situations but approach the issue entirely differently.

Engage in a group brainstorming session

Possible group-think tendencies aside, sometimes the people around you are the best source of ideas. They know the organization and understand the problems.

Furthermore, they are less expensive than consultants since they are already on payroll. Ask for individual suggestions. Sometime a group session where we ask people to come up with “crazy” alternatives is effective. The semblance of a little competition can sometimes unleash new ideas. Remember, today’s innovations were yesterday’s impossibilities.

Read more books and journals; surf the web

The more you read, especially outside your area, the more novel ideas you are likely to come across. Business books are an obvious choice but sometimes great ideas come from works of fiction. Professional journals are a good way of keeping up with new trends. As always, the internet is anarchy and generally fits the adage that you get what you pay for. Still, gems are to be found and modern search engines are amazingly good at ferreting them out.

Focus on your clients or customers

Look at the world from the perspective of the people you serve. How they see your organization is probably very different from how you and your colleagues see it. Besides customers or clients, other great sources of ideas are suppliers. Often these people come by, fill an order and take their commission. You might as well get something for that commission you are paying. Ask them for alternative products or services that are available on the market.

Hire a reputable consultant

Often, you are the local expert at your core business or activity. That is why you are in your position. On the other hand, not all of your decisions relate to your core enterprise. Most businesses engage outside design firms, marketing agencies, web designers or management consultants. The key is to identify the area of expertise that you require. Once done, ask your associates if they can recommend a consulting firm or individual. Usually, smaller firms are more creative and less costly, but creativity is a business.

❖

Of course, you need to be willing to be open to new perspectives. Don’t let your prejudices get in the way. Just because you have a low opinion of someone does not mean they have bad ideas. Also, do not feel intimidated because someone can generate better ideas than you. Especially if that person is a subordinate, you automatically get credit for being smart enough to having hired a creative employee.

Finally, be willing to accept that sometimes, the best options are the obvious ones. A consultant who gives you a report that tells you what you already know, may not simply be lazy or uncreative. It could be that what is obvious to you is indeed the best option. Consider it that your suspicions have been confirmed.

Get a Plan

Whatever its size or complexity, every organization can benefit from having a plan. Whether it is termed a strategic, organizational, or business plan, the point is the same: an organization needs to know why it is doing what it is doing, where it is going, and how it intends to get there.

Without a plan, people make decisions arbitrarily. At best, those decisions will lack consistency and, at worse, they will be contradictory. A plan does not guarantee organizational success or efficiency. Not having one invariably dooms an organization to mediocrity or failure.

Much material outlining how to put together an organizational plan is available both in bookstores and on the internet. Time spent reviewing some of that material would be a good investment.

Essentially, a plan consists of four elements:

1. A general statement of organizational values.
2. A statement of goals and objectives.
3. An outline of how the organization intends to carry out or achieve its goals.
4. An indication of how to measure success.

Four elements of a plan:

Plans vary in complexity but there are advantages to keeping it simple. Complex plans are often difficult to remember and can be highly constraining. As most battlefield generals know, once the action starts, little goes as expected. Often, the best one can hope for is that the troops know what they are fighting for, that they remember the overall goals and objectives, and that the line officers are sufficiently trained to react to unexpected tactical challenges and setbacks. Thus, there is a lot to be said for keeping things simple.

Well-crafted mission, vision or value statements can be inspiring; poorly crafted statements do little more than provide a source of levity.

Statement of Organizational Values

It is currently in vogue among management gurus to spend a great deal of time identifying the fundamental values underlying our organization. Typically, we outline organizational values in one or more of: a mission statement, a vision statement and a values statement. Well-crafted statements can be inspiring, and make for eloquent poster boards that can be placed on office walls and annual reports. Poorly crafted statements do little more than provide a source of levity. As always, the best practical advice is to keep things simple and straightforward. Simple, unambiguous statements are easy to remember and easy to follow.

Essentially, a statement of value should outline the reason for the organization's existence. This is known as the mission statement. For many organizations, such as fire departments, the mission may be obvious. Your raison d'être is to save lives and property or, in a broader sense, to create a safer community. The mission statement is where you answer the great existential question, "What is your purpose?"

Value statements should also provide some expectation of where the organization plans to be in the next three to five years. What, in other words, is the midterm vision for the organization? Perhaps you see yourself as becoming the regional standard for performance.

Finally, a values statement suggests something about your core beliefs. These are meant to be foundational and inspirational. For Google, it was, "Do no harm." In your case, it may be, "Serve the community." While this might seem trite, it is useful to recall the core value when decision making starts to focus too much on what is in the best interest of the organization. In this instance, what you do is not about the organization; it is about serving your community.

Statement of Goals and Objectives

An organization's statement of goals and objectives contains the targets it sets for itself. Organizational goals are the broader targets for which one is aiming; objectives are the midterm step one sets to achieve those goals.

Broad goals may be such things as setting targets for average response times, reducing the number of fire fatalities in the community, or increasing the unit's capacity to handle a broader range of service demands. To achieve the goal of reducing average response times, it is often necessary to make a list of objectives that form a series of intermediate steps. For example, one objective might be to better integrate the dispatch system into a regional data base. Another might be to ensure that a minimum number of weekly training drills is performed by each fire hall.

Implementation Procedures

As we noted earlier, a list is not a plan. Simply outlining the organization's goals and objectives is a necessary but not a sufficient part of the planning process. A true plan involves a discussion of how we can carry out the goals.

What is the mechanism or what are the procedures that are being put in place to meet the desired outcomes? For example, one objective may be to reduce on-the-job injuries. We may link this to the overall goal of increasing worker safety.

Napoleon's goals and objectives

Sometimes it is easy to confuse the concepts of goals and objectives. Too often, the two are used interchangeably. While related, the two are distinct notions. A good example is to consider Napoleon Bonaparte's intentions in 1799.

Goal	Objective
Rule all of Europe	Become head of state in France
	Conquer Italy
	Conquer Spain
	Defeat Prussian Army
	Defeat the Austro-Hungarian Army
	Incorporate Poland into the French Empire
	Conquer Russia

Ironically, Bonaparte achieved all of his objectives except for the last. Despite this impressive achievement, he ultimately failed to achieve his overarching goal.

> Too often, strategic and business plans identify what the organization intends to achieve, but not the means by which it hopes to meet those intentions.

The important issue under implementation is: How do we make this happen? Obviously, the mechanism we choose depends upon the circumstances. Perhaps more training for operational personnel is necessary. On the other hand, people may have adequate training but they have not had sufficient opportunity to practice safety procedures. Another mechanism might be to upgrade the existing safety equipment to current or possibly forthcoming standards.

This applies to all of the goals and objectives identified in the plan, whether they are "soft" objectives, such as increasing employee morale, or "hard" objectives, such as reducing fatalities and injuries. Implementation procedures are the actionable items in our plan. Too often, strategic and business plans identify what the organization intends to achieve but not the means by which it hopes to meet those intentions.

Put another way, if goals and objectives are the nouns in a sentence, implementation procedures are the action components or verbs.

Measuring Outcomes

Measuring outcomes is essentially keeping a scorecard. Before you can do this, however, it is necessary to indicate within your plan what specific performance indicators you are going to use. You should closely link those indicators to the specific objectives you have identified and, in a general sense, to the overall goals outlined in the plan.

Obviously, clear quantitative measures are easiest to use, such as changes in calls for service, response time, proportion of home inspections conducted, and so on. However, qualitative measures should not be overlooked. Indicators of community satisfaction, for example, may be hard to quantify but are crucial performance elements for service providers.

Typically, outcome measures will cover a spectrum of issues, ranging from internal performance metrics, to levels of service provision, to financial accountability. Many discussions on strategic plans suggest creating a table where we list operational objectives in one column and their corresponding measures of success in the next.

These are judgement calls, but complex objectives usually require more varied indicators than simple, one-dimensional objectives.

Because goals are longer term and higher level notions than objectives, it is often more difficult to identify specific measures. Furthermore, goals often require a more qualitative assessment than do intermediate objectives. One thing to keep in mind, however, is that while there ought to be a consistency between the outcome measures of objectives and goals, there need not be a perfect correspondence. It is possible to meet most or all of one's objectives but not one's goals. Similarly, the failure to meet one or more objectives does not necessarily mean that the organization has missed its overall goals.

Practical strategic or business plans sometimes contain other items or provide more detail on certain dimensions. For example, some plans say who is responsible for carrying out certain objectives.

Details might also be put in place about what forms the organization's competitive advantage or how it differs from similar organizations or competitors. Whether these items are relevant depends on the particular environment and circumstances in which the organization finds itself. Regardless, those components become part of the crucial list of elements to which we refer when we need to make a critical decision.

Often we pose questions or decisions vaguely. A good decision maker will define and clarify the issue and relate it to the organization's plan. Having done that, one can then ask subsidiary questions such as: Does the issue warrant action? If so, when should we carry it out? Is the matter urgent, important or both?

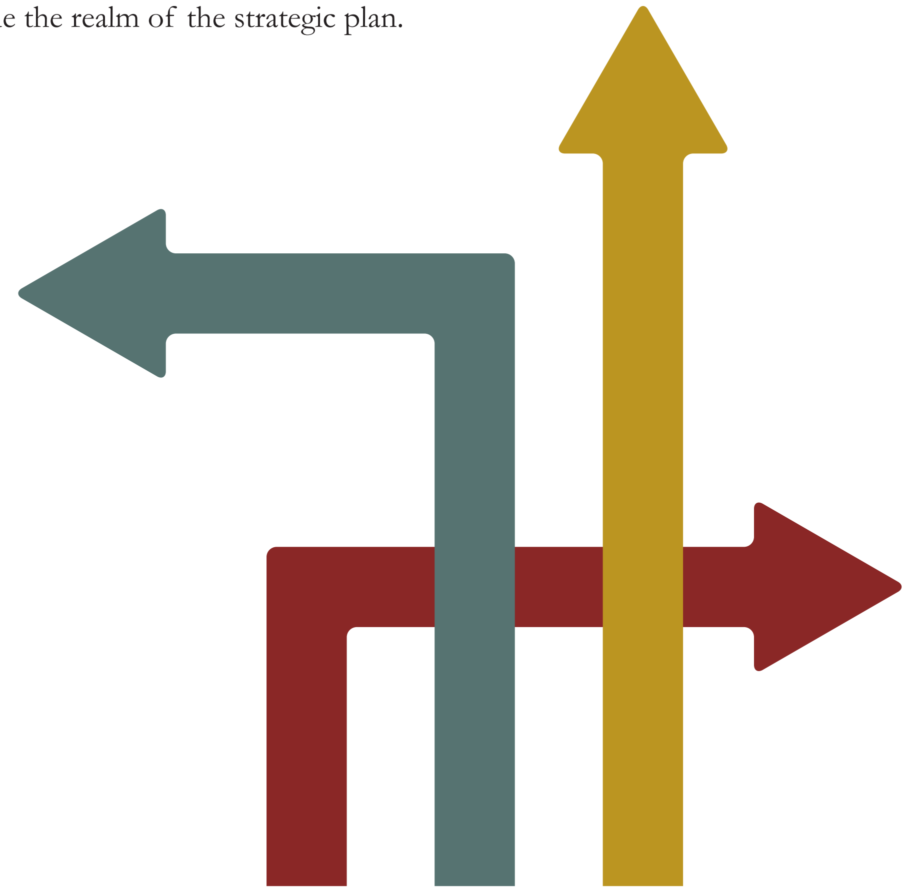
Evidence-based Decision Making

Good evidence-based decision making is tightly linked to an organization's plans. This does not mean that occasionally we must make important decisions that are beyond what we planned to do. Environments change and new issues arise. The world is not static.

Good managers, however, need to be sufficiently flexible to deal with those situations. Regardless, going through a planning exercise often provides a broad enough perspective or sufficient guideposts that "out of the blue" challenges can be placed within the general framework of our plans.

The primary benefit of a good plan is that it allows decision makers to be able to justify why they are assessing the choices they are considering. Raising the criticism that certain options have been considered is easy. In fact, for many decisions there may be an almost infinite list of possible options. We can reduce that list substantially if we point out that the suggestions may have merit, but are outside the realm of the strategic plan.

A good plan, then, lets us know what questions or issues are relevant, what options are worthy of consideration, and consequently, what evidence we need to consider in weighing those options.



A good plan lets us know what issues are relevant, what options are worthy of consideration, and what evidence we need to consider in weighing those options.

Thinking Critically

Clarity of Thought

Evidence and data alone are not sufficient for making good and useful decisions. How we formulate an argument or explanation is just as important as the quality of the information we might bring to bear. When we consider evidence-based decision making, we need to keep two aspects in mind. First, as in making any type of case, the underlying arguments need to be based on sound logic. An argument that can lead to more than one conclusion is generally not very useful. Second, how most people think evidence or proof shores up an argument is typically not the most powerful way of making a case.

If there are two things that seem to characterise humanity, it is that people like to argue and, even when someone shows that their position is false or illogical, they generally won't change their world view. Humans are stubborn beasts with a tendency to defend any coveted untruth against the best of reason and evidence.

Evidence seems to abound that argumentation is one of humanity's most favoured social activities. Go to any sports bar on a Saturday night and you will see what seems to be inexhaustible evidence.

Humans are stubborn beasts with a tendency to defend any coveted untruth against the best of reason and evidence.

Then, there is the internet. Its rise has been the greatest venue for half-baked ideas, conspiracy theories and their supporters since the invention of walls and graffiti. Fundamentally, evidence and sound logic rarely sway people. When was the last time, for example, someone listened to you make a case and said, "Thank you for pointing out my logical fallacies. I see that I was wrong on this issue and I will from now on change my perspective on the matter." A positive outcome is typically one where they change the topic; a negative outcome is where they turn away muttering something about you and your kind having always been idiots.

The fact is, there are some discussions to which no solution exists, either logical or empirical. Arguments over the existence of God; who is the best looking actor or actress; or, whether Aunt Helen made the world's best muffins will never be resolved.

Generally speaking, matters of values are issues that are based on emotional preferences.

On the other hand, there are situations where evidence and rationality sway us (or, at least, some of us). Economic issues, for example, typically command our more rational sentiments. Matters of health, and life and death—immunizing your children against the measles, for instance—tend to elicit a rational response.

Although, it is admitted that charlatans abound and thrive in those domains as in all others.

The focus of this chapter is on those instances where, either individually or in groups, we are willing to consider rational and evidence-based input into our decision-making processes. Since those instances appear rarely in the affairs of humans, it is obligatory for us not to miss the opportunity for making a sound decision by using faulty logic.

Logical Fallacies

Logical statements are generally of the form, if A leads to B and B leads to C, then the occurrence of A will lead to C. Logical fallacies are ones where inherent gaps, contradictions or simple irrelevancies in arguments go unacknowledged or unchallenged. Some logicians and philosophers have made careers listing almost infinite varieties of fallacies (again, see the internet). For the most part, however, logical fallacies fall into a small group. Learn to identify these and you will be less likely to be led astray, whether intentionally or not.

Appeals to Authority

None of us has the capacity to generate all human knowledge from scratch.

As youngsters we are taught that what our parents, teachers and other "experts" say is generally true. It is an accumulation of knowledge passed from one generation to the next that distinguishes humans from other beings. This has allowed us to develop antibiotics, to build skyscrapers and to distribute spam to those little boxes we call cell phones. Without accepting knowledge passed on from authorities, civilization could not exist.

However, while we may be willing to accept the received wisdom from our resident Yodas, we should not be blind to the fact that Yoda may be wrong. There is nothing wrong with asking for further evidence to back up some authority's claim.

> Be suspicious of opening lines such as: "But, it has always been done that way," or "The experts agree that..."

While we do not have the time to question all authority, certain appeals should raise your suspicion.

Typical openings that should cause you to be suspicious are lines such as:

- "But, it has always been done that way."
- "Everyone knows that's the way it is."
- "What do you (we) know? So-and-so is an expert in these matters."
- "Science tells us that . . ."
- "The experts agree that . . ."

In such instances, there is nothing wrong with saying that, "If that is the case, then there should clearly be some hard evidence to back it up. Perhaps we should check it out in more detail." Or, "Gee, that's interesting because some (scientists, experts, etc.) say just the opposite. How are we to resolve this?"

Usually, appeals to authority are code for either, "I am too lazy to check this out," or, "I am blowing smoke."

Personal or Ad Hominem Arguments

Ad hominem is Latin for against the person. Essentially, ad hominem arguments are ones where someone attacks the person making the statement personally. Usually, the person's sanity, morals or parentage is called into question. An ad hominem argument is an attempt to "blow-off" the proponent by undermining their credibility. Among some more polite ad hominem attacks are such statements as:

- "What do you expect from a couple of fascists (socialist, liberals, academics, whatever)?"
- "That's a typical statement from someone who is clearly out of touch with today's realities."
- "That's a typical male (feminist) response."
- "Gee, you would think s/he is an expert in the matter the way s/he is going on."
- "So, how many years have you been in the field?"

The key here is to separate the argument or assertion from the speaker. Just because one has a low opinion of the other person, doesn't necessarily mean that what they have to say is wrong or irrelevant. It may be difficult at times, but trying to respect the idea is essential if not the person presenting it.

The "Red Herring"

Red herrings are irrelevant issues that someone brings up in a discussion. For example, it is asserted in a council meeting, and may be the case, that too much money is being spent on travel, toys for administrators or overtime. Someone then asserts that this would not have happened if we had invested in the appropriate technology a couple of years ago.

The problem here is that inappropriate spending that has gone unchecked is due to a lack of financial oversight. Effective financial oversight has existed before the time of the Romans and long before computers were available. Investing in the appropriate technology may help in the oversight process but does not ensure oversight in itself. Examples are bountiful of solutions that have merely added to the problem rather than solving it.

Pink Herrings

True red herrings are items that are clearly unrelated to the issue at hand. Sometimes, however, someone may raise an issue that is sufficient to address the problem but is not necessarily a solution. We might refer those to as pink herrings. Perhaps the biggest pink herring is for administrators to argue the problems exist in their organization because of a lack of financial resources.

Certainly, money can purchase resources. All too often, however, more money just leads to more of the same. Money, itself, doesn't necessarily solve the problem. Proper oversight, a more effective use of existing physical and human resources, or a more creative approach to the issue may be more effective than simply throwing more money at the problem. What is necessary, is that existing or future resources are directed toward developing or enhancing mechanisms related to the problem.

The key to addressing red herrings is to ask how the herring is related to the problem being considered. How will the technology be used to enhance oversight? Is the appropriate software available? Are the auditors properly trained in the equipment to be able to enhance their performance? Computers, after all, only do what we tell them to do.

Similarly, we need to address the open-ended call for money questions by asking how the money will be used. The answer will likely be to purchase more equipment or hire more people. The subsidiary question then becomes: In what way will that equipment or those people enhance a process that is currently broken or ineffective?

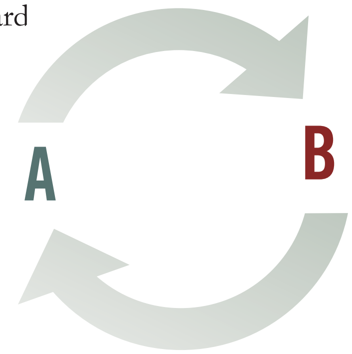
Circular Arguments

Circular arguments are those of the form that A causes B because B is the result of A. Circular arguments abound, particularly in political debates. A favourite of teachers is students who come after an exam and assert that they can't get a C because they are A students. (So, explain how you earned the C if you are an A student?)

Another good example is sometimes found in salary negotiations. Employees will sometimes insist that they need to get a larger increase than their colleagues because they have historically been the highest paid person in the unit. If you don't give the raise, how can they be the highest paid? Usually, most ratcheting effects that we see in labour negotiations are based on circular reasoning. Group A has it in their contract that they are to have a 10 per cent premium on the rest of the jurisdiction because of the high cost of living in their area. Group B argues that to remain competitive, they need to be within 10 per cent of Group A regardless of productivity or other factors. A change in the compensation of any one group automatically ratchets the pay of the other.

Sometimes we use the term begging the question to describe a circular argument. The form of the argument is essentially the same: "You know, the reason that is illegal is because it is against the law."

Similarly, an often heard comment in city councils is that a particular group will not support tax increases because they have made it part of their platform. When asked why that is part of the platform, the answer is that tax increases are not in the interest of the people.



To break the circularity, we need to know why a body passed the law in the first instance: what was its supposed purpose? Likewise, we need to know in what way not increasing taxes benefits the electorate. What is the exact economic mechanism supposed to be at play?

Other Fallacies

People call upon many other logical fallacies when rationality and evidence fail them. These range from the teenager's perennial appeal to popularity: "But everyone at school has one," to appeals to nature: "That is just not natural." Parallels, of course, abound in the professional sphere. Every municipality or department in the region has a Nouveau Widget so, obviously, we need one too. A current bureaucratic favourite is the rationale for why we keep a current practice or why things don't change. The cliché du jour is: "It is what it is," which has replaced the formerly abused, "Well, that is the nature of organizations." All of these are logically non starters.

Causal Linkages

One clinker of a fallacy we did not discuss previously goes by the formal name of *post hoc, ergo propter hoc*, which means "after this, therefore because of this." Those who might have studied statistics will recognize the variant of the "correlation does not prove causation" fallacy.

Just because two things appear associated, doesn't necessarily mean that one causes the other—or that, in fact, they are causally connected in any way. The possible absurdity of assuming that because two things are correlated they are connected is presented in the fire engine fallacy. The story here is that a Martian comes to Earth and notices that wherever it sees a fire, there is invariably a fire engine at the scene. The Martian therefore concludes that fire engines cause fires.[1]

Obviously, association or correlation is somehow related to causation. The question is, how can we identify or recognize a causal relationship when we see one? The issue is important because causal thinking and causal imagery have become entrenched in our everyday view of the world. Whenever we see something we do not quite understand, our first inclination is to ask, how did that come about? In other words, what was the cause?

Just because two things appear associated, doesn't necessarily mean that one causes the other–or that, in fact, they are causally connected in any way.

From an historical perspective, formal causal thinking is a relatively recent idea. Most scholars use David Hume's writings as the starting point for explaining what is a cause and how we might identify one. Hume was a Scottish philosopher who lived in the early to mid-1700s. Without belabouring the issue, Hume identified three necessary conditions for a causal relationship. The first condition is that the cause and the effect must be coincidental or "conjoined," as he said. This is the correlation part where two things generally appear together.

The second condition is that the cause must come before the effect. Therefore, if the Martian had been around a little longer, he would have noticed that the fire occurred first and that the fire engine generally turned up later. Thus, it was fire that caused the engine and crew to respond; fire was not a consequence of the existence of fire engines.

The third element of causation is the most difficult issue and that is what we call the condition of non spuriousness. Non spuriousness means that the cause isn't just enough or sufficient to cause the effect, but that it necessarily produces the effect or outcome. This is sometimes easier to understand in the negative. What non spuriousness means is that no third factor is resulting in the apparent cause and effect to be appearing together. An example might be a situation where fires don't cause fire engines to appear. Rather, it is an arsonist who sets a fire and then calls a fire engine to the scene. The underlying causal mechanism in this instance is the arsonist.

Spuriousness means that a relationship between two or more factors is coincidental. The real cause is an underlying third factor. The problem here is that even if we take away the apparent cause, the effect will remain. Thus, with fires and fire engines, if it is the call of the arsonist that makes the fire brigade appear, then they will appear whether a fire occurs or not.

From an evaluator's or a scientist's perspective, non spuriousness is generally the most difficult factor to control. Observing that two events generally coincide is not difficult, nor is it difficult to see that one event generally precedes the other.

Hume's conditions for a causal relationship

1. The cause and effect must be coincidental.
2. The cause must come before the effect.
3. There is no underlying third factor resulting in the cause and effect to be appearing together.

The difficult issue is assessing whether some other underlying mechanism is driving both of those events. Essentially, we have devised two ways to deal with the spuriousness issue. The first is to try to develop explanatory theories to explain how or why something should cause something else. In formal terms, we need to find what we call a causal mechanism. Logically, why should X produce Y? As we say in the trade, "What's the story?"

The second way of dealing with the non spuriousness issue is through the physical manipulation of conditions. That is, can we physically reproduce the effect ourselves. We call this manipulation an experiment.

Over time, we have developed a series of experimental designs or ways of manipulating situations so that we can isolate what we believe are the cause and effect factors from other possible or spurious influences. We will highlight those techniques in a later chapter.

In summary, then, it is sufficient at this point to consider that all three conditions must exist for us to be reasonably confident that something is truly the cause of something else. Those are the elements of coincidence or correlation; temporal sequencing where the cause precedes or comes before the effect; and, the condition of non spuriousness where no other underlying mechanism is generating both the apparent cause and the effect.

Unfortunately, we conduct much research that does not consider all three of those issues. That is why, for example, we often hear of some medical survey where some factor (say, pomegranates) are supposed to reduce the risk of cancer. Typically, the study is correlational such that someone conducts a survey and it is found that people who eat pomegranates have a lower incidence of cancer. We can probably determine that the consumption of pomegranates preceded the onset or non onset of cancer.

> Unfortunately, we conduct much research that does not consider the three conditions that must exist for a causal relationship. Thus, we see studies that, for example, claim that a certain food lowers the incidence of cancer without considering the presence of other factors.

What those studies generally do not do is to control for spurious or confounding factors. For example, pomegranate eaters may be also less likely to smoke, get more exercise, eat a healthier diet and generally have a healthier lifestyle than non pomegranate eaters. Those factors are likely the real causal agents. Including pomegranates in the diet or not is irrelevant.

Of course, once we start to believe that pomegranates are related to cancer, we can generate any number of possible causal explanations after the fact. For example, we might argue that high levels of vitamin C or antioxidants in pomegranates fight the onset of cancer.

Linking Evidence to Explanations

A common mistake people make is that by collecting sufficient evidence, one can "prove" that a hypothesis or theory is correct. In fact, the relationship between an explanation and what forms evidence is complex.

To prove a relationship we generally need to use data or evidence in two ways. First, when we consider an explanation, we must find one that is consistent with at least most of the evidence or facts that we have to date. If an explanation does not explain most of what we know, it is unlikely to be a good candidate for what we need.

Once we have narrowed our plausible explanations to ones that make sense logically, and ones that generally fit the existing evidence, we need to conduct secondary tests to see whether those explanations hold up under critical circumstances. Obviously, we have selected an explanation that fits the known facts, so simply collecting more data under the same circumstances likely won't give us more hard evidence.

Going back to our Martian example, seeing ever more instances of fires and fire trucks appearing together does not provide more proof that one causes the other. On the other hand, a few instances where fires occurred with no fire trucks about soon disproves the hypothesis.

To prove a theory:

1. We must find an explanation that is consistent with at least most of the evidence we have to date.
2. We must then conduct secondary tests to see whether those explanations hold up.

That is perhaps the single most important point that Hume made in his discussion of causation. It is very difficult to prove something is true; it is much easier to show that it is not true.

One example Hume used was that just because the sun has risen in the east and set in the west since time immemorial, it does not "prove" that this will necessarily happen tomorrow. On the other hand, all we need is one instance where the sun doesn't rise in the east to disprove the pattern. As contrived as that example might be, it does make the point about the relative imbalance between evidence that appears to show a relationship and evidence that appears to dispel a relationship.

To bring this into the language of evaluators and scientists, we can talk about the idea of a working hypothesis.

Working and Null Hyphotheses

A working hypothesis might be something like: oxygen is necessary for combustion to take place. However, just because we see a thousand instances where combustion and oxygen coexist, we really have no absolute proof that oxygen is necessary for combustion. Instances may exist where gasses or elements other than oxygen can support combustion. Thus, to provide evidence that oxygen is necessary for combustion to take place, scientists look for instances where other factors might support combustion. This leads us to what we call the null hypothesis: the notion that oxygen is not necessary for combustion. If we fail to reject or falsify the null hypothesis (so, in fact, other things support combustion) then we must logically reject the working hypothesis that oxygen is necessary for combustion to take place.

It is that strategy that scientists use to test hypotheses and theories. We cannot prove the working hypothesis directly. Instead, we create a null hypothesis that is the opposite of the working hypothesis. If we find support for the null hypothesis (that is, we find something other than oxygen that supports combustion) we toss out the working hypothesis. Or, at least, we need to seriously reconsider what is says. If we do not find support for the null hypothesis (we do not find an alternate support for combustion), we have very strong reasons to believe that our working hypothesis is valid.

Of course, it is possible that we have not yet found other elements that support combustion other than oxygen but those are simply waiting to be found. As we find that fewer and fewer alternatives pan out, the greater credibility we have in the working hypothesis.

Ultimately, we have discovered that halogens such as fluorine and chlorine can support what we term redox reactions that are similar to combustion. Still, the chemical reaction is different, so we still retain our belief that oxygen is necessary for combustion to take place.

The practical consequence of this knowledge is that we have developed mechanisms to deprive incendiary events of oxygen sources. Water, carbon dioxide, nitrogen and other materiel "smother" fires by depriving the reaction of a ready source of oxygen. By having developed a good understanding of the mechanism relating O_2 to the chemical process, we have a direction for pursuing other agents to either retard or extinguish potential fires.

Notes

1. Which would be a totally sound conclusion if they had been exposed to Ray Bradbury's classic 1953 novel Fahrenheit 451.

An Example of a Working Hypothesis and a Null Hypothesis

Because wood is more combustible than steel or concrete, it is commonly believed that fire-related fatalities in wood-framed buildings are higher than in concrete or steel-framed ones.

Some researchers have claimed, however, that with the installation of sprinkler systems, there should be no difference in fire fatality rates across the two designs. For those who believe that wood-framed buildings are more dangerous than concrete or steel-framed ones, every news report of another fatality in a wood-framed building fire is further evidence their position is correct.

However, common sense may not hold here. Wood-framed buildings are more numerous and, proportionately, fewer may have sprinklers installed. Consequently, it only appears that they are more dangerous.

A research project is being conducted looking at fire fatality rates in combustible (wood-framed) and non-combustible (concrete or steel-framed) buildings of less than five storeys with sprinkler systems installed.

At the beginning of the study, the researchers need to state the working hypothesis and the null hypothesis. In this case, we might state them as follows:

Working hypothesis:

Fatality rates among sprinklered *wood*-framed buildings of less than five storeys are higher than among sprinklered *concrete or steel*-framed buildings of less than five storeys.

Null hypothesis:

There is no difference in fatality rates between sprinklered *wood*-framed buildings of less than five storeys and sprinklered *concrete or steel*-framed buildings of less than five storeys.

The researchers test the null hypothesis. If the evidence is consistent with it, they conditionally assume that it is true and essentially reject the working hypothesis. If they find the evidence is not consistent with the null hypothesis, they reject the null hypothesis and have strong reason to assume the working hypothesis is true.

Collecting Evidence

Environmental Scans

Evidence-based decision making is influenced by the plans we create to help us set priorities, and by concerns that affect our organization's ability to fulfill its mandate. When issues arise and decisions have to be made, we need evidence to help us decide the likely impact or effectiveness of our decisions.

A common strategy for gathering this information is through an environmental scan. Simply, an environmental scan gives us an informed, comprehensive picture of the current circumstances in which our organization exists. It makes us aware of the internal and external realities, important issues, and trends that are affecting the organization. Information of this kind helps confirm or refute our perceptions. It can guide us with future programming, strategic priorities, and budgeting. An environmental scan can also be useful in determining future strategies and in developing appropriate, well-informed responses.

What benefits do organizations receive from conducting an environmental scan? Why should we spend the time and energy to conduct one? Among the most prominent are the following.

An environmental scan makes us aware of internal and external realities, important issues, and trends that affect our organization.

Environmental scans can provide:

- A fresh, objective look at issues within the organization's goals and mandate, with an eye toward how to rank them most effectively;
- An opportunity to access existing research, information, statistics, and other data that someone else took the time to collect;
- An opportunity to involve community stakeholders, organizations, individuals, and groups in decisions that affect them, by giving them an opportunity to provide input, perspective, and advice;
- An opportunity to discover the strengths and assets in the larger community to address the issue;
- A framework or point of comparison to understand the assets and strengths possessed by your own organization; and

- An opportunity to learn how your organization's programs and practices are affecting other organizations, agencies, individuals, or groups, and to what degree your programs and practices are effective in fulfilling your organization's mandate.

Conducting an environmental scan is a sequential process that involves gathering information from secondary sources, including existing research reports, statistics, or other information. This is supplemented by first-hand or primary sources of information, from individuals or groups that you will contact yourself. Analysis of this information leads to establishing where your organization fits within the broader social ecology.

Unlike many other management procedures, there are few formal guidelines for conducting environmental scans. What we will do, however, is to provide you with an overview of the procedure and some suggested tools for moving forward.

Types of Environmental Scans

There are essentially two types of environmental scans. The first approach is a less formal type of scanning that you conduct yourself, based on your own knowledge and what you or an assistant can gather sitting at your desk. The first step is to write out what you know about how others are dealing with similar situations. In other words, you are looking to see how others in your social environment do things. Generally, people who are more connected with their colleagues, who read the trade literature, and who regularly attend conventions and workshops tend to find this process easier.

A second part to this might involve a more formal review. Depending on the issue, you might seek out journal or news articles that have been written on the topic. A good place to start is to check the internet. Search engines such as Google, Bing and Webcrawler can retrieve a tremendous amount of information very quickly. One of the big challenges in using general search engines is that it is sometimes difficult to identify the exact search terms you need. Consequently, the search generates more chaff than wheat.

Using Internet Search Engines

There are some tricks to using search engines. If you are fortunate enough to have access to a municipal librarian or a local college or university library, there are usually experienced people who can provide some assistance. Some tips for narrowing Google searches are provided in the accompanying box.

Either online or by visiting a local library, it is also possible to search the professional literature. Trade magazines and journals often provide coverage of general issues.

For more detailed sources of information, it might be necessary to enter the formal research or academic literature. This latter step can be a little daunting at times since there is a lot of variation in how technical articles are written. Some are very accessible while others require extensive prior knowledge of the topic. The key is not to become discouraged.

Sometimes it is worthwhile looking further afield.

In this case, official websites of the US Fire Administration and the UK Government can offer a wealth of information. The box on the next page gives a brief listing of some of the major fire journals. A few key web links are also provided.

If you require information on characteristics of your community or other statistics, a great deal of information is available on the Statistics Canada website.

Effective Searches on Google

1. **Be Specific.**
   Find pages within sites using *site:[website URL]* and your search phrase, find authors using *author:[name]*, and type *intitle:[word]* to find a page with that word in the title.

2. **Format.**
   Use *filetype:[pdf or other extension]* to find images and all sorts of files (such as docs and jpgs).

3. **Broaden Your Search.**
   Use an asterisk (*) as a wildcard search operator to fill in the blanks. For example, "*Fire services* *"

4. **Limit your search by excluding unwanted terms.**
   Put a minus sign in front of terms you wish to exclude. For example, *alarms -burglar* will exclude the term "burglar" from your search. If you want to limit your search numerically, use the range (two dot) indicator. For example "*used trucks 2010 .. 2014*" will limit the listing to those years.

5. **Use specific search engines.**
   Google scholar, for example, is an excellent way to find both academic and other articles on selected topics. Webcrawler looks across a series of search engines. Also check out the website for Amazines (www.amazines.com) for a database of free articles.

Sources of information

Professional Journals

Alarm: Modern Fire Protection and Security Systems Bulletin
American Fire Journal
British Fire Services Association Journal
Building and Environment
Chartered Institute of Public Finance and Accountancy. Fire Statistics.
Actuals and Estimates (UK)
Combustion Science and Technology
Combustion and Flame
Engineering Structures
Experimental Thermal and Fluid Science
Fire and Materials
Fire Australia
Fire Prevention and Firefighter Safety Research
Fire Technology
Fire Risk Management
Fire Journal (Australia)
Fire Magazine
Fire Technology
Fire Safety Journal
Industrial Fire World
International Journal of Wildland Fire
Journal of Structural Engineering
Journal of Hazardous Materials
Journal of Applied Fire Science
Journal of Fire Sciences
Journal of Fire Protection Engineering
Journal of the Fire Protection Profession
Proceedings of the Combustion Institute
Professional Safety: Journal of the American Society of Safety Engineers
Progress in Energy and Combustion Science

Online Sources

Hemming Groups – suppliers of equipment: www.hemmingfire.com

Fire Information Group (UK): www.figuk.org.uk

A listing of free online fire journals: http://nysfireinfo.pbworks.com/w/page/27061249/Free%20online%20fire%20journals

Statistics Canada: www.statcan.gc.ca

UK Government Fire Incident Statistics: https://www.gov.uk/government/organisations/department-for-communities-and-local-government/about/statistics

US Fire Administration: www.usfa.fema.gov

While there is currently no single national fire database in Canada, individual provinces keep substantial amounts of information.

To do a scan most effectively, make sure you have collected information in more than one way. By doing this you can check and cross-reference to see if the same issues and concerns are surfacing through your various sources of information.

Occasionally, it is worthwhile conducting a formal process where others in the organization are involved. In this instance, you might consider bringing in an outside facilitator and conducting a formal scan. The process of doing a formal scan is outlined in the second part of the chapter on SWOT analyses. The primary difference between an environmental scan and a SWOT analysis is that the focus or range of issues considered by an environmental scan is generally much broader. SWOT analyses are generally limited to issues relating to challenges and opportunities confronting the organization.

Framing Your Environmental Scan

The information that you will be gathering is influenced by the question you are trying to answer. To frame the environmental scan, we can start by asking some focused questions, such as the following.

- What is the key issue?
- What do we need to know about the issue?
- What are the trends and drivers affecting these factors?

Once the question has been carefully framed, and the research has been gathered from primary and secondary sources, then the analysis begins.

First, we need to consider what themes, concepts, issues, or concerns surfaced in the secondary research. In other words, how have other groups, organizations, communities or fire departments been affected by this issue elsewhere? How have they ranked those concerns?

Compare the results of your surveys with the qualitative data that is emerging from your focus groups. Consider what people have been saying in the one-to-one interviews. What common themes are emerging? How are the results showing consistency and repetition?

Try to determine how these people have ranked the concerns that also showed up in your secondary research. Do they see it the same way? Or have they raised different thoughts, ideas, or concerns that haven't shown up in the secondary research?

Once you or your team has agreed on the ranking of the issues, beginning with the most serious and urgent, then you can begin to consider the strategies, program activities, and practice that will help you address them. You will also need to consider the budget implications involved in meeting these strategic priorities.

As we noted, environmental scans are often accompanied by a SWOT analysis, which determines the internal and external strengths, weaknesses, opportunities and threats that are affecting the organization's ability to fulfill its organizational mandate. The SWOT analysis is explained more fully in the second part of this chapter.

Example: Responsibilities of Australian Fire Services

In 2010, the Australian Bureau of Statistics (ABS) conducted an environmental scan of emergency services organizations in Australia.[1] These organizations included police and fire services and others. As we might expect, the original document was lengthy and highly detailed. Among the items addressed in the study, however, were the responsibilities of various agencies in different states. The information presented in the report was drawn primarily from the various state websites. Most environmental scans conducted by local agencies would not be as detailed as the one conducted by the ABS.

Let us assume, however, that you wish to know how your service mission compares with other agencies in other jurisdictions. Someone in your organization has suggested that it would be interesting to examine the range of activities for which Australian fire service agencies are responsible. You could do what the ABS did and scan the various state websites. Or, you could be fortunate enough, as we are in this instance, to find an existing document that already contains the raw material.

On the following page we have extracted the roles reported by fire services from three states only—New South Wales, Queensland and Victoria—to provide an example of what the results of a scan might look like. As noted, the original document reports on more agencies and is much more detailed. Based on this and other information, you might decide whether you might wish to refocus the service components of your own department. Clearly, several options are available. Depending on your department's locations (a major metropolitan area as opposed to a small, rural community) you might wish to broaden your range of activities. Another option would be to identify the four or five key areas in which all other departments engage and focus on those as your core functions.

Again, what you get out of an environmental scan is determined by the initial question you are trying to resolve.

Partial Environmental Scan of Fire Service Responsibilities for Three Australian States

New South Wales

- Fire services - managing fire emergencies in major cities, metropolitan areas and towns across regional and rural New South Wales.
- Rescue services - rescues at road, household and industrial incidents.
- Hazardous materials (Hazmat) - protecting 100% of the state population from hazardous materials incidents.
- Bushfire services - supporting the Rural Fire Service of NSW during and after bushfires in NSW and preventing bushfires through hazard reduction strategies.
- Urban search and rescue - urban search and rescue (USAR) is a specialist capability to locate, provide medical assistance to and remove victims who have been trapped or affected by a structural collapse.
- Counter-terrorism services - the Counter-terrorism and Aviation Unit manages the planning, development and implementation of counter-terrorism and aviation capability.
- Fire investigation - the Fire Investigation and Research Unit (FIRU) provides a range of investigative and research services to both internal and external customers including research into human and structural behaviour during fire, and the associated impacts for performance based building design.
- Building fire safety - the NSWFB plays an important role in building safety, both legislatively and as a community service.
- Community education - the NSWFB uses the expertise and experience of our firefighters to educate others in ways to prevent and prepare for emergencies.

Queensland

- Rescue - road accident and other types of rescue.
- Chemical and hazardous material management.
- Community awareness and education on fire and road safety issues.
- Building fire safety inspection, investigation and prosecution
- Administering legislation relating to fire and safety, hazardous materials facilities and hazard mitigation.
- Rural land management advice regarding the role and use of fire.
- Fire scene investigation.
- Alarm monitoring and response.
- Commercial training in firefighting, fire safety and evacuation procedures.

Victoria

Emergency response:

- Suppression of all types of fires;
- Urban search and rescue, including road accidents;
- Emergency Medical Response (EMR) First Responder Program;
- Emergencies on waters in Port Phillip Bay and the metropolitan river system
- Industrial accidents and hazardous material handling and storage incidents;
- Supporting other combat agencies in emergencies
- Chemical, biological and radiological emergencies; and
- Strategic, expert advice to the State Government on major events and anti-terrorist activities.

Non-emergency services:

- Input into the development of Australian Standards, Codes of Practice and Regulations affecting community safety;
- Delivery of community safety activities including education to increase awareness and preparedness;
- Conduct building regulation related inspections of fire and life safety systems and maintenance compliance;
- Development of fire safety and emergency plans for major events;
- Fire investigation and cause analysis, and the provision of data to the community and external authorities (within Privacy Act);
- Review and inspection of the dangerous goods handling and storage practices and fire safety systems of major hazardous materials sites;
- Representation on councils for fire prevention planning and community risk management;
- Attendance at and participation with local councils in municipal emergency management planning exercises;
- Provision of expertise, technical advice and skills acquisition services to interstate and international organisations;
- Servicing and sale of fire safety equipment; and
- Participation in appropriate fundraising activities.

SWOT Analyses

A SWOT analysis is an assessment of an organization's Strengths, Weaknesses, external Opportunities, and Threats. It can be a useful complement to the environmental scan. As we discussed earlier in the chapter, the environmental scan will provide you with primary and secondary information to determine pressing issues and concerns related to your research questions. When that information is combined with the results of the SWOT, you will be better equipped to identify your strategic priorities and future directions.

The SWOT adds to the results of the environmental scan by engaging various members of your organization in a discussion of the strengths and weaknesses that currently exist within your fire department. Looking outside the department allows you to consider opportunities that could be seized to advance the interests of the organization. The SWOT also explores threats: those external factors, realities, or trends that can make the ongoing functioning of the department more challenging.

A SWOT analysis is sometimes conducted as a group session with a facilitator. This might be preceded by a survey that each member of the group completes in advance, so they have a chance to consider their own assessment before group discussion begins. Even simpler, give each group member a blank SWOT template that they can use to jot down their thoughts in advance, and then have them bring it to the meeting.

Conducting a SWOT Analysis

While it is possible to conduct a SWOT analysis by yourself, the real benefit of the exercise is usually seen when several members of the organization are involved. One of the paradoxes managers face is that on the one hand, employees and others expect leaders to lead but, at the same time, they expect to be part of the decision-making process.

As with any activity, consultation has a price. While employees are being consulted they are not doing their normal activities. Furthermore, group dynamics can generate unexpected results. Group politics come into play and red herrings can occupy a significant amount of time. For those reasons, it is often beneficial to have an outside facilitator lead the exercise. The advantages of consultation, however, are numerous.

First of all, groups tend to generate crucial ideas that a single manager or even a management group might overlook. Second, people from different segments of the organization interact with different audiences, suppliers, clients or customers, regulators and, sometimes, competitors.

Example: Completed SWOT Analysis

Strengths
- High staff morale
- Effective leadership
- New Equipment

Weaknesses
- Aging staff and pending retirements
- Recruitment challenges
- Lack of diversity

Opportunities
- Collaboration with neighbouring community concerning disaster management planning.
- Access to senior government funds to expand disaster response strategies.

Threats
- New city councillors who are not well informed of the issues.
- Municipality plans to implement significant across-the-board budget reductions.

This gives them different perspectives on the organization, particularly with regard to outside influences. Third, even participants who do not see their input reflected in the final product generally feel they have had some say in the process. This typically has a positive effect on morale and often creates more "buy in" when choices have to be made and different options are implemented.

In a group situation, one of the first questions when conducting a group analysis is: Who will participate? It is helpful to have a diverse cross-section of individuals to ensure the most comprehensive assessment. While no guarantee, this helps to increase the likelihood that no crucial aspect is overlooked. As a general rule, the SWOT analysis should be done by no less than mid-level management, and preferably even a higher level of leadership. In addition, the analysis should include representative employees from throughout the organization. Front-line supervisors should be included. Again, while not always the case, leaders in the organization often have greater insight into those external and internal issues that need to be considered. This comes from their experience as well as their relationships with a wide variety of people inside and outside the organization.

Before starting the analysis, it is often worthwhile providing the team with the environmental scan results to read in advance of the SWOT analysis meeting. Ensure you include the guiding research question that is behind the environmental scan and SWOT process, as that will create the framework for the discussion. Create helpful ground rules for the discussion.

SWOT Discussion Ground Rules

- Focus on one quadrant at a time.
- Listen to understand, and acknowledge what you are hearing others say. Avoid interrupting or criticizing the contributions of others.
- Establish reasonable time limits to keep the discussion moving forward.
- Respect each other—it's acceptable to have differing points of view and perspectives
- Agree on how distractions will be managed, for example, cell phones, interruptions from support staff, etc. It is suggested that cell phones be turned off and administrative staff interrupt only for emergencies.
- Confidentiality: What can be shared outside the room? Where will the information go at the end? How will anonymity be protected?
- All team members should participate.

As the group considers the issues and concerns that have resulted from the environmental scan, ask them to consider each quadrant in turn as a means of assessing how those issues and concerns could be more fully addressed or understood.

As you proceed through your SWOT analysis, keep these factors in mind:

- SWOT analysis is a subjective process, not a science. However, the quantitative and qualitative data that emerged from the environmental scan will help the participants trust that the results are well-founded.
- Keep it simple, focusing on a few issues only. If other issues emerge, they can be addressed later through a subsequent process. Without these limitations the process may bog down with too much data and information to be dealt with at one time.
- Be realistic about the strengths and weaknesses of the organization. Create safety and transparency so participants will be honest.

In summary, the SWOT analysis combines with the environmental scan to create strategic plans that are realistic, researched, and supported by internal personnel and external stakeholders. Evidence-based decision-making benefits from using tools such as these, leading to plans and decisions that will be solidly grounded in facts and research, and guided by a wide array of perspectives and input.

Notes

1. Australian Bureau of Statistics (2010) NDMP Data Dictionary Project Reference Guide of Phase 1 Attachment 1: Environmental Scan/Literature Review. http://www.fire.nsw.gov.au/gallery/files/pdf/projects/data/attachment_1.pdf

   See: http://www.fire.nsw.gov.au/gallery/files/pdf/projects/data/summary_report.pdf for full report.

Statistics

A Tool for Decision Making

Statistics is probably one of the most misunderstood of disciplines. Most university students dread having to take it, and most professors who teach it often do so with great reluctance. Furthermore, the topic is often reviled as a tool of charlatans. As Mark Twain once claimed, "There are lies, damned lies and statistics." Yet, used appropriately, statistics can be one of the most useful and powerful tools in the decision maker's toolbox.

My suspicion is that statistics' bad name stems from two sources. First, many people see it as an outcropping of math—with which most of us had a less than excellent experience in high school. Second, most people who teach statistics are not themselves statisticians and, while they may come to master the technical details, they rarely grasp the underlying logic. Statistics does entail some math, but most of that math is no more complicated than being able to balance one's chequebook. The key to understanding statistics is to see it as a way of organizing and making sense of a world dominated by uncertainty. In fact, one definition of statistics is that it is the *science of decision making under conditions of uncertainty*.

The key to understanding statistics is to see it as a way of organizing and making sense of a world dominated by uncertainty.

What is key for most decision makers is not to get tangled in the details of statistical analysis, but, instead, to understand the fundamental principles or logic behind the activity. Those fundamental principles are few and, generally, quite simple. Once understood, however, the principles of statistics can be used to great advantage, even if one doesn't have a detailed knowledge of the underlying math or technical aspects.

Statistics consists of two basic activities. The first is the collection of data in an attempt to describe something. The second is the use of data to help make decisions or inferences. The first activity we call descriptive statistics; the second, we call inferential statistics.

A Discussion of Measurement

We often refer to the process of observing and recording data as a measurement. What distinguishes the way statisticians view measurement from most other people is that statisticians assume all measurement contains an element of error. In other words, in the world of statistics, having something measured with one hundred per cent accuracy is more good luck than good management. From a statistical perspective, error in measurement has two basic sources: inherent error or instability, and operational error.

When we speak of inherent error or instability, we are referring to the property of the thing we are measuring. For example, if you were to ask someone to tell you on a 100-point scale how satisfied they were with their job (assuming 0 is total dissatisfaction and 100 represents total satisfaction), they might respond 71. If you asked the person the same question on several different occasions, they would likely give you a range of answers somewhere close to 71.

The reality is, most people have a general idea of their level of job satisfaction but have a hard time giving a precise number. Furthermore, while they may be mostly satisfied with their job, their exact level of satisfaction would vary according to numerous factors ranging from the time of day, to whether they just had an altercation with their superior, to the weather.

Inherent error relates to what we are measuring—e.g. a length of hose, which may be affected by temperature.

Operational error relates to how we are conducting the measurement—e.g., a problem with the measuring device or how we read it.

While relatively stable in a range, most people's actual level of job satisfaction is inherently unstable.

The same applies to physical measurements. Assume one has a metre of hose. If we measure the hose several times, the actual length will vary around one metre as the hose shrinks or contracts due to temperature fluctuations, how tightly we stretch it out, and so on. The measurements may be close to a metre (even within fractions of a millimetre), but it is unlikely that on repeated measures they would turn out to be exactly one metre. Perhaps if we could control the tens or hundreds of factors that influence the fluctuations in the length of the hose we could get an exact measurement. In reality, we can rarely do this.

Consequently, from a statistical perspective the length of the hose is inherently variable.

To the notion of inherent variability, we can also add operational error. Our device for measuring the hose might be off a little. The ruler may slip. We may make an error in copying down the observation and record one metre and eleven millimetres instead of one metre and one millimetre; or, between making the measurement and recording it, we might forget the actual number. Occasionally, we well get it right on and the measurement error will be zero. Furthermore, unless some bias exists in our measurement process—a tendency to consistently under or overestimate the measure—repeated measures of the same item will average out to the true measure.

The point is that, try as we might, it is generally difficult, if not impossible, to have totally accurate measurement.
Believing we can do so is simply fooling ourselves. Furthermore, for most situations, "close" is good enough. What does it matter if the hose is a millimetre longer or shorter than a metre? One thing that makes statistics powerful is that statistics assumes some error will appear in our measurement.

What is also great about statistics is that, when used appropriately, we can estimate how much error exists in the measurement process.

From the statistician's perspective, people who believe that total accuracy in measurement is possible are like ostriches with their heads in the sand. It is far better to admit that error in measurement is everywhere, so why not admit it and try to get an estimate of the size of that error? How can we do that? The answer is that we need to either take several measurements of the same item, or to measure several items assumed to be the same.

From the statistician's perspective, people who believe that total accuracy of measurement is possible are like ostriches with their heads in the sand.

Descriptive Statistics

Remembering the characteristics of a single item is relatively easy, whether that item is a person, an event like the eclipse of the moon, or a fire truck. Similarly, most of us can easily recall the characteristics of several items. The larger the number of items becomes, however, the more difficult it is for us to remember the specific individual items that make up the group. For example, we may recall the ages of all of our colleagues in a fire hall. Recalling the age of all firefighters in a region is virtually impossible. If we want to be able to say something about the ages of firefighters in a region, we need to somehow aggregate or summarize the data. This is where descriptive statistics come into play.

What descriptive statistics do is summarize the characteristics of a group so that we can make sense of a mass of information. Even if we could remember them, listing the ages of 600 firefighters is not a very useful exercise. Descriptive statistics allows us to identify certain useful characteristics of the list. Often, the first two things we want to know about a list or bunch of observations are what is typical and how much variability is there?

The most common measure of typicality is the arithmetic average or mean.

> Descriptive statistics summarize the characteristics of a group so we can make sense of a mass of information.
>
> We may measure *typicality* by determining the average or median age in the group.
>
> We may measure *variability* by determining the youngest and oldest ages in the group, the spread of ages within the group, or how much the results deviate from the average.

We might find, for example, the average firefighter in our region is 38 years of age. Other measures of typicality include the median and the mode. The median is that point in the age distribution below and above which half of the ages fall. The median age might be 35. In other words, half the firefighters in our region are above age 35 and half are younger. The mode is another term for the most common age. The mean, the median and the mode are the most commonly used measures of typicality. We can also think of those measures as a central anchor point for the list or distribution of ages.

Measures of variability give us an idea of how widely a bunch of measures range or vary. It is one thing to know that the average age of a firefighter in our region is 35; it is something else to know that most are between the ages of 30 and 40 as opposed to 25 and 50. The most common measures of variability are what we term *range* statistics and *variance* statistics.

Range statistics are simple measures of the distance between two points. For example, among our firefighters, the youngest may be 24 and the oldest 58. The range would simply be 58-24, or 34 years. This range measurement is based on the difference between the minimum value in the distribution and the maximum value. Min-max ranges are interesting but can sometimes be misleading. For example, the oldest person in a region might be 65 while most of the other "elderly" employees are less than 55. Here, we sometimes call the 65-year-old an outlier.

To deal with distributions that have the odd extreme case, we sometimes use a statistic known as the *interquartile* range. To get the interquartile range, we need to figure out the age of the person who is at the 25th percentile point of the distribution, and the age of the person who is at the 75th percentile. The interquartile range is simply the difference between those two numbers. Again, like the min-max range, the interquartile range gives us an idea of the spread of the ages.

Besides ranges, we often use statistics known as *variability* statistics to give us some notion of how the data are spread or disbursed about the measure of central tendency. The two most commonly used variability statistics are the variance and something called the standard deviation. At first sight, these statistics may appear a little daunting but conceptually, they are quite simple. The key in understanding them is not to focus on the math but to consider the underlying ideas.

See the following pages for examples of typicality and variability.

An Example of Typicality

Even simple descriptive statistics can be useful in decision making. Let's examine the case of two fire stations. Each got six calls for service last Monday. The response times for each call is shown in the accompanying table.

For measures of typicality, we can calculate the average or arithmetic mean, the median and the mode. The average or arithmetic mean is simply the sum of the response times divided by the number of calls. The median is that point below and above which 50 per cent of the numbers fall. The mode is the most commonly recorded response time.

	Station No. 1	Station No. 2
	5	3
	6	5
	7	5
Median Point		
	7	8
	8	9
	9	12
Sum	42	42
Mean	7	7
Median	7	6.5
Mode	7	5

The data in the boxes represent the actual response times of the calls in minutes. Even from this limited amount of information, there are several points of interest. First, both stations took a total of 42 minutes responding to the call. This resulted in an average or mean response time of 42/6 or 7 minutes.

Examining the numbers, however, it appears that Station No. 2 had one call where the response time was 12 minutes. In statistical language, we call exceptional values such as this outliers. The arithmetic mean is very sensitive to outliers. This is easy to visualize if we replace the 12 with a value of 20. All the other values stay the same but the mean would shoot up to 8.3 minutes.

A measure that is much less sensitive to outliers is the median (or midpoint, as it is sometime called). As we have noted, the median is the value that breaks the distribution into the upper and lower 50th percentile. In the table, the median or midpoint is indicated by a break in the column listing. For Station No. 1, the median or midpoint of the distribution is 7. For Station No. 2, the median is half way between the values of 5 and 8, which is 6.5.

That Station No. 2 has a lower median than mean is a consequence of the fact that, except for the outlier value of 12 minutes, Station No. 2 generally had lower response times than Station No. 1. Because we are only dealing with a small number of values, this is easy to see. It would be less obvious with a large data set. Regardless, the principles hold.

An Example of Variability

In this example, we will use the response time data from the previous box. We have seen that the typical or average response times are about the same for both stations. However, looking at the raw data suggests that there might be more variability in the response times in Station No. 2 as opposed to Station No. 1. The fact that the mean and the median were slightly different provides numerical support for this view.

	STATION NO. 1		
	Time	Deviation from Mean	Deviation Squared
	5	-2	4
	6	-1	1
	7	0	0
	7	0	0
	8	1	1
	9	2	4
Mean	7	0	1.7

	STATION NO. 2		
	Time	Deviation from Mean	Deviation Squared
	3	-4	16
	5	-2	4
	5	-2	4
	8	1	1
	9	2	4
	12	5	25
Mean	7	0	9

One measure of variability is the range. Station No. 1 response times go from a minimum of five minutes to a maximum of nine, providing a range of four. Station No. 2 response times go from a minimum of three minutes to a maximum of 12 minutes, providing a range of nine. This supports our intuition.

Another two commonly used measures of variation are the variance and the standard deviation. While seemingly complex, these measures are conceptually simple. In the second column of numbers, we have subtracted the mean from each individual response time. For example, in Station No. 1, the first deviation is 5-7=-2. We do that for each of the individual response times.

In column three, we simply square the deviations from the means (that is, multiply the value by itself). When we do this for all of the observations, we discover two things. First, the average of the deviations from the mean is zero. This will always be the case because the mean is in the "middle" of the distribution and the positive deviations will cancel out the negative ones. That is why we calculated the third column: the squared deviations.

The mean or average of the squared deviations is known as the variance. The variance for Station No. 1 is 1.7 and for Station No. 2 it is nine. This suggests that there is much more variation in the response times of Station No. 2 than for Station No. 1. The variance is a statistic that is used a lot in statistics. In slightly more advanced statistics, our goal is to try to explain why there is more variance or variation in one set of numbers than another. Perhaps, across the two stations, the traffic patterns are substantially different. The difference might also be due to variations in the performance of personnel. Those are notions or hypotheses we might want to test.

Since squared values generate big numbers, we often compare the square root of the variances. This brings the values back to the size of the original measurement (raw numbers as opposed to squared ones). The square root of the variance is known as the standard deviation. The standard deviation for Station No. 1 is 1.3 and for Station No. 2 it is three. This suggests that the variation in the response times in Station No. 2 is slightly more than twice that of Station No. 1.

An Aside

We can also use variance statistics as an estimate of how much error in measurement exists. For example, two crews may take 23 minutes on average to complete an activity. The variance for one crew might be eight minutes and for the second crew three minutes. Based on the average both crews appear equal in performance, but the variance measures suggest that the second crew is much more consistent and, in that sense, better. From a management perspective, the interesting question is why one crew is more consistent in its performance than the other.

Subsequent investigation may show that the first crew has to perform the action under a variety of conditions while the second crew faces fewer environmental challenges. It may also be that the first crew lets things "slide" for a while and then turns on the juice to get their numbers back up to an acceptable average.

Regardless, knowing differences in variances can sometime tell us more than simply knowing differences in averages or central tendency.

Inferential Statistics

The second leg on which the discipline of statistics stands is what we term inferential statistics. Inferential statistics help us to draw conclusions and make decisions. Unlike for most descriptive statistics, the math behind inferential statistics can get complicated. Consequently, we will restrict our focus to the logic underlying inferential statistics and examine how they can be used to help us make decisions. Learning inferential statistics by oneself from a book is typically not easy.

For readers who have no background in the area, it might be worthwhile investing is a one-semester course in a local college. Otherwise, understanding the concepts is sufficient; just leave the details to an expert.

Inferential statistics are used for many purposes. However, the two primary ones are to be able to estimate or infer the characteristics of a population from a sample, and to estimate whether significant differences exist between two or more populations or samples.

Population Estimates

Let's start with the issue of making inferences from samples of populations. If we wanted to know the proportion of the population of a city that uses space heaters, we could contact each household and pose the question. Collecting information from everyone in a jurisdiction is known as conducting a census. In a city of 300,000 households, that could be an expensive and time-consuming proposition. That is why censuses are done only rarely and under limited circumstances. Fortunately, early in the twentieth century, statisticians figured out how to estimate the characteristics of the whole (a population) from a sub group or sample.

The key to being able to do this, however, is in the way in which the sample is drawn or collected from the population. Essentially, "any old sample" doesn't cut it. The sample has to be taken from the population in a particular way. There are some variations on the theme, but let us keep this simple and consider the basic case. What we want is something statisticians call a simple random sample. A simple random sample is one where each household in the population has an equal chance of being selected, and that chance of being selected is independent of the other selections. Let us break that down into the constituent parts: random selection, equal chance, and independence.

Random selection

This implies the households in the sample are chosen using a *chance* mechanism – things like coin tosses and computer random number generators. In other words, someone cannot choose the households based on availability or door colour. Random selection implies that a listing of households (say a city directory) exists where the households are listed or numbered from 1 to 300,000. For a sample of 1,200 households, we would use a random number generator to give a listing of 1,200 numbers between 1 and 300,000. Once we have those numbers, we would then identify the households that hold those positions or numbers on the list.

Equal chance

This implies that each household has the same chance or likelihood of selection. Lists with duplicate addresses or lists that omit certain a type of household (say, all apartments or all households in a particular neighbourhood) mean some households either have a greater likelihood of selection, or no chance of selection.

Independence

This implies that the selection of one household does not determine or affect the selection of another. For example, the person selecting the sample might notice two houses on the same block or two houses next to each other appear on the list. Thinking they might be too much alike, s/he drops one household in favour of another selection. That is not acceptable. The selections that appear must be included despite anything else.

If we follow these rules, then estimating the characteristics of the entire population from the sample is possible. Some other things need to be considered, such as the size of the sample, but those are details that are best discussed with a professional. If we follow the basic rules outlined above, we can estimate what proportion of the population of households uses space heaters within a given likely range.

In other words, the sample estimate will be close to what actually exists in the population but will probably not be the exact figure. What differentiates statistical sampling from other procedures, however, is that it is possible to estimate the range within which the population figure will likely fall. Thus, we could conclude that the likely proportion of space-heater-using households we would see is X percent within plus or minus Y percentage points in, say, 19 surveys out of 20.

The uninitiated often disparage statistical estimates for not being able to provide exact values. But, as we discussed earlier, the fundamental assumption in the world of statistics is that all measurement entails error, so the best we can do is come up with a point estimate and a reasonable notion of its level of accuracy. This is something no other procedure can do. With a statistical estimate, you get an idea of whether an estimate is precise enough to be useful or too variable for practical purposes.

Many different ways of generating estimates are available, but you have no way of knowing if they are close to the actual value in the population or somewhere out near the planet Mars.

Significant Differences

Another primary use of inferential statistics is to be able to estimate whether two samples are similar or different. For example, over a year, a Fire Chief might wish to know whether differences in response times exist across fire halls. Typically, data such as response times are collected though an automated dispatch system. At the end of a period, calculating the mean or average response time is possible. As discussed earlier, the mean value will be an estimate based on error-prone data and there will be a distribution of values around that estimate. Thus, the question is, if the response time of one hall is eight minutes and another one is nine, does that one minute difference reflect a real difference or is it simply within the realm of possible measurement error?

Some differences are big and substantively meaningful and do not require statistics to help us make a decision. For example, if the difference in response time were 10 minutes, then we know a real and important difference exists. However, when we get to one minute, it is not clear that the difference is real or just within the realm of normal variability.

What statistics can do is let us know whether that difference is within or outside that range of normal variability. If it is outside, then we say that the difference is *statistically significantly different.*

We should note, however, that just because something is statistically significantly different, it does not necessarily mean that it is substantively different. On the other hand, if something is not statistically significantly different, then we should assess the difference as being within the normal range of variation and, consequently, not substantively significant either.

Inferential statistics are of even more use when we have multiple comparisons to make. Typically, a city may have 10 or more fire halls. Are the differences across all 10 significantly different? More advanced techniques can help us to figure out what factors might be related to those differences. That brings us to our final topic in this chapter and that is the role of statistical modelling.

Statistical Modelling

For most decision makers, the real power of statistics lies in the ability to model social, natural and mechanical processes. Statistical models allow us to examine complex processes where multiple factors might affect a particular outcome. For example, statistical models have been used to model response times to incidents allowing for traffic and weather patterns. Based on the result of those models, optimal placing of fire halls and response routes can be determined.

Similarly, the physical and socio-demographic characteristics of neighbourhoods can be used to model which neighbourhoods or what types of dwellings are most likely to experience fire incidents.

Models can also be used to find whether changes in the number of personnel dispatched or whether certain mixes of equipment are more effective than others.

Currently, one of the more active and dynamic areas of modelling is in trying to predict the behaviour of forest fires. Forest fires and brush fires have always been particularly unpredictable and dangerous phenomena. With increases in global warming, it is likely that forest fire incidents will increase in frequency and severity. Places where forest fires have been relatively rare events will likely see increases.

Traditional research has focussed on the impact of terrain and weather on the development and evolution of fires. More recent activity involved modelling the impact of volatile organic compounds given off by different plants on the intensity of the fire.

In statistical modelling, most of our focus is on trying to explain variation. Thus, we go back to one of our basic statistical concepts—that of the variance. So, for example, we might wish to ask: What are the factors that likely affect the different variations in response times between Station No. 1 and Station No. 2?

Based on the outcomes of those and other modelling exercises, it is possible to identify what form of intervention works and what doesn't work. Improving the allocation and efficiencies of their resources is also possible for decision makers.

Whatever the complexity of the model or underlying process, statistical analyses enable us to figure out with an estimable level of accuracy, many useful results. Among the key questions that we can answer are the following.

- Does the overall model accurately reflect the process we are trying to describe or emulate? In other words, is it statistically significant?
- How much of the variation in the outcome factor is explained by the model?
- Which elements in the model are statistically significant and which are not?
- What is the relative impact or rank ordering of various components of the model on the outcome factor?
- Are those impacts large enough to be meaningful from a substantive or policy perspective?
- How do the various sub components in the model interact with one another as to their impact on the outcome?

As we indicated, statistics is not the magic bullet for all decision making. Used appropriately, however, statistical techniques can provide a great deal of insight into the questions we are examining.

Decision making is a complex process, and the best processes are those where we use the many tools at our disposal to help come up with an answer. Often, trade-offs have to be made. Something may be statistically significant but not substantively significant. Similarly, just because one choice is more effective than another doesn't mean that it can be justified socially or economically. Regardless, knowing whether something has a "real" impact or not is a good starting point.

Experimental Designs

How Do We Know What it Means?

A basic notion underlying this book is that making decisions based on evidence has advantages over other forms of decision making. By *evidence*, we are referring to observable and measurable "facts" or data. While we argue that it is generally a good thing to have facts, a single fact or bit of data or piece of information is fairly meaningless in itself. The reason for this is that *nothing has meaning except in comparison with something else*.

For example, assume you are on a trip to India and you see a pair of shoes on sale for 2,859 rupees. If you are not familiar with prices in India, you might ask yourself whether this is a good price or not. The "fact" that the shoes are 2,859 rupees is irrelevant to you unless you have something with which to compare it. That comparison might be with another product or with the average hourly wage in India or with the equivalence in another currency. Currently, 2900 rupees is approximately equivalent to $50 Canadian. It is only by making a comparison that the relative value of the shoes takes on meaning.

To understand the meaning of a fact, we need an appropriate point of comparison.

Similarly, you might discover that a source can supply water at a pressure of 100 psi. Whether that is good or bad, useful or useless depends on a context or point of comparison. For example, one-inch attack hoses generally require 200 psi to be fully operable. By having this point of comparison, we can evaluate the value or meaning of a source at 100 psi. In another context, an appropriate reference point might be that most residential water supplies are in the 60-80 psi range and will only occasionally reach 100 psi. A pressure source of 100 psi may be adequate for watering lawns but typically leaves something to be desired when fighting substantial fires.

The point being made is that to understand the meaning of a fact, we need an appropriate point of comparison.

Within the framework of evidence-based decision making, a key question we have to ask ourselves is: What is the most appropriate point of comparison? A complementary question might also be: What is the best way in which to make that comparison? The answer is to use a standard framework that program evaluators and applied scientists call experimental designs. Experimental designs are simply different approaches to helping us make an appropriate comparison.

The remainder of this chapter will focus on some basic experimental designs that we use to assess the value of information or data related to a question about which we need to make a decision. In applied research, designs can become very complex. No matter the complexities of the design, however, there are a few fundamental principles that underlie the value or the merits of the design.

The "Counterfactual"

When we do or observe something, the question is: What would have happened if the event had not occurred? What if the Axis powers had won WWII; what if the party in power had not won the last election? What would have happened if insurance companies provided all fire services instead of municipalities? The comparison is with some theoretical model. It cannot give us proof of something, but as a mental exercise, it forces us to identify the important elements of a policy or program. What are the relevant or active components that are making the difference or that we expect to have an impact?

Einstein referred to this mulling of counterfactuals as *thought experiments*.

Thought experiments consist of conducting an analysis in our heads to think through the potential impacts and consequences. What differentiated Einstein's thought experiment from simple fantasizing or theorizing is that he also focussed on how we might test the thought experiment using real situations and observable data.

As an example of a thought experiment, we might consider the issue of putting out a small fire. We know that dousing a rubbish fire with water will put it out. Still, how does water do this? Is it because of its properties as a liquid? In our thought experiment we might consider using other liquids such as gasoline.

Experience tells us that does not work, so clearly it is not the property of being a liquid that is important. Maybe it is the "smothering" effect of water that is important. We consider other elements or situations that smother or exclude oxygen. We recall that putting a lid on a pan of burning oil, or using baking soda on a small fire, works to extinguish the flame. Through this thought experiment we conclude the principle is that we don't need a liquid to prevent further combustion; rather, we need something to cut off the source of oxygen or something that absorbs oxygen.

By thinking it through, we have come to a conclusion that makes sense. In itself, though, what makes sense logically does not always work out in the observable world. What we need is hard evidence based on repeatable observations—evidence that lies not just in our heads but evidence that can be seen, shared and evaluated by others.

What Makes Up Good Evidence?

When we engage in evidence-based decision making, the fundamental question is: What makes up appropriate evidence? If we think of science as a mechanism for finding the "real" explanation of something, then thinking of it within the context of a court case is possible. In the courts, as in science, there are varying amounts of evidence provided.

What makes sense logically does not always work out in the observable world. What we need is hard evidence based on repeatable observations.

Even if it is fundamentally true, we perceive some evidence as more valid, more reliable and more relevant than others. So it is in science. Good evidence stands up to the rigours of a good cross examination. Still, what makes up good evidence?

One characteristic of good evidence is how rigorously people have tested it. Within the framework of science, the basic mechanism for testing an idea is the experimental design. Experimental designs are physical applications of logic, so let us examine the logic underlying experimental designs.

Assume for a moment that we wish to assess the impact of residential smoke alarms on death rates in fire incidents. One approach would be to take a community and install smoke alarms in all residential structures. We could then see if a difference existed between the rates of death before and after the introduction of the alarms. Unfortunately, any difference might be the result of other factors (recall our previous discussion of spuriousness).

For example, by coincidence, rates of smoking in bed might have dropped, or the winters might have been milder so fewer people were using space heaters. We know both factors are related highly to fire incidences.

Ideally, we would like to be able to observe the same community with and without smoke alarms simultaneously. In other words, we would assess the effect of a smoke alarm program based on the difference in outcomes for the same community with and without participation in the program. Nevertheless, we know that this is impossible. Something cannot be in two states at the same time. At any moment the community either participated in the program or did not participate. The inability to observe the same entity in two different situations simultaneously is known in science as "the counterfactual problem." That is, how do we measure what would have happened if the other situation had existed?

If we cannot assess what would have happened if the opposite or counterfactual situation occurred, then how can we decide if smoke alarms have an impact and not something else? The approach scientists and program evaluators take is to find a comparison group that is as close to the treatment group as possible. How close that comparison group is to the treatment or experimental group determines how much credibility we can have in our results.

The inability to observe the same entity in two different situations simultaneously is known as "the counterfactual problem." That is, how do we measure what would have happened if the other situation had existed?

There are many ways of finding or creating comparison groups, some of which are better than others. The adequacy of a comparison group is something that evaluators spend much time and energy considering.

For example, we might find a "sister" community not far from the target community and use that as a comparison. On the other hand, we might decide to hand out smoke alarms to every second residence, or to residences on the south side of the community but not on the north side. We might even consider comparing our target community with all of the other communities in the province or state. All of those approaches and more provide a point of comparison against which we can judge the potential impact of smoke alarms in the target community.

The problem, however, is that all of those options have possible limitations. Some conditions or circumstances make the target and the comparison group inherently different. Sometimes we can see those differences. For example, in selecting a "sister" community, it may be that the residences in that town are older and built to different code standards. That might be an obvious difference, even to a casual observer. Often, however, the differences are not obvious.

The remainder of this chapter will focus on the different ways we might identify valid comparison groups to accurately reproduce or mimic the counterfactual. Identifying such comparison groups is the crux of any impact evaluation, no matter what type of program we are evaluating. Simply put, without a valid estimate of the counterfactual, we cannot establish the impact of a program with any degree of certainty.

Comparisons With Targets (The One-shot Test)

One of the simplest designs we have is to compare our population of interest with a particular goal or standard. Often, policy guidelines are based on legislated standards or targets set from studies of best practices. Targets can vary according to the context. For example, a community might target a 25 per cent reduction in incendiary incidents over a five-year period. A truck parts manufacturer may implement a six-sigma regime, where one expects that fewer than 3.4 defective parts per million will be manufactured. Human resource policy may also dictate that organizations should strive to hire a certain percentage of individuals belonging to minority groups.

The key, then, is to compare our population of interest with a target that is theoretically doable or achievable.

Once we implement an action, the question becomes whether we have met the target or goal. If we achieve the target, we have reason to believe that the action (which is generally a policy or program implementation) has been successful. Of course, we will use a statistical procedure to help us determine whether we are close enough to the target to be equal to the target.

The methodological literature sometimes calls this approach the *one-shot test*. That is, an action, policy or program is carried out, compared with a standard and, if it meets the standard, we generally assume the action was successful. The evidence might seem reasonably convincing. Unfortunately, one-shot tests have their limitations. We can see one major limitation in the following example.

The one-shot test does not account for alternate explanations for a result.

Example: One-shot Test

Suppose a community has a fire death rate of nine per million population and wishes to reduce it to five per million over a three-year period. The Fire Chief might decide that handing out free smoke alarm is the most cost-efficient way of achieving this goal. He carries out the program and three years later, the death rate is 5.1 per million which, given the size of the community, is statistically equivalent to the target of 5 per million. Can we infer that the smoke alarm program is behind the reduction in fatalities? The evidence seems compelling.

In fact, an alternate explanation for the reduction might exist. The free smoke alarm campaign generated substantial publicity in the local press. Firefighters and volunteers went from door to door distributing the smoke alarms. A notice left at the door asked citizens not at home to pick them up at various retail outlets. Together, the campaign generated substantial awareness of issues relating to residential fire safety. Because of the publicity, people in the community became more aware of the need for fire safety and made other changes in their homes. Some cleared clutter from around furnaces, fewer people used space heaters after going to bed, and more people planned escape routes should fire occur in their houses.

In other words, by heightening awareness of domestic fires, the community members took actions that would have reduced the likelihood of fatalities regardless of whether they had installed the smoke alarms.

The point here is not to argue that smoke alarms do not work in reducing fatalities. The point is that there may be alternate or coincidental explanations why the target was met. How much credibility those alternate explanations might have depends on different factors. First, does it make sense logically that the alternate explanations might hold? If previous publicity campaigns resulted in no noticeable impact then we might wish to stick with the smoke alarms as the effective mechanism. On the other hand, if publicity campaigns in other communities had resulted in substantial drops in death rates, we might be more supportive of the alternative explanation. A further explanation might be that fire death rates were declining overall for a variety of reasons, such as longer-term changes in building code, overall heightened awareness, decreases in smoking rates, and so on. Consequently, the death rate would have declined regardless.

Before-and-after Designs

A variation on the one-shot or target design is the *before-and-after design*. Again, we have a group or community of interest where we are looking to make an impact. We measure the situation beforehand, apply some intervention and then look at the outcome later. The assumption here is that any difference between the after and before results is due to the impact of the intervention. Unlike the one-shot design where the comparison is a policy goal or target, the implicit comparison in this design is the after results with the before baseline.

The before-and-after design shares most of the strengths and weaknesses of the one-shot design. Specifically, we can never be sure if it is the intervention that had an impact or simply some coincidental effect. For example, a jurisdiction might want to reduce the automobile accident rate among young drivers. The way they decide to do this is by dropping the legal Blood Alcohol Concentration limit from .08 to .05 for drivers under the age of 25. Examining the data from the three years before the introduction of the legislation with the data from three years after, an evaluator notices that accident rates have indeed dropped for younger drivers.

Again, we might consider the change in legislation to be the precipitating factor. On the other hand, it is possible that rates of drinking and BAC levels among young drivers have not changed. The difference is simply due to the increased vigilance of the police, who are targeting younger drivers in an attempt to enforce the new legislation. It is likely similar police vigilance without the change in legislation would have produced similar results. That is, the important factor is not the legislation, but simply enhanced surveillance by the police that serves to act as a general deterrent to young drivers.

Looking Past the Limitations

The limitations of these designs does not mean the evidence collected is irrelevant. We would have good reason to believe the results if we impose these interventions in many communities and under different circumstances with similar outcomes. Also, carrying out an intervention and then revoking it can tell us a lot. If the intervention results in the desired outcome and the revocation results in the original baseline outcome, then we have a more powerful argument that the intervention is the causal factor. What we need to remember is that evidence is rarely absolute. It has varying degrees of reliability or credibility associated with it. Just as in the courts, some forms of evidence are more credible than others.

Given the inherent weaknesses of these designs, we might ask what approaches we can take to address the problem. So far, the gold standard among evaluators and scientists is the *classical experimental design*.

The Classical Design

A rule of thumb in science is that nothing is perfect and certainty is an illusive goal. On the other hand, a lack of certainty in one's death is rarely a reason for playing Russian roulette. Similarly, a one per cent risk that one will lose all of one's assets in the stock market generally results in a different form of investment behaviour than if the risk is above 80 per cent. So, if we do not have perfection, what is the current ideal or gold standard for experimental designs?

To date, evaluators and scientists have relied on the two-group, before-and-after design to provide the most valid and the most reliable evidence. We start with the before-and-after design mentioned above. We then complement it with a comparison or control group that serves as the counterfactual. In other words, we have one group exposed to a treatment and one group that is not. If the group exposed to the treatment exhibits a significant change and the comparison group does not, then we have very strong reasons for believing the intervention had an impact.

The key to the strength of classical design experiments is to ensure the comparison (control) group is equivalent to the experimental group.

The key to the strength of this design is to ensure the comparison group is equivalent to the experimental or treatment group. This harkens back to our earlier discussion of the counterfactual where, ideally, we would like to see the same elements exposed to the treatment and not exposed simultaneously. This situation is physically impossible. However, we can ensure that both the treatment and comparison groups are initially as alike as possible. How do we do this?

One way is to take pairs of identical people (or communities or what have you), and divide them into two groups. However, unless the pairs are exact clones, we can never be certain that they are identical on all relevant characteristics. Fortunately, while we can rarely work with clones or identical matches, we can divide subjects into two *statistically* equivalent groups. As we have noted previously, statistically equivalent does not mean truly identical, but it does mean that, on average, no statistically significant difference exists between the two groups. In other words, for all practical purposes, they are close enough to being identical.

The method for ensuring statistical equivalence is to take an initial group and randomly assign them to the treatment and the comparison groups.

By random assignment, we mean using something like a coin flip (with a fair coin) or a random number generator to make the assignment. With a large enough initial group, the resulting two sub groups will be statistically equivalent. That is to say, any significant differences among individuals across the groups will cancel themselves out. To a point, the larger the initial group, the more equivalent the two sub groups will appear.

Any systematic factors that might affect the outcome will be distributed across the two groups. Thus the two sub groups will be the same on all relevant characteristics, except that one is exposed to the intervention or treatment and the other is not.

Avoiding Sample Selection Bias

Situations where we have not randomly assigned subjects to treatment and comparison groups have the potential for what we call *sample selection bias*. What this means is that the treatment and comparison groups might differ on a relevant factor. For example, we might conduct a study of residences that have smoke alarms with those that do not.

If fatalities are lower in the alarmed residences, it may not be that most or all of the difference in fatality rates is due to the alarms. It is quite possible that people who chose to install alarms are simply more safety conscious than people who chose not to do so. In other words, those who purchase alarms are also the same people who do not smoke in bed, who take care not to overload electrical circuits, who purchase fire extinguishers for their kitchens, and who plan safe escape routes with their children.

Usually, any situation where people or subjects volunteer or select into the treatment group should be considered suspect. Subjects often volunteer for a program because they are more motivated or see the treatment as potentially more beneficial. Sample selection bias can only be addressed if the evaluator or researcher has done a random assignment to the conditions. Having said this, it is imperative that the researcher engages in true random assignment. It is not unknown for some researchers to select those they think will be the most cooperative or most likely to succeed to be in the treatment as opposed to the comparison group.

Less Than Ideal Variations

Sometimes we cannot randomly assign members of a group to policy or program intervention and others to the control. A situation where this often arises is when governments decide to legislate policy. By their nature, social policies are implemented throughout a jurisdiction and not randomly assigned within particular areas. What happens, for example, if the province of British Columbia wishes to introduce a new set of response standards? Obviously, we can apply the before-and-after model, but we know that has limitations. Are there ways of using the framework of the classical design to overcome those limitations?

Matched Comparison Designs

The answer is, some approaches are less ideal than the classical model but perhaps more convincing than simply using the before-after approach. Since we have no ability to randomly assign jurisdictions to different response standards, one approach is to find potential clones. That is, jurisdictions with different standards that we know (or, more likely assume) to be similar in all or most relevant aspects. For British Columbia, we might consider choosing Washington and Oregon States and the Province of Alberta as comparators.

The assumption here, of course, is that these jurisdictions have different response standards but have similar geographical and socio-demographic characteristics to British Columbia.

We call this approach the *matched comparisons procedure*. We attempt to find matching jurisdictions that are as similar as possible to the experimental one(s) to provide a relevant control group. Again, the issue of sample selection bias might arise, since there is likely something different about jurisdictions that decide to implement a policy over those that do not. Just as with the simple before-and-after approach, we need to regard these results with greater suspicion than those obtained from the gold standard of the classical design.

Regardless, matched comparison designs have produced convincing evidence that certain practices are effective. Perhaps one of the best examples is the early research into the use of daytime running lights on automobiles for reducing accidents. On the flip side, matched comparison studies have also suggested that some policies do not have the intended impact. A good example here is the research into the relationship between capital punishment and homicide rates.

The preponderance of the cross-jurisdictional evidence suggests that while capital punishment may assuage our feelings for revenge, it does little to reduce actual amounts of homicide.

We need to make a decision and the stronger the evidence, the more likely the decision is the correct one. We could be wrong, but even wrong decisions help us know what doesn't work. Doing the same thing over and over makes no sense if the results do not change. When it becomes obvious that our current practices do not have the desired impact, logic suggests we should try something different. Eventually, we are likely to find something that does work. An important factor is that we must be willing to change our view when faced with contrary evidence.

> Too often, we rituatlistically engage in the same behaviour even when the evidence shows it doesn't generate the outcome we wish.

Too often, we ritualistically engage in the same behaviour even when the evidence shows it doesn't generate the outcome we wish. For centuries, physicians engaged in blood letting because, despite the evidence, it seemed to make "common sense" at the time. The fact that many patients were unnecessarily weakened by the practice and subsequently died, was not a consideration.

The Essentials

The important point behind this discussion is that how evidence is collected—the framework or design used to generate the data—is an important element in helping us determine how credible the evidence might be. Among the key factors is the notion that nothing has any meaning unless it is in comparison with something else.

In other words, everything needs a comparator for us to be able to make sense of it. An intervention or an action only makes sense in comparison with another action or a non action (doing nothing). That comparator is known as the counterfactual.

Since something cannot be in two different situations at once, we must look for the most appropriate comparison. As we have seen, clones are hard to come by, so the best approach we have devised to date is the randomized experiment where subjects or objects of interest are randomly assigned to a treatment group and an appropriate comparison or control group. The randomization process helps ensure that there will be no systematic sample selection bias.

In some cases, random allocation to treatment and comparison group is not possible, so we try to create situations that come as close to that ideal as possible.

Evidence generated by these approaches should always be considered suspect but, if the approach appears sound and there are few logical alternative explanations for the effect, then we are generally willing to give the evidence reasonable weight until we find something superior.

Even with the best designed experiments, however, the results are not always equally credible. The design is one element we consider; the magnitude of the impact or size of the effect being produced is another factor. Obviously, interventions that produce large effects provide better reasons for using the evidence for a decision than small or marginal effects. But that leads us to other considerations such as cost-benefit or cost-effectiveness analyses—the topics of our next chapter.

Costing Analysis

Basic Concepts

Costing analysis comes in one of two variations. The first instance deals with the costs associated with doing something. For example, the decision to purchase a vehicle involves not only the capital cost of that vehicle, but also maintenance such as the cost of repairs, consumables such as gasoline, and support costs such as insurance. Depending on the circumstances, additional support costs may arise, such as those associated with having to build a new garage or rent a parking space. If we are looking at the true cost of ownership, we should also factor the depreciation of the vehicles (hopefully, we will recuperate some capital cost when we sell it in a few years) plus the interest of the funds used to purchase the vehicle.

The other form of costing analysis is what we term a *cost-benefit* or *cost-effectiveness analysis*. In this instance, we weigh the costs associated with the decision with the value of the expected benefits. For example, a station might choose to invest in further training. The question then arises: What is the return on that investment? If the training relates to how to fight fires in high rise buildings in a community where none exist, the return on investment might be considered zero.

Costing studies allow us to identify the total cost of a decision and the associated benefits.

In fact, it is a straightforward cost situation. On the other hand, if the training relates to health and safety matters, the returns may appear in lower accident and injury rates, fewer sick days, lower insurance rates, more efficient or productive employees and higher employee morale. We can weigh the relative value of those benefits against the cost associated with the training sessions to estimate the relative return on investment.

A fundamental idea of economics is the notion of *opportunity cost*. Assuming you have a limited budget, deciding to do one thing necessarily precludes another. For example, given a department's capital budget, the decision is made to purchase a pick-up truck. By making that choice, the alternatives—an SUV, a sedan, a motorcycle, and so on—are foregone. That is to say, the opportunity to select an alternative is no longer available. Not only is the physical choice of the next best alternative not available, we give up the benefits associated with that choice.

Costing studies allow us to identify the total cost of a decision and what the returns or benefits associated with that decision might be. Furthermore, we can also examine what we consider the expected cost and returns associated with the second or third best choices, and compare those to our preferred decision. Sometimes this exercise results in our seeing a "lesser" alternative as superior to our initial preference.

Monetary costs are not, nor should they be, the only factors that we consider when we make a choice. Political and other social considerations influence how we make choices. However, monetary costs are important and are typically easy to quantify. Most products and services have a monetary or market cost associated with them. Also, social and political costs are often closely linked to economic decisions. As with formally assessing monetary costs, using the general costing framework to assess the impacts of non monetary decisions is also possible. The only difference is that in those situations, the costs and returns are often more difficult to quantify. Regardless, decision makers can and do use qualitative data to weigh the impact of those types of decisions.

No matter whether we do a straight costing analysis, cost-effectiveness or cost-benefit analysis, there are five overall steps to consider.

Steps to Consider

1. Identifying the component in the unit's operating or strategic plan to which the question or analysis relates.
2. Setting out the objectives that we intend the decision to achieve.
3. Identifying the options or choices that are available.
4. Conducting a financial (cost-benefit or cost-effectiveness) analysis of the option selected or the options under consideration.
5. Preparing an accounting statement summarizing the results.

These steps may appear to be a restatement of what we have mentioned previously. This is the case. However, we need to see effective evidence-based decision making as part of a broad framework that starts with a consideration of what we are doing and why, what are the alternatives, and what evidence can we bring to bear to help us make a decision. Unless we know what we are doing and why, it is almost impossible to identify the appropriate information. Without knowing that, we may collect much data but we likely won't be collecting much evidence.

Cost Analysis

Straight costing studies involve estimating the total *life cycle* cost of a particular piece of equipment or service. By life cycle, we are referring to the period during which we use the product or service. For example, a motor vehicle might have an actual average life expectancy of about 12 years before it is ready for the scrapyard. A person or a business might decide to buy a vehicle, keep it for five years and then sell it. In that instance, the product's life cycle is five years.

The key to conducting accurate cost analyses is to ensure that we include all of the appropriate costs. Generally, for equipment or capital goods, these fall into the following categories:

- depreciation,
- interest on capital,
- maintenance fees (consumables and repairs),
- licensing or regulatory costs, and
- operator costs.

While analysts will often exclude operator costs from the analysis, those need to be considered, even if the final decision is to exclude them. If the equipment is meant as a replacement component, then the operator costs would carry over from the previous piece of equipment. However, suppose a fire department has decided to purchase an additional ladder truck or to include an Emergency Medical Services vehicle in its inventory.

The key to conducting accurate cost analyses is to ensure that we include all of the appropriate costs.

That additional vehicle may require extra personnel, the cost of whom we need to factor into the analysis.

Some of you may wonder why we have just included depreciation in our list of items instead of the initial capital cost. Here the assumption is that the piece of equipment will be sold at the end of the life cycle. Consequently, the capital cost component here is the difference between the purchase price and the selling price. This is what we call depreciation.

Different pieces of equipment depreciate at different rates, but it is common for that to be about 20-30 per cent per year. We calculate depreciation on the outstanding value, so a $10,000 piece of equipment that depreciates at a rate of 20 per cent would be worth $8,000 after the first year. The second year's depreciation would be $8,000 x .2, or $1,600. Thus, the total depreciation after two years would be $2,000 + $1,600, or $3,600, and the residual value of the equipment would be $10,000-$3,600, or $6,400.

One item often forgotten in costing studies is the interest on the purchase. Interest rates are sometimes called discount rates in the literature. The need to consider interest is generally easy when one borrows the money to make the purchase, since the bank or financing company will include that charge. However, even where the equipment is purchased outright, we should include the "rental" value of the capital. The reason for this is that if we had not made the purchase, we could have invested the money for a given return or used it for some other purpose. This, in effect, is another form of opportunity cost.

Obtaining Reliable Cost Estimates

Whether it is the total cost of hiring someone or purchasing a piece of equipment, the key to good costing studies is to ensure we include all items, and obtain the most accurate and reliable cost estimates of those items. Because all organizations work in different environments, typically we gain the best information from experience. Looking back over your unit's financial records can be revealing. Because they reflect actual experiences, it is easy to see where unexpected costs (and savings) arose. Do not write those off as unique or one-time occurrences; put those in as line items in your analysis.

Where drawing on institutional experience is not possible, one can often obtain information from other sources.

Often, suppliers will give cost comparisons with competitors' products. Beware, however, that those analyses often selectively include or exclude "inconvenient" line items. Make sure that you are comparing the proverbial apples with apples. Where you find missing items, make sure to ask for supplemental information.

Many independent agencies also conduct costing analyses of various items. Look especially to professional or trade associations. Similarly, governments and public organizations will often make their budgets and costing studies available. Much of that can be found online or in a local library. Sometimes a simple phone call can result in a gold mine of data.

An example of a straight costing study is presented in the box on the next page. Here, we are looking at the cost of owning and operating a typical, full-size pick-up truck over a five-year period. The cost of the operator is not included in this example.

Straight costing studies are done to estimate life cycle costs to decide the affordability of a purchase. They are also useful in comparing different products. For example, one brand of pick-up might have a higher capital cost but lower maintenance costs than another. The question then becomes: Which is the better choice?

The Cost of Purchasing a Truck

Five-year cost of purchasing and operating a pickup truck:

Item	Cost ($)
Purchase price	27,500
Selling price	9,349
Depreciation	18,151
Financing	1,737
Fuel	15,855
Insurance	4,075
Taxes and licensing fees	3,650
Maintenance	3,085
Repairs	2,343
Total cost	48,896
Cost per kilometre	0.49

Assumptions:

- 20,000 km driven per year
- 2.7 per cent APR financing cost with $2,750.30 down payment
- gas $1.25/l.
- mileage at 12.66 l/100 km.

Similar analyses can be used to decide whether it is less costly overall to purchase a used vehicle as opposed to new, or to lease as opposed to purchasing outright. Obviously, for these different scenarios, we must make different assumptions regarding expected life cycle, operating costs and depreciation. It might also be worth repeating that the values used in costing studies are generally estimates. As we discuss in the chapter on statistics, all values are estimates. The key, with a little research and experience, is to minimize the error. However, many expected items, such as the selling price of the vehicle and the actual cost of operation, are based on assumptions that are out of one's control.

We have considered the cost of capital goods but we can conduct similar analyses for personnel. The same general principles apply. Typically, we focus on a person's salary when deciding to hire someone, but ancillary costs can be substantial. When pensions, taxes, insurance, benefits and other compensation-related issues are considered, it is common for those to add an additional 15-30 per cent to the total salary cost. This is above the cost of training and maintaining the person. Maintenance costs include the person's working space and any equipment and supplies they may need to do their job. In the previous example, we noted that equipment typically needs an operator. So, too, do people often need equipment to do their jobs.

A Note on Cost-effectiveness

In the previous analysis, our attention was on the total cost of owning and operating a vehicle over its life cycle. Knowing the total cost of something is an important consideration in decision making. Often, however, knowing the total cost does not tell us the whole story. Most equipment or other items generate some form of output or product. For a truck, the output is transportation. In that instance, knowing the cost per kilometre is often a more valuable piece of information than the total cost.

In the example provided in accompanying box, the expected cost of the pick-up per kilometre is about $0.49. We term the price or cost of something per unit of output as its cost-effectiveness. While cost-effectiveness is clearly related to total cost, we should treat it as an independent issue for decision making. Often, differences in total costs might be irrelevant. It is the per-unit cost that is important. One reason unit costs differ from total costs is the fact that total costs consist of two components: fixed or sunk costs, and variable costs. Fixed costs are associated with the one-time cost of purchase. Variable costs generally relate to operating and maintenance costs. A piece of equipment may have a higher fixed cost but, if it is more efficient than a lower priced piece, it will generally have lower unit costs.

A key element in cost-effectiveness analyses is being able to identify the appropriate output measures and being able to measure them appropriately.

The same applies to personnel costs. Higher salaries to people who are more productive, who are less likely to miss work and who provide a better quality of service can outweigh “savings” accrued by outsourcing to lower-cost jurisdictions. What is important is how many items are produced, how many people are served, and the quality of that output or service.

A key element in cost-effectiveness analyses, however, is being able to identify the appropriate output measures and being able to measure them appropriately. Again, this is where examining the organization’s operating or strategic plans becomes important. It is in those documents that the organization’s objectives and operational purpose should be outlined. Either directly or indirectly, an organization’s effectiveness is related to the product or service it is meant to deliver.

Cost-benefit Analysis

Cost-benefit analyses are generally extensions of simple cost-effectiveness studies. A primary difference is that cost-benefit analyses look at a broader range of returns on the investment. Most cost-benefit analyses include effects (benefits) that are not easily quantifiable or outcomes that have a broader social impact.

Cost-benefit analysis is grounded in welfare economics. It differs from most branches of economics since the focus in not just on decisions of consumers and firms, but on public decisions that affect the economic interests of a broader community. Consequently, cost-benefit analyses often focus on issues such as quality of life or quality of the environment. A fundamental challenge for those doing cost-benefit analyses is how to measure the benefits so they are comparable across issues. Among commodities, apples are not electrical transformers. However, a market for both exists and it is possible to place a monetary value on both. Currency is a common exchange unit that allows the producers of apples to purchase transformers even when the producers of transformers have no interest in exchanging their product for apples.

The difficulty with many public goods and services is that there is no open marketplace in which the monetary value of those items is established.

Moreover, for ideological reasons, many people refuse to assume a monetary value on public goods. A common refrain, for example, is that "You can't put a price on the environment" or "You can't put a price on a human life." The fact is, we do both. The problem is that no independent or *indifferent* market exists to set those prices. Regardless, this is an essential weakness of cost-benefit as opposed to straight costing analyses.

Revealed and Stated Preferences

While the philosophical issue of whether you can truly value a human life may not be answerable, welfare economists have two broad tools at their disposal. They term one approach the *revealed preference* method. Revealed preferences relate to how people actually behave when confronted by a qualitative phenomenon. For example, comparing a particular property with similar ones could reveal the "eyesore value" of having a fire hydrant on a front lawn. How much parents value education for their children might be suggested by what proportion of their income they are willing to spend on a child's tuition.

The second tool in the economist's repertoire is what we call *stated preferences.* Stated preferences are just that: what someone is willing to tell you they would pay for something.

We may judge people's value of environmental elements, for example, by how much of a tax increase they are willing to support for clean air or nature conservatory initiatives. Typically, stated preferences are determined through surveys and similar procedures.

While both stated and revealed preferences have their merits, both have their limitations. Using how much life insurance a person has to assess how much they value their lives might appear like an excellent revealed preference. However, how much they can buy is limited by how much insurance they can afford. Furthermore, a person may value their life highly but not be willing to see relatives "benefit" from their death since life insurance goes to the beneficiary and not the insured. Stated preferences on various aspects have been studied extensively by sociologists for the past century. Their overwhelming conclusion is that what people say and what they do varies considerably.

Still, cost-benefit analysis is one of the few techniques we have to assess the broader impact of various policies and programs. For its limitations, it helps us to clarify the issues, identify the constituent components, and bring some evidence to bear on the issue. It has gained general acceptance in the public sector and is mandatory in many government shops. For example, the Treasury Board of Canada has mandated that any regulatory framework put in place by the federal government must be based on a cost-benefit analysis.

We can use cost-benefit analysis to:

- decide whether a proposed project should be undertaken
- decide whether an existing project should be continued
- choose between alternative projects

The purpose of this is for "departments and agencies [to] assess regulatory and non regulatory options to maximize net benefits to society as a whole. Hence, all regulatory departments and agencies are expected to show that the recommended option maximizes the net economic, environmental, and social benefits to Canadians, business, and government over time more than any other type of regulatory or non regulatory action."[1]

In summary, we can use cost-benefit analysis in various ways. For example, to:

- decide whether a proposed project or programme should be undertaken;
- decide whether an existing project or programme should be continued; or,
- choose between alternative projects or programmes.

Components of a Cost-benefit Analysis

In setting up and executing a cost-benefit analysis, several steps need to be followed. These include:

1. **Define the problem.**

   Again, this is a statement of the issue with a link back to your operational or strategic plan.

2. **Identify any constraints or limiting factors.**

   This is a discussion of what administrative requirements and other challenges you might face. These include a listing of financial limitations, managerial or personnel challenges, environmental and other regulations, and any other factors or "hurdles" you might need to address.

3. **List the alternatives.**

   Every initiative has alternatives, including doing nothing or staying the same. For example, if the issue is whether to close a particular fire hall or not, it may be informative to looking at amalgamating with another unit, integrating other services such as community policing or ambulance services, or expanding the operation to incorporate other halls.

4. **List the benefits.**

   For the alternatives outlined, what is the return on investment? Is there a monetary return or an increase in productivity or effectiveness? Perhaps, the matter is not one of generating further revenues, but one of reducing or avoiding costs. Are there health, safety or environmental benefits to be gained? The issue might be related to overall quality of life. Are there savings to be had in equipment, time or personnel?

5. **How are the costs and benefits to be quantified?**

   Clearly, market or monetary values of goods and services are the easiest with which to work. We have already outlined the challenge of providing market values. Still, finding a shadow or proxy price for a given cost or benefit may be possible. Social scientists have developed ways to estimate the value of a human life. The cost associated with noise levels or high traffic volume in a community, for example, can be estimated by differences in housing values between noisy and quiet communities or between those with high and low traffic volumes.

Often, we can find ways of assessing the value of tough-to-monetize issues by searching the appropriate literature. We have already discussed techniques for conducting more focussed online searches. Using the expertise of economists and other social scientists in local colleges and universities might also be possible.

Once we have conducted these steps, we can put a report together summarizing these elements and presenting the relative costs and benefits.

Net Present Value

As the saying goes, "A bird in the hand is worth two in the bush." So it is with money. One reason we charge interest on borrowed money is that by giving capital to a borrower, the lender faces an opportunity cost. That money cannot be used for anything else. To compensate the lender for the opportunity cost, borrowers must pay interest. For example, when you buy a locked-in savings certificate with a five-year redemption, you get back more than you invested. A $1,000 certificate invested at 3 per cent would be worth $1,000 x 1.03 x 1.03 x 1.03 x 1.03 x 1.03 = $1,000 x 1.03^5 = $1,159.

We can also consider the opposite. What would an endowment of $2,000 that you are to receive in five years be worth to you today? In other words, what would you be willing to pay for the benefit of having the cash right now?

This is the principle behind reverse mortgages. A bank or financial institution will give you a fraction of your home's value today if you allow them to sell it at market value and keep the proceeds several years hence. This is the opposite of the previous problem. In these instances, we call the interest rate the *discount rate*. At a three per cent discount rate, that future $2,000 endowment would be worth $2,000 x $1/1.03^5$ = $2000 x .863 = $1,725 today.

We term this current value on a future amount its *net present value* or NPV. The NPV is the opposite of the future value. Since programs and capital goods have an expected life cycle, it is common to standardize costs to today's value, that is, the NPV. Another way of thinking about NPVs is to consider them as equivalent to constant as opposed to real dollars when we are trying to control prices for inflation.

In these examples, we have discussed what economists call the *private time preference rate*, since the focus is on an individual. Within the public sphere, the choice to invest public funds in a particular program often precludes investments in other programs of benefit to the public. Within the public or welfare sphere, economists generally call the deferred value the social opportunity cost. While the terminology differs, the underlying principles are similar.

Benefit-Cost Ratios

For programs extended over time, we need to amortize both cost and benefits. Occasionally, the duration of the costs may be different from the duration or life expectancy of the benefits. Consequently, to make things comparable, analysts will calculate the NPV of both costs and benefits.

We term the ratio of the benefits to costs as the *benefit-cost ratio* or BCR. Assuming the NPV of the benefits of a program is $13.5 million and the net present value of the costs is $10 million, the BCR would be:

$$\mathbf{BCR} = \frac{\text{(NPV Benefits)}}{\text{(NPV Costs)}} = \frac{13.5}{10.0} = \mathbf{1.35}$$

Ideally, the BCR should be greater than one. Anything less assumes that the costs outweigh the benefits and, all other things being equal, the option should not be chosen. If we chose to evaluate several alternatives, the one with the highest BCR would normally be our choice. If a program with a lesser BCR is selected, then it is likely that we should have included the reason for that selection on the benefit side of the ledger.

Example: Home Sprinkler Systems[2]

In 2007, the US National Institute of Standards and Technology (NIST) conducted a cost-benefit analysis on the installation of sprinkler systems in newly-constructed single-family housing.

The report put out by NIST noted that a previous study conducted in 1984 concluded that, at the time, sprinkler systems in single-family dwellings were not cost-effective. They backed this assertion by noting that residential sprinkler installations were rare. Sprinkler technologies had changed over the past two decades and had become less costly and more viable as options within private homes.

The primary costs were associated with the purchase and installation of the sprinkler system. The expected benefits for residents included "reductions in the following: the risk of civilian fatalities and injuries, homeowner insurance premiums, uninsured direct property losses, and uninsured indirect costs."

The analysis was fairly sophisticated. The fact that sprinkler systems had a long life cycle was considered, with the costs and benefits being amortized and discounted to a reference period. Three different house types were considered, including a two-storey colonial with a basement, a three-storey townhouse, and a single-storey ranch house. They estimated the present value costs of installing a multipurpose network sprinkler system at $2,075 for the colonial, $1,895 for the townhouse, and $829 for the ranch-style. An example of the items considered in the sprinkler systems is outlined on the next page.

NIST Residential Sprinkler Cost Analysis Considerations

Sprinkler System Cost Component	Quantity	Units	Bare Material Cost/Unit	Total Bare Material Cost	Labour Cost	Combined Material & Labour Cost
Material						
Fire Sprinklers						
F1/Res 49 (155 °F) (68.3 °C) Recessed Pendent Assembly, White	22	each	$25.03	$550.55		
Pipe and Fittings						
½ in (12.7 mm) PEX - white, 1000 ft (304.8 m) coil	1	1000 ft.	270.00	270.00		
½ in (12.7 mm) PEX - white, 100 ft (30.48 m) coil	1	100 ft.	27.00	27.00		
1 in (25.4 mm) Copper Branch Manifold, 12 outlets	1	each	32.23	32.23		
PEX Ring ½ in (12.7 mm) (blue	150	each	0.06	8.25		
PEX Brass Tees, ½ in (12.7 mm) PEX x ½ in (12.7 mm) PEX	10	each	1.45	14.50		
Accessories						
Hangers (½ in [12.7 mm], 5/8 in [15.875 mm], ¾ in (19.05 mm) PEX	3	each	5.95	17.85		
Total Bare Material Cost				920.38		
Labour						
Design Cost (4 h at $40.00/h)					$160.00	
Labour Cost (12 h at $50.31/h)					603.72	
Total Labour Cost					763.72	
Totals						
Total Material and Labour Cost						$1684.10
Total Material and Labour Cost without cold water system				(99.00)	(100.62)	1484.48

Where possible, generic product descriptions have been substituted for product trade names. Material prices do not include any markup to cover overhead and profit. Labour cost is based on manufacturer's estimation that it would take a 2 person crew 12 h total to install the system. The sprinkler fitter and plumber trades are estimated at $50.31/h (2007 National Construction Estimator, accessed at www.get-a-quote.net). Design cost of $40/h is provided by manufacturer. Extra sprinkler heads and cabinet exceeding the minimum requirements of NFPA 13D were removed from original estimate. For the estimates without the cold water system, one-third of the combined pipe and 2 h of installation labor are subtracted.
Source: This table is based on NISTIR 7277, Economic Analysis of Residential Fire Sprinkler Systems (Brown 2005, page 12); however, the labor cost has been changed to $50.31.

The researchers summed the expected benefits (costs) across fatalities averted, injuries averted, direct uninsured property losses averted, indirect costs averted and insurance credits accrued. They estimated that the expected difference in present value was $2,919 for the colonial-style house, $3,099 for the townhouse, and $4,166 for the ranch-style house. In other words, those were the net benefits after costs. All three components, the costs, benefits and net benefits are outlined below.

NIST Residential Sprinkler Costs, Benefits and Net Benefits

	Colonial	Townhouse	Ranch
Benefits			
Fatalities Averted	$3725.57	$3725.57	$3725.57
Injuries Averted	224.74	224.74	224.74
Direct Uninsured Property Losses Averted	79.64	79.64	79.64
Indirect Costs Averted	15.93	15.93	15.93
Insurance Credit	948.41	948.41	948.41
Benefit Subtotal	4994.29	4994.29	4994.29
Costs			
Installation (50 % Markup)	2075.08	1895.17	828.66
Costs Subtotal	*2075.08*	*1895.17*	*828.66*
Net Present Value	$2919.20	$3099.11	$4165.62

Summary

While costing studies are but one way of generating data for evidence-based decision making, they are often one of the more commonly used tools. Essentially, costing studies do three things for us. First, when done properly, they link the outcomes we wish to measure with the goals and objectives of our operational and strategic plans. They essentially help us focus on the question about whether the activity is within the organization's mandate.

Second, costing studies help us focus on the many line items that make up actual costs. Often, "back of the envelope" or convention-based costs omit many ancillary costs associated with our activities. For example, it is common for costing studies to omit interest payments or costs associated with the need for extra personnel. By focussing on a detailed analysis, we are more likely to ensure that we include those items. Furthermore, exhibiting the results of a costing analysis to colleagues and others provides the opportunity for independent observers to identify potentially missed items.

Third, costing studies provide a transparent and fairly mechanical way of helping us decide on options. The assessments are relatively objective and focussed. The assumptions underlying the costs can be scrutinized, as can the values associated with individual items. The transparency of the process provides for a more defensible decision: one that is replicable by an independent observer. Furthermore, unlike purely value-based decisions, decisions based on evidence force critics to generate alternate values or analysis to validly criticize the analysis. Even if someone can put forward alternate evidence, a net benefit still exists since that evidence will contribute to a more accurate assessment of the situation. In the end, a better basis for a decision is put forward.

Notes

1. Treasury Board of Canada (2007) Canadian Cost-Benefit Analysis Guide: Regulatory Proposals. Ottawa: Government of Canada. Catalogue No. BT58-5/2007. http://www.tbs-sct.gc.ca/rtrap-parfa/analys/analys-eng.pdf
2. Taken from D. Butry, M.H. Brown and S.K. Fuller (2007) Benefit-Cost Analysis of Residential Fire Sprinkler Systems, Gaithersburg, MD: U.S. Dept. of Commerce, National Institute for Standards and Technology.

Making Decisions

Using Evidence

We make decisions all the time in our private and in professional lives. Mostly, those decisions are based on what we learned in our training, on conventional wisdom, or on traditional practices. Often, questioning common practice only leads to rediscovering the wheel. Yet, there are many circumstances where traditional practice and common knowledge doesn't work. We may not achieve the results we want, or our practices lead to less-than-efficient outcomes. For some reason, however, humans are reluctant to change. We are a conservative species. We become comfortable doing the same thing repeatedly, even when we are not happy with the outcome. As the Alcoholics Anonymous Handbook states, however, "Insanity is doing the same thing, over and over again, but expecting different results."

Historically, we can forgive decision makers for pursuing timeworn rituals. After all, as rainmakers knew, if you danced often enough, it would eventually rain. Modern weather forecasting has become sufficiently accurate, however, that rainmaking is no longer a viable profession.

Evidence-based decision making makes the process transparent—it is no longer a closed, magical process, but one where observers can follow the logic and follow the evidence.

The reason for that is meteorology has accumulated sufficient systematic knowledge that it is possible to predict local temperatures, precipitation and other phenomena with a high degree of certainty. Meteorologists have accomplished this by turning to scientific research and other forms of systematic study.

The reliance on systematic study and data collection, which is what underlies science, has made inconsistent inroads in most other disciplines. This is unfortunate since, today, there is a large amount of empirical evidence to help us make better decisions. Furthermore, where existing analyses do not exist, conducting a local analysis to improve our own decision making is often not that difficult. This doesn't mean that one needs to become a scientist—far from it. All we need to do is to use empirical results to be able to build a reliable body of evidence.

Decision making based on evidence will generally allow you to make better decisions. Evidence-based decision making has the advantage of making the process transparent. Outsiders can become privy to the foundations of the decision. It is no longer a closed, magical process but one where observers can follow the logic and follow the evidence.

Evidence-based decision making is using the best available research and information on the outcomes of fire work to carry out guidelines and evaluate agencies, units, and personnel.

We are not suggesting that you can always find an optimal solution to your problem. However, evidence-based decision making helps us to identify options and practices that do not work. In those instances, you are likely no worse off trying something new. Most often, however, a review of the existing evidence or the collection of your own data will help provide a more fruitful direction.

In summary, how can we put the lessons of this book together to formulate a good evidence-based strategy for decision making? Essentially, there are four main steps.

Identify and Frame the Question

The first three chapters of this book are focussed on identifying appropriate questions. Without the right question, no amount of data will help provide an answer. We have stressed repeatedly that good questions need to be put into an appropriate framework. Ideally, you should draw these from your organizational plan or your strategic plan. This helps to focus the issue on the key purpose and objective of your unit. One main reason many organizations fail is that they lose sight of their mandate. They try to be all things to all people. This is simply not doable.

If you lack an organizational or strategic plan, the next best thing is to drill into the issue. Ask several fundamental questions:

- Why are we proposing to do this?
- What are the likely outcomes?
- How does this action relate to the organization's mission?
- What benefits will this action bring to my unit or the people we serve?
- Are there more cost-effective or cost-efficient alternatives?
- Does this action have long-term or short-term consequences?
- What other resources am I likely to need if we pursue this action?

If what you are proposing to do is new or outside the traditional scope of your organization's traditional mandate, consider putting together a focussed business plan to support or justify the activity.

Once you have identified and justified the appropriate question, outline the options. Commonly, two or three viable alternatives are available. In other situations, the range of options and their relative merits is not necessarily obvious. In those situations, consider performing an environmental scan or SWOT analysis. If the issue is crucial, consulting an outside facilitator may be worthwhile.

Gather the Evidence

Often the best source of evidence is your own organization. You keep records of calls for service and your financial accounts. Those and other resources can give you valuable insights. Usually, internal data will provide a good base line or a measure of the status quo.

Outside your organization, other sources of information are available. Professional and trade organizations are a good place to start. Suppliers will also give you information on comparative options and estimates of lifetime service costs. Do an online search. Despite all of the trash on the internet, there are also nuggets to be had. Learn how to use your favourite search engine to eliminate as much of the irrelevant material as possible. Do not be afraid to check organizations in outside jurisdictions. In the UK, the Department for Communities and Local Government is responsible for keeping information on fire incidents. It also produces many annual reports and studies. Similarly, the US Fire Administration is a gold mine of information, as is the National Institute of Standards and Technology.

> Remember, a librarian can be your best friend.

Other excellent sources of information are libraries and your local college or university. Libraries have access to online databases that can search academic articles and other specialized material. Some of this can be intimidating to us if we are not used to using the facilities. *Remember, a librarian can be your best friend.* Contact your municipal librarian or visit a local college to seek expert advice.

Librarians can also help you navigate statistical databases. Most provinces and provincial agencies collect and make available regional data. While most data are available to the public, some is limited to authorized agencies. If you work for a public service agency, it is likely that yours is one of those authorized agencies. The Statistics Canada website is also a valuable source of information.

Some colleges and universities have laboratories and research groups or institutes that focus on fire-related matters. Again, these can often be found through an internet search or by asking a local librarian for help.

Do keep in mind, however, that not all evidence is of equal value. Do not be afraid to be critical, especially if claims are at odds with your department's or your colleagues' experience. While not always the case, if something is too good to be true, it generally is. Ask yourself if the source is trustworthy. Is the agency presenting the data operating impartially or at arms length, or does it have a self-serving agenda? Has the research or the publication gone through an external review process?

Organize the Evidence

Once you gather it, put your evidence together in an organized manner. Costing studies are easily presented in a spreadsheet. Other material can be presented in a table. Be sure to label your information and keep track of where you found it. That way, if someone questions its veracity, you can refer them to the source.

A key element in presenting data is putting it in context. Remember, nothing means anything unless it is relationship to something else.

Is a six-minute average response time adequate for an urban fire department?

> A key element in presenting data is putting it into context.

Is a million dollars an appropriate price for a fire truck? Is our level of training adequate? These questions can only be answered by making reference to a comparable benchmark. What is the price range for goods and services in the marketplace? What are industry norms or standards for performance? Are there best practices against which you can compare your unit or department?

Review the Decision-making Process

Once you have done your analysis, it is good practice to review the entire decision-making process. What have you learned? How could the process be streamlined or made more efficient? The more you engage in evidence-based decision making, the easier it will become. Knowledge is cumulative. You will soon determine the best sources of information. You will discover how to make the process more efficient and how to minimize the likelihood of getting sidetracked.

While evidence-based decision making generally takes longer than other approaches, it has its benefits. Decisions based on hard evidence are more resilient in the face of scrutiny.

The more you engage in evidence-based decision making, the easier it will become. Knowledge is cumulative.

Taking a request to city council with strong external evidence is more likely to result in a positive decision. Presentations that show prior examples of success or that have reliable estimates of returns on investment are powerful. Finally, if someone challenges you, it is fair play to say that you have provided evidence to support your request. If they disagree, then ask them to show you the numbers.

Making the Right Decision

As a professional in the fire service or other public service, you make crucial decisions every day that balance need with available resources. How should you approach these decisions, and how can you justify the decisions you make?

In this manual, Professor Paul Maxim, Fire Chief and Professor Len Garis and Professor Emiritus Darryl Plecas explore the what, why and how of evidence-based decision making.