

# THE EIGHTFOLD PATH

The analytic work in problem solving generally proceeds in a certain direction, from defining the problem at the beginning all the way to making a decision and explaining it at the end. But remember, this is a process much given to reconsidering, reviewing, and changing your mind—in other words, retracing your steps on the path before starting out once more. Also, in some cases, the client or, perhaps, the political situation has already narrowed and focused the analytic task to such a degree that you need not even bother thinking through some of the steps. The exposition that follows lays out a generic process that must be adapted to particular contexts.

## STEP ONE: DEFINE THE PROBLEM

---

Your first problem definition is a crucial step: it gives you both a reason for doing all the work necessary to complete the project and a sense of direction for your evidence-gathering activity. And in the last phases of the policy analysis, your final problem definition will probably help you structure how you *tell your story*.

It is easy to get the problem definition step wrong. Analytic looseness can creep in, creating a muddle. For example, at a congressional hearing about regulation of social media, lawmakers expressed the following concerns about Facebook:

- Facebook is too big and needs to be broken up.
- Facebook does not exercise sufficient care about sharing the data with outside organizations.
- Facebook collects too much data in the first place.
- Facebook is causing ideological polarization.
- Facebook is vulnerable to political exploitation and is not doing enough to curb hate speech and fake news.
- Facebook is promoting addictive message products to children.<sup>1</sup>

Might some of these concerns form the basis for a usable problem definition? Possibly, but all of them would require much greater precision. Many of the

aforementioned concerns point not to problems but rather to policy options. Still others are merely claims (which may or may not be accurate) about Facebook or social trends. The following suggestions can help improve problem definitions.

## Think of Deficit and Excess

**Semantic Tip** It often—but not always—helps to think in terms of deficit and excess. For instance:

- “There are too many homeless people in the United States.”
- “The demand for agricultural water is growing faster than our ability to supply it at an acceptable financial and environmental cost.”
- “Utah’s population of school-aged children is projected to grow by 10 percent over the next seven years, and our ability to develop the physical facilities in which to educate them is not growing nearly as fast.”

It often helps to include the word *too* in the definition—as in “too big,” “too small,” “growing too slowly,” or “growing too fast.” These last two phrases (about “growing”) remind us that problems deserving our attention don’t necessarily exist today but are (at least potentially) in prospect for the future, whether near or distant.

However, it does not help to think in terms of deficit and excess when your problem is an already well-structured decision choice—for example, “Dump the dredging spoils either in the Bay or somewhere out in the Pacific Ocean.” Nor does it help if your challenge is to invent *any* way to accomplish some defined objective—for example, “Find some grant funds to close the anticipated gap between revenues and expenditures.” These decision- and invention-type challenges are problems *for* the policy analyst but are not the substantive sort of problems we are addressing in this section.

## Make the Definition Evaluative

The idea of a “problem” usually means that people think there is something wrong with the world, but *wrong* is a debatable term. Even if everyone accepts the same facts, not everyone will agree that the facts you (or others) have defined as a problem really do constitute a problem, for each person may apply a different evaluative framework to these facts. For example, some people believe that the fact of a growing foreign-born share of the US population is a problem; others believe it is not. Unfortunately, there are no obvious or accepted ways to resolve philosophical differences of this type.

A common philosophical as well as practical question is this: “What private troubles warrant definition as public problems and thereby legitimately raise claims for amelioration by public resources?” It is often helpful to view the situation through the “market failure” lens (Friedman 2002; Weimer and Vining 2017, chap. 5).<sup>2</sup> In its

simplest formulation, market failure occurs when the technical properties of a good or service have one of the following effects:

- Making it hard to collect payment from all the potential beneficiaries—for instance, the large number of people who profit, albeit indirectly, from national defense or advances in basic science (“public goods”). Trust in others, trust in institutions, and the practice of civil discourse, even in heated political contests, are also “public goods,” albeit of a novel sort.
- Making it hard to collect from the beneficiaries of consumption the true economic cost of making use of the good or service—such as the fresh air that vehicle owners use as a sink for their auto emissions (“negative externalities”).
- Making it hard to collect from all the beneficiaries for the true economic value of producing a good or service—such as all the citizens and co-workers who benefit from the education received by individuals looking mainly after their own self-interest (“positive externalities”).
- Making it hard for consumers (and sometimes suppliers) to know the true qualities of the good or service they are acquiring—for instance, many repair-type services, including those performed by physicians as well as those performed by auto mechanics (“information asymmetries”). Insurance and other such “selection markets” often fail because those who select in have knowledge about their own vulnerabilities that the sellers’ side does not. This can lead to disappearance of a seller’s side altogether.<sup>3</sup>
- Making competition on the producer side socially too costly to permit competition altogether, such as a dozen companies building a dozen power plants to serve the same service area. That classic example of social cost is now supplemented by that of competition that undercuts broad and inclusive information networks (“natural monopolies”).

Distributional outcomes can also legitimately raise claims for amelioration by public resources. Areas of possible concern include low living standards, income or wealth inequality, stalled economic mobility, and the existence of discrimination against women and racial and other minorities. The policy analyst’s tool kit can contribute to the understanding of such problems in several ways. Policy analysts can assess whether and to what extent proposals intended to reduce employer discrimination will increase opportunities for minority job applicants. Policy analysts can also measure the efficiency costs of redistributive policies—and identify the most cost-effective ways to combat poverty. More generally, policy analysts can project the distributional consequences of policy options and identify the potential benefits and harms of policies to marginalized populations. Finally, as noted later in Part I, policy analysis can incorporate equity concerns into their analyses as an evaluative criterion.

In addition to market failures and distributional concerns, there are “legal-framework failures” that affect injured parties’ abilities to sue under usual common-law (torts, contracts, property) rules. If you get food poisoning eating at Greasy Spoon, how do you prove that Greasy Spoon was actually responsible? If some factory dumps smoke into the air you breathe, a court would probably find you have no property right to clean air in the first place. If an oil slick on the factory floor causes you to fall and break your arm, it may be too costly and uncertain to sue the employer—who may not be found liable anyway. Market failures often overlap with legal-framework failures and may be called “market failures,” by analysts, but that characterization can obscure what is essentially at issue and suggest market-related solutions that may be blunt or irrelevant.

Finally, less formal social institutions, such as families and neighborhoods, can fail, leading parents to neglect or abuse their children, for instance. A related type of failure is that of self-management practices like avoiding addiction. Other related types of failure have to do with community climates, such as consensus on language(s), what to consider “legal tender,” enforcement of norms on civil discourse, and a climate of trust and trustworthiness across strangers. We can call these “social capital failures.”<sup>4</sup>

One important implication of trying to find and label various social “failures” is that, absent such identification, it is usually hard to see how government can usefully improve matters. Worse, government may be led to intervene in a way that exposes the society to a variety of “government failures.” While the theory of “government failure” (Weimer and Vining 2017, chap. 8) is not as well developed as the theory of market failure, scholars have identified seven main sources of government failure:

- Problems of direct democracy, such as a majority imposing very high costs on a minority. These are complemented by other problems such as the relative underweighting of interests that have little chance of tipping the balance of power in elections.
- Problems of representative democracy. Between elections, the costs of different interests gaining access to elected representatives and senior administrators differ substantially. The ones with lower costs of collective organization and seeking access do much better in gaining their preferred political outcomes. Political scientists often speak of these as “concentrated” versus “diffuse interests.” Concentrated interests might also have an easier time raising campaign contributions when elections roll around.
- With representative government, time horizons often differ substantially among those whose careers depend on being in government, like representatives always sensitive to votes in the next election or senior managers awaiting retirement, versus the general citizenry. Those with relatively short time horizons prefer to put off costs and risks until their careers are safely over and someone else has to take responsibility.

- Problems due to “overhead agencies” of government, such as procurement, space management, personnel, women and ethnic or racial minority preferences, contracting, legal review, fiscal control, and so on. These have accumulated over the years mainly with the objective of reducing waste, fraud, abuse, and unwarranted discrimination. But their combined presence can slow down government and sometimes forces cumbersome work-arounds to get the job done.
- Problems of multiple veto points and demands for “democratic representation.” One of the main—and justified—complaints about government is that “nothing ever happens.” Or, slightly more optimistically, “it just takes forever.” In the interests of “democratic participation,” fearful and aggrieved interests are given, and also create, countless opportunities to hold things up till they are satisfied or are beaten down.
- Problems of risk aversion among bureaucrats. The “public choice” school of bureaucratic behavior sees elected “principals” struggling to control rogue bureaucrats (appointed and civil service alike) who are in democratic theory supposed to be nothing more than the principals’ “agents.” This model fits some bureaucrats some of the time. However, an equal, if not greater, problem among public managers is the desire to avoid subsequent criticism for their actions. They are fearful rather than imperious. Much delay occurs while the public managers worry, consult, and study.
- Problems of administrative standardization. Private markets can cater to every human particularity of taste and need. But governments need to simplify, homogenize, and standardize. To the extent administrators can do this effectively, standardization is not a failure but a success. Otherwise governance would be impossible. On the flip side of success, though, individual citizens are made to feel like cogs in a machine. Additionally, their real needs and tastes are ignored in favor of a hypothetical “everyman.”

Just because a government failure is occurring does not necessarily mean that shrinking the government’s role would improve social welfare; doing so might make a bad situation worse. Many policy interventions are responses to failures of social institutions or distributional concerns, and some government programs accomplish enormous good for society. Policy design matters a lot. For example, nationally uniform programs that involve direct financial transfers to beneficiaries are often more robust and less vulnerable to problems of administrative coordination and delay. The point in recognizing and analyzing government failure is not to resolve debates about the appropriate size and scope of government, but rather to help craft and implement solutions that will further legitimate policy goals at reasonable cost. In an era in which the government has a presence in virtually every aspect of the economy and society,

an important task for the policy analyst is to identify whether and to what extent current policy interventions are underperforming relative to feasible alternatives or otherwise falling short. Policy analysts thus have a key role in diagnosing and remedying both market failures *and* government failures.

### Using Issue Rhetoric

Usually, the raw material for the evaluative aspect of your initial problem definition comes from your client and derives from the ordinary language of debate and discussion in the client's political environment—language that we call generically “issue rhetoric.” Such rhetoric may be narrowly confined to a seemingly technical problem or broadly located in a controversy of wide social interest. In either case, you have to get beyond the rhetoric to define a problem that is analytically manageable and that makes sense in light of the political and institutional means available for mitigating it.

Use the raw material of issue rhetoric with care. It often points to some condition of the world that people don't like or consider “bad” in some sense, such as “decaying infrastructure,” “corporate welfare,” or “wage stagnation.” These evaluations do not necessarily need to be taken at face value. You will sometimes wish to explore the philosophical and empirical grounds on which you, your client, or others in your eventual audience should or should not consider the alleged condition “bad.” Furthermore, issue rhetoric may point to some alleged—but not necessarily real—cause of a troubling condition, such as “gentrification” or “globalization.”

Issue rhetoric often has a partisan or ideological flavor. Although most ordinary Americans do not possess a consistent ideology, issue rhetoric is created by the more passionate and often more articulate individuals whose views tend to be uniformly extreme in one direction. The great ideological divide in most developed democracies concerns the role of government assistance and regulation in solving problems relative to reliance on self, kin, and neighbors. Self-reliance is generally presumed to be the ideal, but this is a rebuttable presumption. “Liberal” issue rhetoric typically offers many rebuttals, usually involving distrust of “the market,” but only some of these rebuttals are grounded in realistic understanding of how markets do and do not work. “Conservative” issue rhetoric sometimes offers thoughtful defenses of “the market” but can also fall silent when favored business interests seek protectionist legislation. Because government as an institution is the chief alternative to private and community problem solving, liberals and conservatives alike ideologize the question of just how competent and trustworthy it is. Selective perception abounds on both sides of this argument, especially in today's polarized environment.

A growing source of issue rhetoric today is populism. Populists divide the world into “the elite” and “the people.” Right-wing populists cast blame on cultural elites and credentialed experts. Left-wing populists argue that plutocrats are “rigging the economy.” Populist rhetoric can sometimes give voice to the frustrations of ordinary people but rarely suffices to settle concrete issues of problem definition and policy design.

Policy analysis typically bridges all political ideologies by reliance on the normative standard of “maximizing welfare” and on social science theorizing and evidence about the comparative advantages of different institutions for different purposes. Thus you want not simply to echo the issue rhetoric in your problem definition, but to use it as raw material for a provisional problem definition that you hope will prove analytically useful.

Note also that some issue labels may signify more than one problem. Depending on the audience, for example, “teenage pregnancy” may connote any or all of the following conditions: sexual immorality, the blighting of young people’s and their children’s life chances, exploitation of taxpayers, and social disintegration. Usually you will want to determine a primary problem focus, to ensure that the analysis does not get out of hand. But if the problems aren’t too complicated, you may feel willing to define more than one.

### “Uncertainty” Is the Problem That Evaluation Addresses

If you are evaluating how well some policy or program has been working, what “problem” are you working on? How does evaluation fit into the Eightfold Path framework?

Like all policy analysis, your work here is answering questions about the future. True, you are looking at the past, but the intention is to use your conclusions to shape future action. Depending on your assessment of past performance, the typical future action could be to expand the program, cut the program back, kill the program altogether, start it up in some additional site, or modify it in some way. But what exactly is past performance, how does this performance measure against evaluative criteria, and what aspects of program design and implementation seem to have produced that outcome? It is this uncertainty that your evaluative work is addressing; therefore “too much uncertainty” is the problem.

### Quantify If Possible

Your problem definition should, insofar as possible, include a quantitative feature. Assertions of deficit or excess should come with *magnitudes* attached. How big is “too big”? How small is “too small”? How about “too slowly” or “too fast”? With regard to homelessness, how many homeless people are there in the United States? Or in the case of agricultural water, how many acre-feet of water are used now, and how does that amount compare with the demand in some specified future year (given certain assumptions about water pricing)? Exactly what is “our ability to develop physical facilities for water storage,” and how do we expect it to grow, or shrink, over time?

If necessary, gather information to help you calibrate the relevant magnitudes. (See the discussion under “Step Two: Assemble Some Evidence.”)

In many or most cases, you will have to estimate—or, more likely, “guesstimate”—the magnitudes in question. Sometimes you should furnish a *range* as well as a *point*

estimate of magnitudes—for example, “Our best guess of the number of homeless persons in families is 250,000, although the truth could lie between 100,000 and 400,000.”

Even if you cannot come up with good numbers yourself, qualitatively defining a *metric* that might be used to quantify the problem helps you make your problem definition more behavioral and concrete. It is better to say, “Too many people with annual incomes over \$60,000 are living in subsidized apartments,” than simply, “Too many relatively well-off people are taking advantage of low-rent public housing.” The \$60,000 value provides desirable texture and information about a threshold number that will serve in the promised analysis.

## Diagnose Conditions That Cause Problems

Some problematic conditions are not experienced as troublesome per se by citizens but are perceived by them, or by analysts working on their behalf, to be causes of trouble. It is sometimes useful to diagnose at least one alleged condition of this type and to define it as a problem to be mitigated or removed—as in “One of the problems in the air pollution area is that states have not been willing to force motorists to keep their engines tuned up and their exhaust systems in proper order.”

**Semantic Tip** Note that this sort of problem definition is not merely descriptive but is also diagnostic. It implicitly asserts that some condition, which people may or may not find troubling on its own, is an important cause of some other condition that is indeed troubling. Problem definitions that pretend to such diagnostic power can be useful, but they can also be treacherous. Suppose, after all, that the causal diagnosis is mistaken or misleading—for example, that states’ unwillingness to enforce engine maintenance routines is *not* in fact a very important cause of air pollution. Because the term *definition* in some contexts connotes legitimate arbitrariness (“I’ll define *justice* to mean . . .”), the causal claims implicit in diagnostic problem definitions can easily escape needed scrutiny. (See “Step Five: Project the Outcomes” for further discussion.)

## Risky Conditions: “The Odds”

“The odds are too high that this nuclear reactor will suffer an accident in the next twenty-five years that will emit excessive radiation.” This sentence does indicate a problem, but it is not something tangible, like “Too many cases of asthma are being reported in this neighborhood.” It refers to risk and is stated in probabilistic language dealing with “the odds.”

**Semantic Tip** Referring to the odds is a useful way to talk about anything that is uncertain in your analysis (not just the problem definition) where the probabilities of outcomes can be approximately described or at least debated. It can also refer to the risk that an alternative will not work out as planned, or the likelihood that a key political actor will remain in office in order to oversee policy implementation. It is an especially useful locution when talking about risks that are particularly resistant to precise quantification—for example, “The odds are that the US nuclear modernization

program is causing other countries to look more favorably on acquiring nuclear weaponry themselves.”

The odds formulation can also be used for specifying criteria. For instance, one could say that one criterion is “Maximize the odds that members of the Freedom Caucus will hold a majority on the Ways and Means Committee following the next election,” or “Minimize the odds that the health department’s new computer system for verifying benefit eligibility will crash upon rollout.”

## Work on Hypothetical Problems—Up to a Point

Often “the problem” is implicit in a statement by the client (or some concerned group) that if only some alternative practice (“solution,” in a sense) were in place, the world would be much better off. The analyst is charged with evaluating the merits of this supposition. For example, “If we’d had an up-to-date purchasing department, we would have anticipated this price increase and stocked up on X beforehand.” Because Purchasing has allegedly been slack with regard to anticipated price movements, public money has been wasted. But there is a potential confusion lurking here: This is a useful problem definition only if the allegation is true. If it is not true, this problem does not exist. Should the analyst go off in search of a solution to a problem that does not really exist?

The simple solution is to conceive of “the problem” as “hypothetical” (or “possible”) rather than actual. It is perfectly reasonable to study a hypothetical problem while not committing oneself to a belief in its reality just yet. Commitment is deferred until the study is completed or nearly completed.

The idea of a “hypothetical” problem implies a troublesome question: Of the billions of “hypothetical problems” in the world, how do we recognize, and characterize, this one? Primarily, it is implied by the statement of the supposed solution. The bundle of hypothetical problems implied by “lack of an up-to-date purchasing department” is not so very large, and it is even further focused by the particular example given by the client, being obliged to pay a price for X that is higher than would have been necessary. In the real world, certain policy areas seem to generate more of these hypothetical problems than others. The leading one concerns waste and inefficiency: “If we do things this way, the results will be more efficient.” Or: “Currently, our procedures waste a lot of time going back and forth, checking and rechecking, whereas that would all be minimized were we to do Y.” Hence, if someone alleges that the failure to do something is a problem, see if you can reframe “the problem” as “(possibly) too much waste.” Note the waste does not have to be of money only—it could be time or opportunities to improve output in a cost-effective way.

## Identify Latent Opportunities

A special kind of problem is an opportunity missed. Is it not rather small-minded to think of policy analysis as devoted merely to the amelioration of problems? Might

policy analysis not rise above the tedious and uninspiring business of patching and fixing? Can we not aspire to a world in which we can identify opportunities to do creative—not to say wonderful—things? “If it ain’t broke, don’t fix it” is a confining idea, and certainly policy analysts, policymakers, and public managers ought not to allow the problem focus to restrict the search for plausible opportunities. Unfortunately, the working agenda of most policy professionals is set by complaints, threats, worries, and troubles—often leaving little time or energy to think about improvements that no one has identified as needful. Still, if latent opportunities are really lying around, it would be a pity to ignore them.

Where do we find opportunities for creative policy improvements that haven’t first been identified by complaints, threats, and so on? Relatively little academic or technical theory is available to answer this question.<sup>5</sup> Box I-1 contains a list that is suggestive.

## BOX I-1: SOME GENERIC OPPORTUNITIES FOR SOCIAL IMPROVEMENT THAT OFTEN GO UNNOTICED

**Reducing administrative burdens.** *Many government programs make it difficult for people to apply for or receive benefits or assistance for which they are entitled. Reducing unnecessary administrative burdens can save people time, effort, and money, and reduce the stress that people may experience when interacting with public agencies.*

**Designing the architecture of choice.** *By varying the ways in which choices are presented to people, it may be possible to overcome cognitive biases that lead to poor decision-making. For instance, flipping the preselected, default choice (the choice people automatically receive if they do not actively indicate a preference) from “opt out” to “opt in” can increase participation rates in organ donation programs and employee savings plans.*

**Social norms marketing.** *People often wish to follow social norms, but they can’t do so if they are unaware of them. For example, many college students believe they drink less alcohol than the average—and increase their consumption to be more like their peers. When the true drinking rate is disseminated, peer pressure to binge is greatly reduced.*

**Internalizing the social effects of individual decisions.** *Many opportunities exist to improve social welfare by removing incentives for individuals to ignore the spillover costs of their decisions. For example, introducing congestion tolls would reduce traffic congestion by discouraging drivers from using roads during peak hours.*

**Operations research strategies.** *By means of sequencing, timing, prioritizing, matching, clustering, and other such rationalizing arrangements, it may be possible to use a fixed stock of resources to achieve higher productivity than is possible otherwise. For instance, provided that traffic flow conditions are within certain parameters, high-occupancy-vehicle (HOV) lanes can maximize vehicle throughput in a fixed section of roadway.*

**Cost-based pricing.** Discrepancies between prices and real costs present an opportunity for enhancing social welfare by adjusting prices to better reflect reality. For instance, removing rent controls would bring prices more into line with real housing costs.

**By-products of personal aspirations.** It is possible to structure new incentives or create new opportunities for personal advantage or satisfaction that can indirectly result in social benefit. For example, public-sector employers can offer to share the benefits of cost-reducing innovations with the employees who conceive them and implement them.

**Complementarity.** Two or more activities can potentially be joined so that each may make the other more productive. For example, increased public works construction can combat unemployment.

**Input substitution.** The world abounds in opportunities to substitute less costly inputs in a current production process while achieving roughly equivalent results. For instance, municipalities can hire lower-paid civilians to perform police clerical tasks rather than use expensive uniformed officers.

**Development.** A sequence of activities or operations may be arranged to take advantage of a developmental process. For example, a welfare agency can assess clients for employability and vocational interest before, rather than after, sending them out to search for a job.

**Exchange.** Unrealized possibilities for exchange can increase social value. Policymakers typically design policies to simulate market-like arrangements—for example, conducting pollution permit auctions or reimbursing an agency for services it renders to another agency's clients or customers.

**Multiple functions.** A system can be designed so that one feature has the potential to perform two or more functions. For example, a tax administrator can dramatize an enforcement case in such a way as both to deter potential violators and to reassure non-violators that they are not being played for suckers because of their honesty.

**Nontraditional participants.** Line-level employees of public agencies—as well as their customers, their clients, or the parties whom they regulate—often have knowledge of potential program improvements that could usefully be incorporated into the agencies' policies and operations. The Internal Revenue Service (IRS), for instance, has sought feedback from ordinary tax filers about how to improve federal tax forms.

**Underutilized capacity.** Governments sometimes systematically underutilize resources at their disposal. In many communities, school facilities are used for relatively limited purposes for only part of the day and for only part of the year—although school officials would be quick to warn that tapping this capacity without harming school functions is not always easy.

## Avoid Common Pitfalls in Problem Definition

Problem definition is a step beset by at least three pitfalls.

**Semantic Tip** *Defining the solution into the “problem.”*

Your problem definition should not include an implicit solution introduced by semantic carelessness. Projected solutions must be evaluated empirically and not

legitimated merely by definition. Therefore, keep the problem definition stripped down to a mere description, and leave it open where you will look for solutions.

- *Don't say:* "There is too little shelter for homeless families." Inadvertently implying that "more shelter" is the best solution may inhibit you from thinking about ways to prevent families from becoming homeless in the first place. *Try instead:* "Too many families are homeless."
- *Don't say:* "New schools are being built too slowly." Simply assuming that "more schools" is the solution may inhibit you from thinking about ways to use existing facilities more efficiently or even to try forms of "distance learning." *Try instead:* "There are too many schoolchildren relative to the currently available classroom space."

A tip-off that you're probably smuggling an implicit solution into the problem definition is to hear yourself saying, "Aha, but that's not the real problem; the real problem is . . ." While there are better and worse ways to conceptualize a problem—or to solve a problem—it stretches ordinary usage too much to say that one problem could be "more (or less) real" than another.

### Accepting Too Easily the Causal Claims Implicit in Diagnostic Problem Definitions

We suggested earlier that conditions that cause problems may also be problems themselves. However, the causes must be real, not merely assumed. You have to evaluate the causal chain that goes from the situation itself to the bad effects it is alleged to cause, and to convince yourself that the causal relationship is real. For instance, for some people, cocaine use is not a problem in itself, but it may become a problem if it *leads* to crime, poor health, family disintegration, and so on. But does it lead to these outcomes, and to what degree? The evidence on this question should be evaluated very carefully before you decide that it's okay to work with a problem definition that sounds like "too much cocaine use." (See Appendix F on the use of experimental methods in policy analysis.)

It is easy to get causal attributions wrong and then follow a suboptimal path of searching and reasoning. Consider the problem of low vaccination rates. In some parts of the country, a significant fraction of parents delay or refuse to vaccinate their children out of a fear that vaccination causes autism. When local vaccination rates fall below 90–95 percent, communities lack "herd immunity," leaving children at risk of contracting vaccine-preventable diseases like measles. Given this, it is tempting to work with a problem definition like "too many parents believe vaccines cause autism" and focus on options to reduce vaccination misperceptions, such as public health information campaigns to educate people about the true benefits and harms of immunizing their children. But the presumed causal chain that goes from parental beliefs to

vaccination rates may be wrong. And in fact studies have shown that pro-vaccination messaging may not be effective and can even backfire.<sup>6</sup> Further, research has shown that imposing strict school vaccination rules (which make it harder for parents to enroll children who haven't received required immunizations) significantly boosts vaccination rates.<sup>7</sup> In other words, changing people's minds about vaccinations may not be required to change their behavior.

### **Ignoring the Context of the Problem**

Context makes a difference. Possible solutions that work in one place fail in another, and vice versa. This is not just because many public policy problems are very difficult, and often overwhelming, but because some contexts are favorable to a particular strategy while others are not. If context is likely to matter—and, even though it mostly does, that is not always the case—and problem contexts differ, problem definition should recognize this by specifying the contextual conditions under which the problem is likely to be encountered. For example, if public trust and confidence in the police is significantly influenced by neighborhood crime rates, the problem of “too little satisfaction with the police” should be defined at the local rather than national level.

### **The “Last Mile” Problem**

Until now our focus has been on what we really want to accomplish. What market failure or government failure do we wish to solve? What value do we seek to advance? These are important questions, but our answers will count for little if preferred policy solutions cannot be implemented. We call the various and sundry administrative, political, organizational, and legal challenges associated with making policies work on the ground “Last Mile” problems.

Consider a government's plan to lay an underwater cable from shore to shore. Eventually, the government will have designed, built, and laid down almost all of it. All except the Last Mile. Your grand design and all of yours and everybody else's heroic labor are of little value unless you can take that cable the Last Mile. In that Last Mile there are permitting agencies, neighborhood preservationists, endangered species habitats, local city council members, taxpayer-protecting city and county budget bureaus, and more that you can't even imagine at this point.

Last Mile problems are ubiquitous in policymaking and largely spring from the cherished institutions and practices of liberal democracy itself. They involve all those moving parts of bureaucratic imperialism and turf protection and all the nonmoving parts as well, which are made up of bureaucratic lethargy, obstructionism, incompetence, and fear. They also involve multiple and often competing goals, rent-seeking, and partisanship.

A program ultimately must be administered, and this typically means turning it over to some agency of government. But which agency? One that will be eager to

administer it, of course, and will in the long run try to protect it from hostile interests both within and without government. The agency should also have the sheer administrative capacity to do the job—for example, the accounting, computational, and human resources to handle what might prove to be complex tasks.

A common Last Mile problem is that the designated agency views its existing program and administrative responsibilities as having much greater priority than some new program landing in its lap. Care must be taken to provide enough new resources so that the agency does not simply go through the motions of trying to implement it.

Certain semipredictable long-run bureaucratic problems can be anticipated and certain safeguards installed, though their effectiveness is always somewhat unpredictable. Regulatory agencies might become captive of the firms and other interests they regulate. *Rent-seeking* is the common name for this process, a formerly technical term meaning profiting by efforts to impose anticompetitive regulatory rules and enforcement procedures on others in the same industry. Rent-seeking also plagues service delivery agencies, although in this case the captors are typically private-sector nongovernmental organizations, or NGOs, and for-profit contractors that accept government funds in exchange for actually delivering, or appearing to deliver, the services.

Another common problem is sclerosis: as it grows older, the agency becomes progressively more inflexible and procedure-bound, lacking in energy and imagination. We also cannot omit “WFA,” being beset by waste, fraud, and abuse.

Federalism is a major contributor to Last Mile problems. National programs that require the buy-in of state governments give opponents multiple venues in which to block implementation. For example, when Congress enacted the Affordable Care Act in 2010, advocates of Medicaid expansion believed they had won. As of 2023, however, ten states have not adopted Medicaid expansion plans for their residents, leaving millions without health insurance.

## **Iterate**

Defining the problem is a crucial step. (See Box I-2 for an illustrative specimen of a problem definition from a policy analysis report.) Because problem definition is hard to get right, however, you may take that same step again and again. Also, your empirical and conceptual understanding will evolve over the course of your analytic work. For instance, you may start out thinking that the main problem is “too many halfway houses for people with mental illness in our city” but end up concluding that the main problem is how badly some of these residences are managed.<sup>8</sup> As you begin to rule out alternative approaches to solving or mitigating your problem, you will probably want to sculpt the problem definition so that, in the end, you and the political system will have some chance of attacking the problem successfully. Finally,

if you are working in an office or agency context, you will implicitly be negotiating a mutually acceptable problem definition with your analyst colleagues and your hierarchical superiors.<sup>9</sup>

## BOX I-2: AN ILLUSTRATIVE EXAMPLE OF “DEFINING THE PROBLEM” FROM A POLICY ANALYSIS REPORT

*The supply of skilled workers is not keeping up with the demand for them (Goldin and Katz 2012). Employers report shortages of workers with occupation-specific skills (Holzer et al. 2011). A recent survey of 2,000 U.S. companies found that 30 percent had been unable to fill skilled job positions for more than six months (Manyika et al. 2012).*

*Many low-income workers would not be able to access vocational training without assistance from government programs. Although the vast majority of vocational training in the United States is provided by employers (Mikelson and Nightingale 2004), employers are less likely to provide training for their lower-skilled positions, which tend to have higher rates of turnover (Lane 2000). Hypothetically, workers could pay for their own training, but many unemployed and low-skill workers do not have the financial resources or the ability to borrow to pay for training.*

*The United States does not currently invest heavily in vocational training compared with other countries, and funding for vocational training has declined over the past decades. Whereas the United States spends less than 0.05 percent of its gross domestic product on vocational training, other industrialized nations invest up to ten times as much. Since 1985 the amount budgeted for key U.S. Department of Labor training programs has declined by about 20 percent in real terms.*

*Even among supporters of vocational training, there is legitimate concern that many people who start programs do not complete them. Within three years of enrollment in a community college, fewer than half of all enrollees have attained an associate’s degree or vocational certificate, transferred to a four-year institution, or remain in college (Horn and Weko 2009). Only about 55 percent of the people who begin two-year colleges obtain either an associate’s degree or a certificate (Holzer and Dunlap 2013). Analysis of data on training vouchers provided by the WIA [Workforce Investment Act] Adult and Dislocated Worker programs found that only 64 percent of workers who enrolled in training programs at community colleges completed a training program within three years (Perez-Johnson, Moore, and Santillano 2011). Although the rate of completion for those enrolled in training at a private training provider was higher, about 15 percent of trainees still did not complete a training program within three years.*

*A second concern is that too many workers who complete training cannot subsequently find a job to use the acquired skills. A study of training vouchers provided through the WIA Adult and Dislocated Worker programs reported that only about 40 percent of the participants found employment in the occupation for which they*

received training (Perez-Johnson, Moore, and Santillano 2011). Similarly, a study of the Trade Adjustment Assistance program found that only 37 percent of people who participated in training funded by that program held a job in the occupation for which they were trained in the fourth year after they were initially laid off (Schochet et al. 2012). These statistics suggest that there is often a missing link between employers and training programs.

---

Source: Sheena McConnell, Irma Perez-Johnson, and Jillian Berk, "Proposal 9: Providing Disadvantaged Workers with Skills to Succeed in the Labor Market," in *Policies to Address Poverty in America*, ed. Melissa S. Kearney and Benjamin H. Harris (Washington, DC: The Hamilton Project, June 2014), 98, [http://www.hamiltonproject.org/assets/files/policies\\_address\\_poverty\\_in\\_america\\_full\\_book.pdf](http://www.hamiltonproject.org/assets/files/policies_address_poverty_in_america_full_book.pdf).

Note: Endnotes and references to figures omitted from excerpt.

## STEP TWO: ASSEMBLE SOME EVIDENCE

All of your time doing a policy analysis is spent on two activities: thinking (sometimes aloud and sometimes with others) and hustling data that can be turned into evidence. Of these two activities, thinking is generally the more important, but hustling data takes much more time: reading documents, hunting in libraries, poring over studies and statistics, interviewing people, traveling to interviews, waiting for appointments, and so on.

The real-world settings in which policy analysis is done rarely afford the time for a research effort that would please a careful academic researcher. In fact, time pressure is probably almost as dangerous an enemy of high-quality policy analysis as is politically motivated bias, if not more so. Therefore, economize on your data collection activities. The key to economizing is this: try to collect only those data that can be turned into "information" that, in turn, can be converted into "evidence" that has some bearing on your problem.

**Semantic Tip** For the logically minded, here are some definitions: *Data* are facts—or, some might say, representations of facts—about the world. Data include all sorts of statistics but go well beyond statistics, too. Data also include, for instance, facts about an agency manager's ability to deal constructively with the press. *Information* consists of data that have "meaning," in the sense that they can help you sort the world into different logical or empirical categories. The prevalence of cigarette smoking in five different countries constitutes data, but these data become information when you decide it is interesting to array the countries comparatively (e.g., from lowest to highest prevalence). *Evidence* is information that affects the existing beliefs of important people (including yourself) about significant features of the problem you are studying and how it might be solved or mitigated. Differential prevalence of smoking, for instance, can become evidence bearing on hypotheses regarding different levels of concern about personal health across countries.

You need evidence for three principal purposes, all of which are relevant to the goal of producing realistic projections of possible policy outcomes. One purpose is to assess the nature and extent of the problem(s) you are trying to define. A second is to assess the particular features of the concrete policy situation you are engaged in studying. For instance, you may need to know—or guess—about agency workloads, recent budget figures, demographic changes in a service area, the political ideology of the agency chief, the competency of the middle-level managers in the agency, and the current attitude of some other agency that nominally cooperates with yours on some problem. The third purpose is to assess policies that have been thought, by at least some people, to have worked effectively in situations apparently similar to your own, in other jurisdictions, perhaps, or at other times. (Sometimes these situations will have been evaluated statistically and sometimes not: see Part IV, “Smart (Best) Practices’ Research: Understanding and Making Use of What Look Like Good Ideas from Somewhere Else.”)

Because each of these purposes becomes salient in different phases of the policy analysis process, the second step on the Eightfold Path, “Assemble Some Evidence,” will be taken more than once but with a different focus each time.

### **Think before You Collect**

Thinking and collecting data are complementary activities: You can be a much more efficient collector of data if you think, and keep on thinking, about what you do and don’t need (or want) to know, and why. The principal—and exceedingly common—mistake made by beginners and veterans alike is to spend time collecting data that have little or no potential to be developed into evidence concerning anything you actually care about. People often do this because running around collecting data looks and feels productive, whereas first-rate thinking is hard and frustrating. Also, when they see you busily collecting data, the people paying for your work tend to be reassured that somehow they are getting their money’s worth.

### **The Value of Evidence**

Since most evidence is costly to produce, you must weigh its likely cost against its likely value. How is its likely value to be estimated? The answer may be cast in a decision-analytic framework (decision trees), though remember that the process of making a decision involves a great many elements prior to the moment of actual choice, such as defining a useful problem, thinking up better candidate solutions, and selecting a useful model. In general, the value of any piece of evidence depends on these factors:

- The likelihood that it will cause you to substitute some better decision for whatever decision you would have made without it (which might have been an “acceptable” decision in and of itself)

- The likelihood that the substituted decision will, directly or indirectly, produce a better policy outcome than the outcome that would have been produced by the original decision
- The magnitude of the difference in value between the likely-to-be-improved outcome and the original outcome

## The Utility of Research

Although evidence is costly to produce, there are instances when a high price may be worth paying. If the evidence required to understand a situation does not exist but could be assembled, analysts may choose to invest in substantial data generation efforts, which may range from qualitative interviews and case studies to quantitative modeling. (Box I-3 provides an example of modeling using a hard-edged engineering approach to estimate the greenhouse gas emissions from global oils. The evidence produced by the model is intended to serve as an input into the design of climate mitigation strategies for the “unconventional” energy resources being unleashed through technologies such as fracking.)

### BOX I-3: AN ILLUSTRATIVE EXAMPLE OF “ASSEMBLING SOME EVIDENCE” FROM A POLICY ANALYSIS REPORT

*The character of oil is changing. Consumers may not notice the transformation—prices have fluctuated, but little else appears to have changed at the gas pump. Behind the scenes, though, the definition of oil is shifting in substantial ways. There is oil trapped tightly in shale rock, and oil pooled many miles below the oceans. Oil can be found in boreal forests, Arctic permafrost, and isolated geologic formations. Some oils are as thick as molasses or as gummy as tar, while others are solid or contain vastly more water or gas than normal.*

*Oil resources were once fairly homogeneous, produced using conventional means and refined into a limited number of end products by relatively simple methods. This is no longer the case. Advancements in technology mean that a wider array of hydrocarbon deposits in once-unreachable areas are now viable, extractable resources. And the techniques to turn these unconventional oils into petroleum products are becoming increasingly complex.*

*As oil is changing, so, too, is the global climate. The year 2014 ranked as the earth’s warmest since 1880. Fossil fuels—oil along with coal and methane gas—are the major culprits.*

*The only way to determine the climate impacts of these previously untapped resources—and to compare how they stack up against one another—is to assess their*

greenhouse gas (GHG) emissions at each stage in the oil supply chain: exploration, extraction, processing, refining, transport, and end use. The more energy it takes to carry out these processes, the greater the impact on the climate. And in the extreme case of some of these oils, it may take nearly as much energy to produce, refine, and transport them as they provide to consumers. Moreover, each oil yields a different slate of petroleum products with different combustion characteristics and climate footprints.

The Oil-Climate Index (OCI) is a metric that takes into account the total life-cycle GHG emissions of individual oils—from upstream extraction to midstream refining to downstream end use. It offers a powerful, yet user-friendly, tool that allows investors, policymakers, industry, the public, and other stakeholders to compare crudes and assess their climate consequences both before development decisions are made as well as once operations are in progress. The Oil-Climate Index will also inform oil and climate policy making.

The index highlights two central facts: The fate of the entire oil barrel is critical to understanding and designing policies that reduce a crude oil's climate impacts. And oils' different climate impacts are not currently identified or priced into the market value of competing crudes or their petroleum products. As such, different oils may in fact entail very different carbon risks for resource owners or developers.

Analysis of the first 30 test oils to be modeled with the index reveals that emission differences between oils are far greater than currently acknowledged. Wide emission ranges exist whether values are calculated per barrel of crude, per megajoule of products, or per dollar value of products, and it is expected that these emission ranges could grow as new, unconventional oils are identified.

There are several critical variables that lead to these variations in oils' life-cycle climate emissions. They include how gas trapped with the oil is handled by producers, whether significant steam is required for oil production, if a lot of water is present as the oil reservoir depletes, how heavy (viscous) or deep the oil is, what type of refinery is used, and whether bottom-of-the-barrel products like petroleum coke (known as petcoke) are combusted. Given these factors, the most climate-intensive oils currently identified—gassy oils, heavy oils, watery and depleted oils, and extreme oils—require special attention from investors, operators, and policymakers.

---

Source: Deborah Gordon, Adam Brandt, Joule Bergerson, and Jonathan Koomey, *Know Your Oil: Creating a Global Oil-Climate Index* (Washington, DC: Carnegie Endowment for International Peace, 2015), 3–4, [https://carnegieendowment.org/files/know\\_your\\_oil.pdf](https://carnegieendowment.org/files/know_your_oil.pdf).

Note: See another excerpt from this report on coping with uncertainty in Step Five.

## The Utility of an Educated Guess

It is surprising how well you can do in many cases by gathering no evidence at all but simply sitting down and thinking something through and then making some serious educated guesses. There is nothing shameful about acting on such guesstimates and thereby conserving your data-collecting time and energies for answering questions for which good evidence is really necessary (see Part II, “Assembling Evidence”).

A helpful check on your thinking, to avoid collecting useless data, is to ask yourself the following questions before embarking on some data collection venture:

- “Suppose the data turn out to look like so-and-so as opposed to thus-and-such. What implication would that have for my understanding of how to solve this problem?”
- “Compared to my best guess about how the data will look once I’ve got them, how different might they look if I actually took the trouble to get them?”
- “How much is it worth to me to confirm the actual difference between what I can guess and what I can learn about the world by really getting the data?”

It is this sort of critical attitude about the value of expensive data collection (especially ad hoc surveys and “needs assessments”!) that often leads good and experienced policy analysts to make do with back-of-the-envelope estimates. However, none of this reasoning is meant to be an excuse for shirking the job of getting good data—and sometimes a lot of data, at huge costs in time and money—when you’ve convinced yourself that the investment really will pay off. There’s an obvious and critical difference between justifiable and unjustifiable guesstimates.

## Review the Available Literature

There hardly exists a problem on whose causes and solutions some academic discipline or professional association is not doing research. It is easy to find journals and various professional publications disseminating empirical results, theories, case studies, and so on. The internet brings much of this literature to your desktop. Studies vary in their quality, rigor, and internal and external validity. Policy analysts should evaluate research with a critical eye, focusing not only on a study’s “bottom line” but also on the strengths and limitations of its research design and, especially, on the relevance of its findings to the problem-solving tasks at hand.<sup>10</sup>

Advocacy organizations often publish a great deal of interesting work and may take special pains to disseminate their findings on the internet. However, because advocacy-based analyses are not, in general, as reliable as more disinterested work, there is a danger of relying too much on such sources just because they are readily available.

## Survey “Best Practices”

The chances are good that the problem you are studying is not unique and that policymakers and public managers in other jurisdictions, perhaps not very different from the one you are studying, have already dealt with it in some fashion. See if you can track down some of these past solutions and extrapolate them to the situation you are studying. Bear in mind, however, that the extrapolation process is complicated (see Part IV, “‘Smart (Best) Practices’ Research”).

## Use Analogies

Sometimes it pays to gather data about things that, on the surface, seem quite unlike the problem you are studying but, on a deeper level, show instructive similarities. For instance, your understanding of how a merit pay plan for compensating managers in the public sector might work could perhaps be improved by seeing how similar schemes work in the private sector. Or, if you are working on the problem of how a state can discipline, and perhaps disbar, incompetent attorneys, you might usefully spend a good deal of your time learning about how the medical profession handles problems of physician incompetence. If you are working on how to reduce neighborhoods' resistance to accepting low-income housing projects, you could usefully look into the literature on community resistance to accepting solid-waste incinerators.

As these examples suggest, some analogies are easier to perceive, and to make sense of, than others. It takes a little imagination to see an instructive analogy and, occasionally, is a little daring to try to convince others to recognize both its usefulness and its inevitable limitations.

## Start Early

You are often dependent on the very busy schedules of other people whom you ask to furnish information or to make time for an interview. It is extremely important to submit requests for information—and especially for interviews—well in advance of when you want to have completed the data collection. (For a useful description of how to conduct literature reviews, library searches, phone interviews, and personal interviews, see Weimer and Vining 2017, chap. 14; see also Part II, “Assembling Evidence.”)

## Touch Base, Gain Credibility, Broker Consensus

The process of assembling evidence inevitably has a political as well as a purely analytic purpose. Sometimes it entails touching base with potential critics of your work so that they will not be able to complain later that you have ignored their perspectives. Conversely, by making yourself known to potential supporters of your work, you may be able to create a cadre of defenders.

A more complex objective, where appropriate, might be to blend policy analysis with the process of improving a policy idea or decision during the course of implementation. (See the following discussion of “improvability” as a practical criterion.) This objective entails obtaining feedback from participants, usually in an iterative process, and sharing some of your own reactions with them. You thereby become more of a partner in the process than an outside observer and diagnostician. An even more complex and challenging role would be for you to become a particular type of “partner,” a facilitator and broker, whether by acting as a conduit from one person to another or by convening meetings and other gatherings.

## Free the Captive Mind

In exchange for access to data and a ready-made worldview, researchers sometimes uncritically accept problem definitions and preferred solutions from kindly informants (not to mention from paying clients or employers). To counter such temptations, be sure to make contact with individuals or factions whom you would expect to disagree—the more sharply the better—with those informants. A time-saving, but only partial, substitute is to ask your kindly informants, “Who might object strongly to your point of view about this, and why might they do so?” In today’s era of polarization, this may require you to reach out to people outside of the professional and social networks in which you typically travel. Remember that you do not have to share someone’s values or outlook to learn from them. By engaging people who have a different perspective on a situation, you may gain a new insight into a problem, a new idea for a policy alternative, or a better understanding of how a proposed solution will work on the ground.

## The Last Mile

Most of the evidence you will need as a policy analyst or designer comes from “past experience.” Demonstration projects and systematic evaluations of ongoing program components, from your own jurisdiction or from elsewhere, are examples. So is informal, casual appraisal of what has happened in similar jurisdictions or with comparable programs.

It is surprising, though, how little inclined policy designers are to consider other jurisdictions. “We are unique, and no one else has our problems (resources).” Or policymakers, giving into their political biases, may decline to examine what can be learned from the experiences of jurisdictions that face the same challenges but have different partisan leanings (in a polarized age, “blue states” may be unwilling to learn from the implementation experiences of “red states,” and vice versa).<sup>11</sup> These are excuses for intellectual laziness; however, some people are simply not up to disentangling the commonalities and the singularities of various situations.

Still, while there can be no such thing as “evidence” about future events—by definition—it is possible to venture estimates about the administrative issues to be faced in the Last Mile simply by interviewing assumedly relevant senior managers about the prospects.

## STEP THREE: CONSTRUCT THE ALTERNATIVES

---

By *alternatives*, we mean something like “policy options,” or “alternative courses of action,” or “alternative strategies of intervention to solve or mitigate the problem.”

For example, the Congressional Budget Office (CBO) in 2013 analyzed four alternatives to the US Army’s plan to develop a new Ground Combat Vehicle (GCV) as a replacement for the Bradley Infantry Fighting Vehicle (IFV): (1) purchase the Israeli Namer armored personnel carrier, (2) upgrade the Bradley IFV, (3) purchase the

German Puma IFV, and (4) retain the current Bradley IFV.<sup>12</sup> See Box I-4 for an illustrative example.

## Beware a Linguistic Pitfall

**Semantic Tip** Specifying alternatives does not necessarily signify that the policy options are mutually exclusive. Policy analysts use the term *alternative* ambiguously: sometimes it means one choice that implies forgoing another, and sometimes it means simply one more policy action that might help to solve or mitigate a problem, perhaps in conjunction with other alternatives. Be aware of the ambiguity in other people's usage, and in telling your story (see Step Eight), be sure that no such ambiguity enters your own usage.

If it is impossible to implement two or more options separately—such as an increase in the length of the school day and a restructuring of after-school programs—it usually makes sense to combine them into a single policy option. Sometimes, though, you won't be entirely sure whether two alternatives are or are not mutually exclusive. For instance, although the mayor may have promised enough money to either fix potholes or provide homeless shelters (but not both), you may have made such a great case for both programs that the mayor may decide to increase the budgetary allocation. See the subsection under “Step Six: Confront the Trade-Offs” that advises you to rank your list of preferred alternatives so that it is up to the decision-maker to decide when enough is enough.

## Start Comprehensive, End Up Focused

In the last stages of your analysis, you won't want to be assessing more than two or three principal alternatives, but in the beginning, err on the side of comprehensiveness. Make a list of all the alternatives you might wish to consider in the course of your analysis. Later on, you will discard some obvious losers, combine others, and reorganize still others into a single “basic” alternative with one or more subsidiary “variants.” For your initial list, though, where should you turn for ideas?

One starting point would be to note the alternatives that key political actors are actively proposing or seem to have on their minds. These may include prominent people's pet ideas, institutions' inventories of “off-the-shelf” proposals that simply await a window of opportunity, and prepackaged proposals that party leaders or political ideologues are perennially advocating. Then you could try to *design* alternatives that might prove to be superior to the alternatives currently being discussed by the key political actors.

## Entertain Out-of-the-Box Solutions

It's good to brainstorm, to try to be creative. You might not produce much better ideas than those that other people have already advanced. But, then again, you might.

One way to coax your creativity is to refer to the checklist in Appendix B, “Things Governments Do.” For each entry on the list, ask yourself: “Might it make sense to try some version of this generic strategy to help mitigate this problem?” Because it is

a comprehensive list, the answer with respect to any single strategy will usually be no. Going through the list systematically is worthwhile, however. Because the list is not very long, with experience, you will need to spend only a few minutes to decide whether any ideas there might be worth considering further. (See also the valuable discussion about generic policy instruments in Weimer and Vining 2017, chap. 10.)

Another approach is to free your mind to consider unconventional, out-of-the-box solutions. To be sure, most of these ideas won't turn out to be workable for one reason or another; yet many good policy ideas used today (e.g., emissions trading programs) were considered odd or impractical when they were first proposed. Also, technological breakthroughs and changing social norms are continuously expanding the set of feasible solutions.

Following are some suggestions for coming up with a better mousetrap<sup>13</sup>:

- **Ask how you would solve a problem if cost were no object.** Of course, in the real world, cost *is* an object. But imagining it isn't can sometimes free the imagination. For example, if cost were no object, you would never wait in line at the post office—you would have an assistant mail your packages. But why couldn't there be a public website where citizens could find out the wait times at every post office—indeed, at every government agency that serves customers?
- **Ask where else it would work.** If a solution works in one context, maybe it will work in another. For example, a mathematical algorithm is used to match medical students with residency programs, minimizing the frequency of “bad” matches. Why shouldn't an algorithm also be used to match foster children or adoptive children with qualified adults who wish to be foster or adoptive parents?
- **Ask why not.** Many people are upset about the high cost of auto insurance. One reason people are right to complain is that insurance rates hardly vary with mileage. A flat-fee policy also discriminates against people with lower incomes, who tend to drive less. But why *shouldn't* insurance rates be charged on a per-mile basis? That would be fairer and more efficient. GPS devices could monitor the number of miles driven. To be sure, there are privacy concerns and other obstacles to overcome. But asking, “Why not?” often leads to creative thinking.
- **Ask whether less could be more.** When people generate ideas for how to solve a problem, they instinctively think about what they can add, such as a new program or activity. But sometimes the best way to improve a situation is to remove a component that is already there, such as a reporting requirement that creates administrative delays without materially improving accountability. As you develop a menu of policy alternatives, remember to consider subtractive options, not just additive ones.<sup>14</sup>

**Semantic Tip** Always include in your first approach to the problem the alternative “Let present trends (or ‘business as usual’) continue undisturbed.” You need to do this because the world is full of naturally occurring, ongoing changes, some of which may mitigate, or worsen, the problem on which you are working. (Note that we are not characterizing this alternative as “Do nothing.” It is not possible to do nothing or to “not decide.” Most of the trends in motion will probably persist and alter the problem, whether for better or for worse.)

To see if “natural” change will affect the scope of the problem, inspect its most common sources in the public policy environment: (1) political changes following elections, as well as changes induced by the prospect of having to contest an election; (2) changes in unemployment and inflation rates that accompany the business cycle; (3) the changing “tightness” or “looseness” of agency budgets caused by overall taxing and spending policies as well as by program features such as automatic cost-of-living increases; (4) demographic changes, such as population migration patterns and population “bulges” moving through certain age levels; and (5) changing technologies. In most cases, however, this “let-present-trends-continue” option will drop out of your final analysis. It follows that if you do your problem definition work well, you will end up with an important problem in your sights that in most cases can be mitigated to some degree by purposive action.

Another frequently helpful alternative is “Learn more.” This can be done by using pilot studies, or by looking around for examples of “smart practices” elsewhere (see Part IV), or by waiting for the future to get less murky, or perhaps by negotiating further with important players to ascertain what they might do under various contingencies. Don’t forget that there is a cost to waiting if, in the absence of further learning, you would have guessed “the right conclusion” anyway. Conversely, there is a cost to premature decision-making or action if you are likely to make a consequential mistake that could be corrected by further learning.

## BOX I-4: AN ILLUSTRATIVE EXAMPLE OF “CONSTRUCTING THE ALTERNATIVES” FROM A POLICY ANALYSIS REPORT

### Summary

*The Federal Housing Administration (FHA) insures the mortgages of people who might otherwise have trouble getting a loan, particularly first-time homebuyers and low-income borrowers seeking to purchase or refinance a home. During and just after the 2007–2009 recession, the share of mortgages insured by FHA grew rapidly as private lenders became more reluctant to provide home loans without an FHA guarantee of repayment. FHA’s expanded role in the mortgage insurance market ensured that*

borrowers could continue to have access to credit. However, like most other mortgage insurers, FHA experienced a spike in delinquencies and defaults by borrowers.

Recently, mortgage borrowers with good credit scores, large down payments, or low ratios of debt to income have started to see more options in the private market. The Congressional Budget Office [CBO] estimates that the share of FHA-insured mortgages going to such borrowers is likely to keep shrinking as credit standards in the private market continue to ease. That change would leave FHA with a riskier pool of borrowers, creating risk-management challenges similar to the ones that contributed to the agency's high levels of insurance claims and losses during the recession.

This report analyzes policy options to reduce FHA's exposure to risk from its program to guarantee single-family mortgages, including creating a larger role for private lenders and restricting the availability of FHA's guarantees. The options are designed to let FHA continue to fulfill its primary mission of ensuring access to credit for first-time homebuyers and low-income borrowers. . . .

### What Policy Options Did CBO Analyze?

Many changes have been proposed to reduce the cost of risk to the federal government from FHA's single-family mortgage guarantees. CBO analyzed illustrative versions of seven policy options, which generally represent the range of approaches that policy-makers and others have proposed:

- Guaranteeing some rather than all of the lender's losses on a defaulted mortgage;
- Increasing FHA's use of risk-based pricing to tailor up-front fees to the riskiness of specific borrowers;
- Adding a residual-income test to the requirements for an FHA-insured mortgage to better measure borrowers' ability to repay the loan (as the Department of Veterans Affairs does in its mortgage guarantee program);
- Reducing the limit on the size of a mortgage that FHA can guarantee;
- Restricting eligibility for FHA-insured mortgages only to first-time homebuyers and low- to moderate-income borrowers;
- Requiring some borrowers to receive mortgage counseling to help them better understand their financial obligations; and
- Providing a grant to help borrowers with their down payment, in exchange for which FHA would receive part of the increase in their home's value when it was sold.

Although some of those approaches would require action by lawmakers, several of the options could be implemented by FHA without legislation. In addition, certain options could be combined to change the nature of FHA's risk exposure or the composition of its guarantees. CBO did not examine the results of combining options.

---

Source: *Options to Manage FHA's Exposure to Risk from Guaranteeing Single-Family Mortgages* (Washington, DC: Congressional Budget Office, September 2017), 1-2, <https://www.cbo.gov/publication/53084>.

## Model the System in Which the Problem Is Located

We often think about alternative approaches to the problem as possible *interventions* in the system that holds the problem in place or keeps it going. Logically, it is not necessary to model the causes of a problem in order to cure it—pharmaceutical manufacturers can testify that many of their successful products work by unknown causal routes on conditions whose causes are not at all understood. But a good causal model is often quite useful for suggesting possible “intervention points.” This is especially true when the problem is embedded in a complex system of interacting forces, incentives, and constraints—which is usually the case.

Consider, for instance, a system that produces “too much traffic congestion” at some choke point such as a bridge or a tunnel. A sketch of the relevant causal model would include the demand for travel along the relevant route, the available alternative modes of travel, the amount of roadway capacity, and the price to users of roadway capacity. An efficient and simple—but usually politically unpopular—intervention might be to increase the price to users so as to reflect the degree to which each user contributes to congestion and increased travel times.

How self-conscious, elaborate, and rigorous should your causal model be? Many social scientists who devote themselves to policy analysis would hold, “The more so the better.” We say, “Yes, but . . .” Self-consciousness is highly desirable. Elaborateness (or comprehensiveness—in this case a near synonym) is desirable because it decreases the risk of missing important causal connections, but it can blur the analytic focus and blunt creativity in designing intervention strategies. Rigor is desirable if it prevents you from relying on unarticulated and false assumptions; its downside is that it may persuade you to exclude factors that are important—for instance, the personalities of certain actors—because you don’t know how to model their effect rigorously or because you have only hunches regarding the facts.

Many models are best thought of as elaborations of a fundamental metaphor. They can be mathematically precise or verbal and evocative. Some commonly used metaphors that are the bases for models of particular value in designing alternatives are discussed in the following sections.

### Market Models

The model of a market in which disaggregated suppliers exchange goods or services with disaggregated demanders can apply to unpriced goods and services. The main idea behind the market model is really equilibration through exchange. Hence, the market model can be applied to many phenomena other than the production and allocation of textbook goods such as widgets or apples.

For instance, you might try to understand the flow of patients into a state mental hospital system in terms of supply and demand: there is a fixed short-run “supply” of available beds in state hospitals and a per diem charge for each, and a complex

“demand” for their use generated by police departments, county psychiatric emergency units, judges, members of the public, and so on.

A standard intervention strategy for improving markets that are not working as well as they might is to find some way to raise or lower the prices faced by either suppliers or demanders.

### **Production Models**

Unfortunately, little academic literature has examined the operating logics of the common types of production systems found in public policy—such as command-and-control regulation, service provision, and all the others, which are briefly described in Appendix B, “Things Governments Do.” (However, see Weimer and Vining 2017, chap. 10, on “generic policies”; see also Salamon 2002.) In any case, the main concern in understanding production systems should be to identify the parameters whose values, when they move out of a certain range, make the systems most vulnerable to breakdown, fraud and abuse, egregious diseconomies, and the distortion of intended purpose. It is also helpful to know about those parameters that matter most when we try to upgrade a production system from mere adequacy to performance levels we might think of as “excellent” (see Part IV, “Smart (Best) Practices’ Research”).

Another way to look at production models is through optimization lenses. Operations research models—such as queuing, inventory management, and Markov processes—are relevant here.<sup>15</sup>

### **Conformity Models**

Conformity models describe a process by which individuals adapt the attitudes and actions of other people around them. Psychologists have identified three sources of conformity: automatic mimicry and imitation, normative influence (doing what others do to increase social acceptance), and informational influence (the crowd is often a good source of information about what is correct or appropriate). An understanding of conformity models can improve the effectiveness of many interventions.

For instance, information policies intended to encourage healthier or more socially desirable behavior—such as posters hung on the walls of an inner-city school warning students of the dangers of dropping out—often fail to change behavior because they inadvertently reinforce the message that the “bad” behavior is prevalent—and people like to do what is “normal” for their reference group. The key to designing more effective interventions is to leverage the tendency of people to think and act like people around them. For example, hotels have been able to significantly boost the percentage of guests who reuse bath towels (reducing water and energy use) by informing guests on signs in their rooms that reusing towels is a typical behavior of other hotel guests.

### **Evolutionary Models**

An evolutionary model describes a common process of change over time. It is constructed of three important subprocesses: variation among competitors, selection, and

retention. Suppose, for instance, that in an agency enforcing health-related standards in the workplace, the complaints disproportionately concerned visible and annoying problems that were not as hazardous to worker health as less visible and annoying problems. In this case, the evolutionary model suggests several plausible intervention points. The agency might try to educate workers to detect and complain about more serious problems, contriving thereby to swamp the less serious problems—thus changing the pool of “competitors.” It might start screening the complaints for their likelihood of being associated with more fruitful targets—thus changing the “selection mechanism.” Or it might attempt to persuade workers, and perhaps their union representatives, to reduce their propensity to complain about matters the agency wishes to hear less about—thus changing the “retention mechanism.” workers’ attitudes.<sup>16</sup>

### Conceptualize and Simplify the List of Alternatives

The final list of alternatives—the one you include in your presentation to your client and other audiences—will almost certainly look quite different from the one you started with. Not only will you have thrown out some that just don’t look very good, but you will also have done some work to *conceptualize* and *simplify* alternatives. (For a good example of such conceptualization, see Appendix A.)

The key to conceptualization is to try to sum up the basic strategic thrust of an alternative in a simple sentence or even a phrase. This is difficult but usually worth the effort. It helps to use very plain, short phrases stripped of jargon. When the Environmental Protection Agency (EPA) was created, the first administrator confronted (a partial list of) alternatives that might have been described as thus: “Let the states do the work; let the feds give them the money”; “Remove impediments to firms cooperating on antipollution research”; and “Sue the bastards” (meaning the large, visibly polluting firms and industries, the prosecution of which would help build political support for the new agency).

The key to simplification is to distinguish between a basic alternative and its variants. The basic element in many policy alternatives is an intervention strategy—such as regulatory enforcement or a subsidy or a tax incentive—that causes people or institutions to change their conduct in some way.<sup>17</sup> But no intervention strategy can stand alone; it must be implemented by some agency or constellation of agencies (perhaps including nonprofit organizations), and it must have a source of financing. Usually the variants on the basic strategy are defined by different methods of implementation and different methods of financing.

The distinction between a basic strategy and variants based on implementation details is especially helpful when you have a lot of possible solutions to consider and you need to reduce the complexity involved in comparing them. Making the distinction puts you in a position to break your analysis into successive steps. In the first step, you might compare, say, three basic alternatives while ignoring the details described

by their variants. Then, once you have decided on one of these basic alternatives, you could turn to comparing the variants.

For example, you want to decrease the prevalence of heroin use in your county by 50 percent over the next five years.<sup>18</sup> You consider three basic alternatives: methadone maintenance, law enforcement pressure, and drug education. Potential variants for each one have to do with the funding sources in that state, federal, and county money can be used in different degrees (although not all mixes of funds available for one approach are also available for the other two). Variation is also possible according to who administers the program(s): nonprofit organizations, county employees, or state employees. Or you might consider variants of scale and scope, such as two possible sizes for your methadone program.

### Points on a Continuum as “Alternatives”

Suppose you are asked to recommend changes in, say, the rental rates for public housing in your city. Theoretically, each penny change in the rent charged could represent an alternative, but clearly that is a mistaken way to consider “alternatives.” A better approach is to make this into a two-step problem. Step one is to establish the upper and the lower limits of an acceptable range of possibilities, and step two is to choose some point within that range. Choosing each of these limits is a small policy problem in itself, complete with criteria, projections, and the like. For instance, equity might require that the upper limit not be “too high,” meaning somewhere close to \$600 per month, whereas affordability might suggest a slightly lower upper limit. Cost recovery requirements might suggest a lower limit of, say, \$450 per month. In any case, suppose that at the end of step one, the acceptable range has been narrowed to \$475–\$575 per month. One might almost say that a good move for step two is simply to take the midpoint of these two limits, \$525 per month. But there might be additional criteria of interest—for example, finding a “reasonable” increment relative to the current rental rate. If the current rate is \$475 per month, a \$50 increment, to \$525, could be seen as reasonable, but so might a \$75 increment (especially if rents have not been raised in several years), which will permit the city’s housing authority to offer some needed services to residents. At any rate, \$25 increments between \$450 and \$575 seem to be the psychologically “right” set of alternatives—not too large and not too small for the range of options to be considered. Thus, in the end, we have narrowed our alternatives down to six, from an initial array of several thousand.

This two-step procedure could be useful for a variety of problems involving near-continuous variables as alternatives—for example, budget allocations, future dates to begin or to discontinue a service, the number of people to be accommodated by some project or program, emission limits for some effluent, fee or fine schedules, or quantity of water to be released from a reservoir.

The great majority of social science hypotheses about what might work to ameliorate a given problem show up in the language of continuous variables, which then need

to be transformed by the policy analyst into policy-compatible discontinuous choices. If, for instance, studies show that the price elasticity of a pack of cigarettes is  $-0.4$ , that tells you about a continuous relationship (within a certain range) between aggregate cigarettes demanded and the price charged. But if you want to exploit this fact to raise cigarette taxes so as to discourage smoking, you need to translate this information into particular numbers—for example, “Raise the tax \$0.25 per pack to \$1.75.”

## Alternatives Should Be Detailed and Behavioral

A recurring question is how detailed to be in the characterization of “an alternative,” especially in the early stages of one’s work.

The usual answer should be “more detailed,” since there is a natural reluctance to commit oneself to particulars, especially if one is likely to change one’s mind eventually anyway. Detail supports clearer thinking—and also clearer communication with others. The more detail, the less room for talking past one another or for agreeing (disagreeing) when, given the underlying interests and the realistic scope for action, the parties are simply disguising their differences from each other and, probably, from themselves.

Actually, it is a certain kind of detail that is most valuable in characterizing alternatives: “behavioral” detail. Say what you expect people to actually be doing, and especially what you expect them to be doing differently from what is being done now. For instance, “Increase facility inspections from one per year to two or more per year.” Or, “Abolish about half the current safety net features now provided in kind in favor of giving cash.”

Rarely are generic alternatives—“Regulate emissions,” “Subsidize abatement,” “Model smart practice,” and the like—sufficient. These arguments for more detail, particularly more behavioral detail, also apply to Step One: Define the Problem.

## Multistage Analysis

Most of this book focuses on a “one-off” decision (and/or design) process that has a beginning, middle, and end. However, many policy choices are—or should be—part of a process. A single choice is not once-and-for-all but part of a developmental sequence of choices mixed with developments unfolding in the policy environment.

This possibility should show up in one’s thinking about the construction of alternatives. Here are some common process-based alternatives.

There is the classic “wait and see” version. The first decision in this process is to let present trends continue and then to monitor what happens as a result. The second decision comes at some time later, to be made in light of what has happened in the interim.

Another version of multistage analysis is contingency planning: “X or Y will (probably) occur. We will wait until that is resolved. If X, then we should choose A; if Y,

then we should choose B.” For example, the Department of Homeland Security might take certain steps to protect the nation’s ports if a threat of a terrorist attack is made, and then take additional steps if a detailed security assessment finds that the threat is indeed credible.

A third version turns on political feasibility. Political feasibility is not only a condition to be assessed; it is a challenge to be addressed, and addressed in such a way as to help shape a group’s choices about policy. The simple version is to begin to form a political coalition around a problem or objective, and to see what policy options find most favor with the emergent coalition.

Fourth, there is learning by doing. A reasonable policy choice is to start small and easy, make some mistakes along with some successes, learn from both mistakes and successes, and scale up over a few years. Arrangements can even be made for systematic evaluation, either at the beginning or later in the evolution of the policy. Unfortunately, this strategy is better in theory than in practice, since changing political environments and personnel turnover make social learning both hard to do and hard to institutionalize.

## The Last Mile

The most common approach to almost all anticipated problems of bureaucratic pathology, alluded to in “Step One: Define the Problem,” is to put effective safeguards in place—or what planners hope will be effective safeguards. Legislative oversight structures are an example. They can periodically call in senior managers to query them about their agency’s effectiveness, efficiency with taxpayer dollars, and so on.

Other examples—drawn from a nearly endless array—are as follows:

- To hasten bureaucratic action, specify milestones and deadlines.
- To ensure mission fidelity and other goals, require periodic written reports.
- To check on likely quality of agency performance, connect with knowledgeable insiders in the agency who share your goals.
- To prevent waste, fraud, and abuse, require detailed documentation of agency activities.
- To deal with agency abuse of power or administrative incompetence, provide for customer/citizen complaints channeled to special agency oversight units or to members of the legislature.
- To correct and deter a wide variety of abuses or mistakes, provide for customer/citizen lawsuits against administrative agencies.
- To ensure that all interests are represented and heard, specify the frequency and nature of public hearings.

To be more precise, these and other such safeguards are intended to *try* to provide external constructive influence on bureaucracies. They do not always produce their intended effects, however. In “Step Five: Project the Outcomes,” we mention uncertainties and probable weaknesses. And in “Step Six: Confront the Trade-Offs,” we note the costs in agency delay, excessive outlays, administrative burdens on both frontline workers and clients, and program distortions that are often imposed by such bureaucratic safeguards.

## STEP FOUR: SELECT THE CRITERIA

---

It helps to think of any policy story (see Step Eight) as having two interconnected but separable plotlines, the analytic and the evaluative. The first is all about facts and disinterested projections of consequences, whereas the second is all about value judgments. Ideally, all analytically sophisticated and open-minded persons can agree, more or less, on the rights and wrongs in the analytic plotline and on the nature of its residual uncertainties. But this is not true with regard to the evaluative plotline—where we expect subjectivity and social philosophy to have freer play. The analytic plotline will reason about whether X, Y, or Z is likely to happen, but it is in the evaluative plotline that we learn whether we think X or Y or Z is good or bad for the world.

This fourth step in the Eightfold Path belongs primarily, though not exclusively, to the evaluative plotline. It is the most important step for introducing values and philosophy into the policy analysis, because some possible “criteria” are evaluative standards used to judge the goodness of the projected policy outcomes that are associated with each of the alternatives.

Of course, the most important evaluative criterion is whether or not the projected outcome will solve the policy problem to an acceptable degree. But this is only the beginning. After all, any course of action is likely to affect the world in many ways, some desired and some not. Each of those effects—or projected outcomes, to apply our Eightfold Path language—requires a judgment on your part as to whether or not, and why or not, it is thought desirable. Our set of criteria embodies such judgments. Because any significant impact cries out for such a judgment to be made, the greater the variety of significant impacts, the richer will be the set of evaluative criteria needed to deal with them.

**Semantic Tip** Evaluative criteria are not used to judge the *alternatives*, at least not directly. They are to be applied to the projected *outcomes*. It is easy to get confused about this point—and to get the analysis very tangled as a result. This confusion is encouraged by a commonsense way of speaking: “Alternative A looks to be the best; therefore, let’s proceed with it.” But this phrasing ignores a very important step. The complete formulation is “Alternative A will very probably lead to Outcome  $O_A$ , which we judge to be the best of the possible outcomes; therefore, we judge Alternative A to be the best.” Applying criteria to the evaluation of outcomes and not of alternatives makes

it possible to remember that we might like  $O_A$  a great deal even if, because we lack sufficient confidence that A will actually lead to  $O_A$ , we decide not to choose Alternative A after all. With that judgment on the table, it will be possible to look for other alternatives with a greater likelihood of producing  $O_A$ .

## Commonly Used Evaluative Criteria

### Hit the Target!

We sometimes want, or need, to achieve a particular goal by a particular date—for example, cut the rate of state water consumption by 5 percent for the first quarter of next year. Or de-lead all painted interior surfaces in a certain neighborhood by December 31 of this year. Stipulating such concrete targets is often useful for political purposes like mobilizing resources and focusing attention, but it can also be very helpful in framing an analytic agenda. The target might originate in a political mandate, or it might simply be an invented analytic construct. If it is the latter, revisiting the target during the course of the analysis might prove necessary, as the initial version is eventually likely to look too high or too low.

### Efficiency

Typically, the efficiency criterion is the most important evaluative consideration in cost-effectiveness and benefit–cost studies. We use *efficiency*, more or less as the term is used in economics, for maximizing the aggregate of individuals' welfare as that welfare would be construed by the individuals themselves—in economic jargon, “Maximize the sum of individual utilities,” or “Maximize net benefits.” Another roughly equivalent formulation would be “Maximize the public interest.”

Although *efficiency* has an antiseptic, technocratic, and elitist ring to it, the insistence here that “utilities” are to be assessed according to individual citizens' construction of their own welfare is thoroughly democratic. Indeed, siding with efficiency—on average, across most policy issues and policy decisions—is a way to produce more humanistic policy results, too. The reason is not that efficiency is so very humane a concept in itself, but that policy decisions failing to consider efficiency very often fail to take account of the welfare of the little guy at all. The little guy may be little, but in a proper efficiency analysis, he at least shows up to be counted. Efficiency analysis imposes a moral check (for whatever that is worth in the real world of politics) on political visionaries eager to relocate entire populations so as to make room for dams, and on special interests eager to impose seemingly small price increases on large numbers of consumers through protectionist measures in order to maintain the incomes of a relatively small number of producers.

A concern is sometimes raised that a focus on efficiency undercuts efforts to promote social justice. Just how big a limitation this analytic antiegalitarianism turns out to be will depend on particular cases, however. One way this can happen is if analysts

estimate people's "utility" by inferring their willingness to pay for some benefit (or to be spared some deprivation). Individuals with less money do not, in an analytic sense, have as much clout as those with more. A second way that a focus on efficiency can undercut efforts to promote social justice is if the costs and benefits of policies are assessed over an artificially short (e.g., five- or ten-year) time horizon. Some safety-net programs have been found to be highly cost-effective, but only after the long-run benefits to children (including employment, earnings, and mortality) are taken into account.<sup>19</sup>

Second, if the values at stake have few or no human defenders, and therefore no human pocketbooks to back an estimate of willingness to pay, the efficiency criterion may underestimate these values even if by some conception of justice they ought to be weighted heavily. In theory, environmental values are the main example, although in fact some environmental values do have human defenders who derive enormous utility from preserving them—a utility that would be accounted for in a proper efficiency analysis.

Although cost-effectiveness analysis and benefit–cost analysis sound alike and are frequent traveling companions, they are not the same, and their uses can be quite different. True, both construe the policy problem as involving some production relationship between resources and objective(s). And both entail thinking about the relationship by using an economizing lens. However, cost-effectiveness analysis is usually satisfied to assess how well a policy achieves the nature and quantity of the desired outputs,<sup>20</sup> whereas benefit–cost analysis goes a step further and tries to evaluate how much those outputs are valued in terms of money or (rarely) actual utility by individuals. The cost-effectiveness approach is more common in policy analysis than is the benefit–cost approach for two reasons.<sup>21</sup> First, it is less ambitious. Second, cost-effectiveness is often responsive to the immediate needs of clients. Cost-effectiveness analysis examines how well a policy “achieves particular objectives set by the government itself,” whereas benefit–cost analysis “defines value in terms of what individuals desire without any reference to any collective-decision-making process.”<sup>22</sup> In other words, cost-effectiveness analysis examines how policymakers can achieve their *own* goals most efficiently. Indeed, a surprisingly large number of policy issues can be simplified and stylized as cost-effectiveness problems, even though on the surface they may not appear to be likely candidates at all for this sort of treatment. Here are two examples:

- The Mudville mayor wishes to respond to business complaints that building permits “take forever” to obtain. Given that you can spend no more than \$500 and are permitted to change the workflow in the city planning office but not personnel assignments, the cost-effectiveness framework might suggest minimizing delay (measured in days) arising from purely procedural and bureaucratic sources.

- Quake City must upgrade the seismic safety of several thousand buildings constructed of unreinforced masonry. You have a twenty-year time span and no immediate budget constraint, but you wish to accomplish the job with minimum disruption to the lives (and incomes) of the residents and small businesses that may be displaced temporarily by the building renovation process. To minimize such disruption, cost-effectiveness analysis might lead you to propose that the work be done in one season rather than another, or that not all grocery stores be closed at once, or that tenants be assisted in organizing mutual-aid groups. A variant of this is that you have a target deadline (see earlier discussion) and a budget constraint, and you want to find the most cost-effective means of achieving the target while staying within the constraint.

A cost-effectiveness framework typically simplifies policy analysis in another useful way, as well: It assumes as *fixed* either resources or outputs, and focuses only on choices involving the other member of this pair. Fixed resources usually involve a money budget or a human or physical asset such as a work team or a set of hospital beds. A fixed output is generally a target of some kind, such as a minimum required pollution abatement level or a maximum acceptable proportion of children failing an achievement test. Analysis then involves finding the best means to manipulate the other member of the cost-effectiveness pair so as to improve productive efficiency. Colloquially, if resources are fixed, you are “getting the biggest bang for the buck,” or if you have a fixed target, you may be “doing no worse with less.”

Now suppose that, once you have figured out some approach whereby you can do no worse with less, you want to broaden your inquiry to explore whether you can make use of this new and better approach to produce a little (or even a lot) more than you had originally planned. That is, instead of assuming that either resources or outputs are fixed, you are prepared to allow the scale of the activity to increase. The analytic challenge is much more difficult now, because at this point you cannot avoid the question of whether the augmented output “is worth it,” given the envisioned cost increment. That question cannot be answered unless you compare the utilities of *both* the cost increment and the augmented output. That is, cost-effectiveness analysis must now rise to the level of benefit–cost analysis.<sup>23</sup>

Here is an excerpt of a 2005 RAND Corporation benefit–cost analysis—concerning the social return on investing in universal preschool in California. The study reached these conclusions:

- Using our preferred assumptions, a one-year high-quality universal preschool program in California is estimated to generate about \$7,000 in net present value benefits per child for California society (public and private sectors) using a 3 percent discount rate. This equals a return of \$2.62 for every dollar invested, or an annual rate of return of about 10 percent over a sixty-year horizon.

- Assuming a 70 percent participation rate in the universal preschool program, each annual cohort of California children served generates \$2.7 billion in net present value benefits to California society (using a 3 percent discount rate).
- These estimates from our benefit–cost model are sensitive to assumptions about the distribution of benefits that accrue to more- and less-disadvantaged children from participating in a high-quality preschool program. When we consider a range of assumptions from the more conservative to the less conservative (where our baseline results above fall in between), we find that California is estimated to gain at least two dollars for every dollar invested and possibly more than four dollars.<sup>24</sup>

### **Equality, Equity, Fairness, Justice**

There are, of course, a great many different, and often opposed, ideas about what these terms do, or should, mean. In addition to thinking hard about these ideas yourself, sometimes you should also take your audience through some of that thinking, as in the following examples:

- Drivers who do not carry liability insurance leave persons whom they injure in auto accidents at risk of being undercompensated. Many of those who “go bare” are relatively poor. Many other drivers purchase their own insurance against exactly this risk (“uninsured-motorist coverage”). A policy proposal to pay for all drivers’ liability insurance out of a fund created by surcharges at the fuel pump was denounced by some observers as “inequitable” to the poor, who were going bare of insurance. Other observers said that those who go bare impose inequitable premium expenses or risks of undercompensation on the rest of society, including many individuals who are poor or not very well off. Clearly, the analyst needs to include a discussion of the idea of equity.
- The current debate over how to address the gender gap in US education (boys are less likely than girls to graduate high school and earn a college degree) is sometimes said to pit the interests of boys against efforts to promote gender equality. This is odd, though, since some experts—and most ordinary folks, too—argue that no education system claiming to be just could tolerate half of the population falling behind. Again, the analyst has a job to do in sorting out ideas and language.

### **Freedom, Community, and Other Ideas**

To stimulate thought, here is a (far from complete) list of more ideas of possible relevance as evaluative criteria: free markets, economic freedom, capitalism, “freedom from government control,” equality before the law, equality of opportunity, equality of

result, free speech, religious freedom, privacy, safety (especially from chemicals, various environmental hazards, and the like), neighborliness, community, sense of belonging, order, security, absence of fear, traditional family structure, egalitarian family structure, empowerment of workers, maintenance of a viable nonprofit sector, voluntarism, and trust in others.

### Process Values

American democracy values process and procedure—that is, having a say in policy issues that affect you, rationality, openness and accessibility, transparency, fairness, and nonarbitrariness—as well as substance. These considerations probably apply to the very design or decision process for which you are doing your present analytic work. Therefore, remember to consult broadly and equitably. In addition to building up legitimacy for your work, you may be surprised at how much you can learn, especially from people who are very unlike yourself socially or ideologically. This does not, of course, mean that you should in the end accord equal deference to all opinions or desires, or keep the consultative process open forever. Some opinions are more creditable than others, and at some point consultation must give way to decision.

Do not make the mistake of thinking that “more participation” or “greater access to the process” necessarily equates to “more democratic” or “more rational.” Greater opportunities for participation may be exploited more heavily by those with more time to participate or by those with special interests to protect or by ideological zealots. Many citizens are understandably busy with their jobs and families, and may not have the free time or resources to attend meetings or otherwise engage on public policy issues. Their ordinary concerns can come out as relative losers even when the policy-making process is seemingly open and democratic.

### Some Evaluative Criteria Deserve More Weight Than Others

As we saw in the case of defining the problem, when values are at issue—as they are in regard to criterion selection, as well—we must reckon how to weight opposing values. There are three general approaches to this problem.

#### The Political Process Takes Care of It

One approach is simply to allow existing governmental and political processes to determine the weighting. Typically, this approach will accord primacy to the analyst’s employer or client, as well as allowing derivative influence to be exercised by those parties in the relevant arena who are in turn important to the employer or client.

#### The Analyst Imposes a Solution

A second approach is for the analyst herself to modify—though not replace—the weighting assigned by the employer or client by reference to some overarching philosophical or political conception. The justification usually offered for this approach

is that because certain interests, and perhaps philosophies, are typically “underrepresented” in government and politics, and because the analyst is in a better position than most other participants in the process to see or understand or appreciate this problem of underrepresentation, the analyst is duty-bound, or at least permitted, in the name of fairness and democracy, to right the balance.

For instance, some observers would argue that were it not for policy analysts, efficiency-related criteria would rarely be heeded and that, as a consequence, analysts should in effect speak up for the taxpayers whose interests may be squeezed out by better-organized advocacy groups. A related argument is sometimes made that certain conceptions of equity—in particular those having to do with the idea that the beneficiaries of publicly provided goods or services should pay for them—are underrepresented except among policy analysts. (These conceptions of equity typically exclude public expenditures deliberately intended to redistribute wealth among citizens.) Other interests that people sometimes claim are underrepresented and therefore need representation by analysts are future generations, children, people who live outside the jurisdiction making the decisions, ethnic and racial minorities, women, the LGBTQ community, the poor, consumers, and animals and plants (ecological entities).

A variant of this approach introduces the idea of an educational process. Depending on circumstances, the analyst might encourage influential political actors—perhaps including the analyst’s boss or principal client—to rethink their existing criteria in the light of facts or arguments the analyst can draw to their attention. In this case, the analyst takes responsibility for opening up a dialogue, and perhaps for trying to infuse it with reason and insight, but then allows the political process to take over.

The analyst can help the dialogue along by making sure that the assigning of weights will be done in the context of confronting the trade-offs, and framed as crisply and clearly as possible. (See “Step Six: Confront the Trade-Offs.”) For instance, “We project that a choice of Route A for the railroad will impose \$20 million more in construction costs than a choice of Route B. Route A is five miles longer than Route B (which is about 2 percent); however, Route A will permit us to spare the homes of about thirty families, half of whom have lived in the area for at least ten years. We are offering a \$300,000-per-family compensation if they have to move, and twenty of the families seem willing to do that. But that still leaves ten families who might be unhappy with this solution. So a lot of your decision here probably rests on just how much weight you give to imposing on these ten families versus saving the \$20 million in construction costs.”

### **The Distribution of “Rights” Precludes Some Solutions and Forwards Others**

If X has a recognized property “right,” you can’t easily override it just because your policy solution would find that convenient; and if Y has a “right” to privacy, you might be inclined to tilt the weighting of criteria heavily in that direction. Generally, claims based on rights are a reasonable guide to choosing “better” policies, and rights-based

criteria deserve some extra weighting. However, plenty of exceptions exist, and it pays to examine, briefly, the whole matter of where rights come from and how policy analysis can make good use of them. This is a controversial matter, of course, and our thoughts on it are certainly contestable.

Typically, rights are specially protected claims of an individual or a group against encroachment by “others,” including society as a whole, though in some cases it is society that claims the rights against component groups or individuals. Sometimes rights are long-standing, well established, consensual, and, within our social context, unquestioned, as in “X has a right to be treated with dignity, irrespective of X’s economic condition.” In these cases, the pattern of rights claims, hedged and limited though they might be, very likely is found to be a good self-help tool for organizing the many and varied interactions in a complex society. But rights are sometimes more emergent than established, and claims based on rights can be quite contentious or in conflict: “I have a right to use my cell phone in any place, private or public,” versus “I have a right not to be disturbed by your loud and obnoxious cell phone conversation, thank you very much.”

It is best to think of *all* rights claims as emerging from a social process of trial and error and contestation, with the ones that seem obviously legitimate to most people being merely the (so far) best established and (probably) most socially beneficial. Claims that particular rights are justified by nature or “divine will” or reason or “our common humanity” are simply rhetoric, because these justifications are always challenged by others in heterogeneous societies. Over the centuries through which these debates have continued, no permanent resolution has occurred, and we think one will never occur, since rights are simply convenient tools of social organization and rights-based claims a consensually accepted way of negotiating the changing landscape of whose interests should be protected to what degree and with what exceptions. The recognition and status of particular rights in both the law and society can change over time. New rights can be recognized (e.g., the right of same-sex couples to marry) and existing rights can sometimes be withdrawn or scaled back (e.g., abortion rights).

From the point of view of a hypothetical social engineer trying to improve social welfare, some rights should certainly be treated as relatively fundamental. If the moral realm were the legal realm, these rights would be considered constitutional. But, like the Constitution, even fundamental rights, such as “the right to free speech,” should evolve to fit new social and technological conditions. Technological change raises questions of privacy and transparency (e.g., confidentiality and fairness), and the past structure of rights is not necessarily a good guide to how to redesign that structure for the emergent situation. A fortiori, this applies to matters of lesser moment—where it is easier to see that rights are constructed rather than found. Forcing others to listen to your cell phone conversation may or may not be a right we wish to create and honor, but it is surely novel, and it needs to be settled by reference not to reason and the like but rather to the balancing of utilities in a strictly pragmatic fashion. The same applies to compensation for takings, decent health care, privacy, abortion, and a host of other

matters that now or in the recent past have been subject to debate over who ought to have what sort of highly protected positions that we dignify and crystallize as rights.

Please do not misunderstand this as an argument in favor of relativism, which in many people's usage is the same as saying that there is no choosing among different rights claims, that one is as good as another. That is not true. Certainly, allowing people to claim a high level of protection for (i.e., a right to) certain values—such as individual autonomy—is beneficial to the running of a modern democratic society. But this right sometimes needs to have exceptions carved out of it to accommodate cases when the exercise of this right imposes excessively on other people. The fine texture of the fabric of such rights is always subject to discussion, and the basis for making these decisions is to be found not in rhetoric or in philosophical speculation but in the analysis of alternative fabrics, each taken as a whole and including all the internal tensions that are bound to be included in them. The evolution of rights in the moral realm—that is, in the realm of private practice and thence public opinion formation—involves the sort of constant tinkering and adjustment we see in the realm of both statutory and judge-made law.

In the end, therefore, claims to weight criteria by reference to which rights ought or ought not to take priority deserve to be treated critically.

### Commonly Used Practical Criteria

Not all criteria that come into play in an analysis are part of the evaluative plotline. Some are purely practical and are part of the analytic plotline. These criteria have to do with what happens to an alternative as it moves through the policy adoption and policy implementation processes.<sup>25</sup> The main ones are legality, political acceptability, administrative robustness and improbability, and policy sustainability.

#### Legality

A feasible policy must not violate constitutional, statutory, or common law rights. Remember, however, that legal rights are constantly changing and are often ambiguous. It is sometimes worth taking a gamble on a policy that might—or might not—be adjudged illegal when tested in court. (In such cases, advice of counsel is clearly in order to help craft the policy so that its survival chances are enhanced.)

**Semantic Tip** As noted earlier, however, remember that rights alleged to be “natural” or “human” are conceptually quite different from legal rights, despite the semantic similarity. Examples are the conflicting abortion stances predicated on right-to-life values or a woman's right to control her own body. Alleged natural or human rights are sometimes controversial in that some people would like to have them recognized as legal prescriptions whereas others would oppose such recognition.

#### Political Acceptability

A feasible policy must be politically acceptable, or at least not unacceptable. Political unacceptability is a combination of two conditions: too much opposition (which may

be wide or intense or both) and/or too little support (which may be insufficiently broad or insufficiently intense or both) for the proposal to win adoption.

A stakeholder analysis can help gauge political acceptability. First, using web searches, interviews, and other research methods, identify the relevant actors—including elected and appointed officials, businesses, trade associations, professional societies, advocacy groups, and ideological organizations—that might plausibly take a stand (pro or con) on the proposal. Keep in mind that actors are more likely to become active if they have been involved with an issue in the past, if they believe a proposal impinges on their interests or ideological goals, and if they perceive that they will bear losses under the proposal (people tend to react more strongly to losses than to gains of equal size).

Second, list the resources possessed by each actor, such as authority, expertise, financial resources, and the ability to mobilize or speak on behalf of others.

Finally, identify the institutional venues in which decisions will be considered, the rules and procedures by which each such venue operates, and the type of claims each venue permits to be heard (e.g., courts require parties to have “legal standing” and to express their preferences in terms of duties and rights).

Do not take a static view of political unacceptability, however. Always ask yourself the question, “If my favorite policy solution doesn’t look acceptable under current conditions, what would it take to change those conditions?”

You may discover that creative political strategizing can change the set of relevant stakeholders, modify their respective preferences or resources, or shift the institutional context in which policies will be made, thereby opening up options that haven’t been seriously considered before.

In assessing strategic limitations and possibilities, it will help to make use of various models of the political process. As we observed earlier, models are based on metaphors, and the ones that are likely to be most valuable in this case are these:

- A game in which strategic actors (both individuals and groups) seek to maximize their payoffs without cheating, given both the rules of the game and the strategic behavior of the other players<sup>26</sup>
- A war, in which partisan or ideological armies seek to defeat and demoralize their political enemies, preserve past victories, and conquer new policy territory
- A theater, in which the actors are elected officials who strive, with or without a basis in reality, to create a good appearance—to themselves, to each other, to the critics, and to the audience (whose approval, ultimately, is all-important)
- A marketplace of slogans, symbols, and ideas, with a mix of honorable merchants and hucksters as sellers and a mix of sophisticates and innocents as buyers
- A school in which elected officials learn how to do good policy design work and sometimes share their results and their methods with their classmates

How exactly is one to make use of such models? Think of them as conceptual lenses. Observe the relevant political process through each of them in turn, and identify the probable pitfalls and opportunities brought into focus by each.<sup>27</sup>

A common obstacle to the adoption of policies that would generate net benefits for society is that the changes will impose concentrated costs on the interest groups who profit from current arrangements. These clienteles will inevitably lobby against the proposals, and they may be better organized than the many people who would gain from the changes. By modifying the incidence of costs and benefits, however, it may be possible to boost the acceptability of a policy option without unduly blunting its effectiveness. Those who would bear losses under the policy might be given direct or indirect transitional assistance, for example. The use of “grandfather” clauses and “phase-ins” may also improve the odds of policy adoption.

Another common obstacle to the adoption of socially beneficial policies in today’s era of polarization and tight partisan competition is that supporting the policies may impose political costs on elected officials. If representatives believe they hold an electoral advantage on an issue, they may prefer stalemate to action. Indeed, elected officials might even *oppose* the adoption of a policy they support on substantive grounds to preserve their party brand. One way to counter the incentives for “strategic disagreement” is to create a “penalty default” (which all relevant actors wish to avoid) that automatically goes into effect if policymakers do not negotiate an agreement. Penalty defaults are already employed in some programs. For example, the benefits of Social Security recipients will be cut by 21 percent in 2034 (under current forecasts of the system’s insolvency) unless policymakers take steps to stabilize the program’s finances before then. As this deadline approaches, the political rewards for compromise will begin to exceed the political costs, making the adoption of a policy fix more likely.<sup>28</sup>

### **Administrative Robustness and Improvability**

Policy ideas that sound great in theory often fail under conditions of field implementation. The implementation process has a life of its own. It is acted out through large and inflexible administrative systems and is distorted by bureaucratic interests. It is also mediated by the incentives, preferences, and capacities of program “targets,” such as low-income mothers who are required to participate in a “welfare-to-work” program.<sup>29</sup> Policies that emerge in practice can diverge, even substantially, from policies as designed and adopted. A policy alternative, therefore, should be robust enough that even if the implementation process does not go very smoothly, the policy outcomes will still prove to be satisfactory.

Some adverse implementation outcomes usually worth worrying about are long delays; excessive budgetary or administrative costs; scandal from fraud, waste, and abuse that embarrasses supporters; and administrative complexities that leave citizens (and program managers) uncertain as to what benefits are available or what or how regulations must be complied with.

Even the best policy planners cannot get all the details right at the design stage. They should therefore allow room for policy implementers to improve on the original design. The most common vehicle for such improvement is participation in the implementation process by individuals and groups whose expertise or point of view was not included in the design phase. However, the openness that makes for improvability can also, by opening the door to hostile political interests, diminish robustness. Hence, a careful evaluation of the current factual situation—personalities, institutional demands and incentives, political vulnerabilities, and so on—is usually in order.

In estimating robustness and improvability, models of bureaucracy can serve as useful conceptual lenses, as suggested earlier with regard to carrying out political analysis. We find the most useful metaphors for bureaucracy to be these, listed in no particular order:

- An automaton enacting preprogrammed routines (“standard operating procedures,” or “SOPs”)
- A person in an environment, driven by survival needs, self-enhancement interests, and, under some conditions, a desire for self-actualization
- A political arena wherein individuals and factions jockey for influence over the organization’s mission, access to its decision systems, and its prerequisites
- A tribe with its own rituals and an array of safeguards against contamination by “outsiders”
- A society of individuals cooperating toward a more-or-less common set of goals
- A structure of roles and interrelationships that are intended to complement one another in a rational division of labor
- An instrument used by “society” for society’s own objectives

### Policy Sustainability

Policies typically must endure for a period of time to achieve their desired impacts, but the elected officials who voted for the policies will not remain in office forever, and their successors may have different agendas. Moreover, the groups who opposed a policy’s adoption may return to fight another day. Rather than a one-shot affair, policymaking is a *dynamic* process in which the consolidation of a new policy may be more challenging than winning its adoption in the first place. Sustainability refers to the capacity of a policy to outlast its enacting coalition, maintain its integrity, and deploy core principles to stave off unwarranted political pressures for debilitating changes.

Policies “stick” not simply because they produce net social benefits but also because they generate “enough” support from key constituencies over time. The most durable policies create “positive feedback” by encouraging citizens, businesses, and groups to

adjust to the new reality. Once that happens, such actors become reluctant to have the policy repealed or fundamentally changed, because they want to protect their investments in the new ways of doing things. Social Security's history illustrates these dynamics. When Social Security was created in 1935, older Americans were the least active age group in politics. As Social Security grew, it built a constituency among older citizens by "(a) giving them the resources of money and free time; (b) enhancing their levels of political interest and efficacy by tying their well-being visibly to a government program; and (c) creating incentives for interest groups to mobilize them by creating a political identity based on program reciprocity."<sup>30</sup> Durable policies not only build supportive clienteles and shape social identities. They also divide opponents in ways that make policy reversal more difficult.

Some sustainability risks to be concerned about include initial passage by a thin or temporary majority; lack of bipartisan support; benefit flows that are too small, too delayed, too uncertain, too invisible, or too stigmatizing to mobilize supporters and encourage self-reinforcing adaptations; capture of program benefits by a relatively underserving or unintended group; and lack of credibility of the government's commitment to the policy.<sup>31</sup>

### "Criteria" as Logical Constructs

Criteria such as efficiency, equity, political acceptability, and robustness are substantive. But we can think of criteria of a purely formal sort, as well. For instance, we can distinguish among criterion values that we wish to maximize, those that must be minimally satisfied, and those of a generally lesser priority for which "more is better."

It is helpful to focus initially on one primary criterion, a *principal objective to be maximized* (or minimized). Typically, this principal objective will be the obverse of your problem definition. For instance, if your problem is that too many families are homeless, then your principal objective will probably be to minimize the number of homeless families. If the problem is that global temperatures are rising too rapidly, a good statement of a principal objective might be "Minimize or reverse the increase of global temperatures." Naturally, there are other criteria to judge outcomes by, such as costliness, political acceptability, and economic justice, and these should all enter into the final evaluation. However, unless you focus—initially, at least—on a single primary criterion and array others around it, you will likely find yourself getting very confused. As you get deeper into the analysis and feel more comfortable with a multiplicity of important objectives, you may wish to drop your emphasis on a primary criterion and work on a more complex "objective function," in the language of mathematical programming. See Box I-5 for an example of selecting the criteria from a real-world report.

### Linear Programming

A mathematical (and now computer-accessible) technique for optimizing choice when you have a principal objective or an objective function and a scarce stock of

resources for maximizing it is called “linear programming.”<sup>32</sup> Often, at least some of the resources—such as the agency budget and the available physical facilities promised by a nonprofit agency—are constrained. Even if the problem is not subject to simple quantitative assessment, analysts often find it useful to take advantage of the logical structure of linear programming to conceptualize their task. The conventional formulation then sounds like this: “Maximize this objective (or objective function) subject to such-and-such resource constraints.”

## BOX I-5: AN ILLUSTRATIVE EXAMPLE OF “SELECTING THE CRITERIA” FROM A POLICY ANALYSIS REPORT

*Approaches to funding highways can be evaluated in terms of equity and economic efficiency. Equity, or fairness, is subjective and can be assessed in several ways. Observers of highway funding often gauge fairness by considering the share of funding that is obtained from taxes paid by highway users rather than from general taxpayer funds, from people in households that fall into various income categories, or from people in rural versus urban households.*

*The economic efficiency of a funding approach depends partly on its effects on users’ travel behavior and partly on what it costs to implement. Charging users for the costs their travel imposes on society would create incentives for people to limit highway use to trips for which the benefits exceed the costs, thus reducing or eliminating overuse of highways and helping identify the economic value of investments in highways. However, the costs of collecting and enforcing such user charges also must be considered in evaluating their net effect on efficiency. . . .*

### Equity

*The equity implications of fuel taxes, the primary current source of HTF [Highway Trust Fund] revenues, are mixed: Fuel taxes satisfy the user-pays criterion, but they tend to be regressive; that is, they impose a larger relative burden on low-income than on high-income households. An analysis of 2004 data on effective tax rates (taxes paid divided by income) that divided all households into five groups of equal size by income showed that people whose households were in the second-lowest and middle quintiles paid somewhat larger shares of their income in gasoline taxes than did people in the lowest quintile or in the top two quintiles. Fuel taxes are less directly burdensome for households in the bottom group of earners, in part because people in some of those households do not own automobiles. However, the diesel fuel tax also imposes an indirect burden through the effect on the prices of shipped goods. Because lower-income households consume larger shares of their income, that indirect effect would add to the overall regressivity of the fuel taxes considered together.*

*Some observers find another equity concern in the fact that fuel taxes disproportionately affect people who live in rural areas. According to data from the Department*

of Transportation's National Household Travel Survey, people in rural households spend more, on average, on gasoline or diesel fuel because their vehicles (including light-duty trucks and older cars) tend to be less fuel efficient than are the vehicles of their urban counterparts and because people in rural areas tend to drive more. The survey data indicate that rural households at all income levels spend more on gasoline and diesel fuel than is spent by comparable urban or suburban households. For example, rural households with income below \$25,000 spent 30 percent more than did their urban counterparts, in part because they drove 13 percent more miles. Relative differences in spending on fuel between rural and urban households were even greater among other income groups.

## Efficiency

In terms of efficiency, two aspects of fuel taxes are positive: First, the costs of collection and enforcement are low, in part because fuel taxes are not collected directly from individual service stations or from users of fuel but from fuel distributors, which collect them from the service stations where the money is collected from fuel purchasers. (In 2008, there were 114,000 filling stations and about 8,000 distributors in the United States.) Second, in combination with state and local fuel taxes, the federal taxes give motorists an incentive to reduce fuel consumption, thereby reducing the external costs associated with that consumption and, to some extent, the costs related to mileage.

---

Source: *Alternative Approaches to Funding Highways* (Washington, DC: Congressional Budget Office, March 2011), <https://www.cbo.gov/sites/default/files/112th-congress-2011-2012/reports/03-23-highwayfunding.pdf>.

Here is an example from the homelessness problem: “Maximize the number of homeless individuals housed on any given night, subject to the constraints of not exceeding \$50,000-per-night total budgetary cost to Agency X, not putting shelters into Neighborhoods A and B for political reasons, and trying to give ‘more’ choice to the beneficiary population as to where they will take shelter.”

**Semantic Tip** If it is possible to sort your criteria according to whether they refer to values to be maximized or minimized, values that stand as constraints, or values that have a more-is-better or less-is-better quality, keep the different statuses of the criteria in mind. Be conscious of them. You can do this with a simple verbal trick: As appropriate, define your criteria as “maximize such-and-such value,” “satisfy such-and-such value constraint,” or “minimize such-and-such value.” For example, minimize tons of carbon dioxide (CO<sub>2</sub>) released; or maximize lives saved per dollar spent. If a criterion label contains no signal as to the better direction to move in, as in “governance structure” or “effect on landlords,” it is almost certainly insufficient.

In any case, to the extent possible, the criteria should be characterized both in conceptual and in operational (typically quantitative) terms. Conceptually, for instance, one talks about “maximizing the reduction of greenhouse gas emissions from publicly

owned buildings,” whereas operationally, one talks about “minimizing the tons of greenhouse gas emissions per month from publicly owned buildings.” In this case, the operational definition is a close proxy for the more qualitative conceptual definition. Frequently, however, something of a gap exists, since what is measurable may only imperfectly reflect the conceptual characterization. For instance, minimizing “the hassle factor” to the citizen in recycling his garbage is conceptually meaningful but hard to express quantitatively. It is really about the psychology of effort, the degree of belief in the desirability of the goal, and the degree of frustration involved in preparing one’s garbage for pickup. In this case, the best you could probably manage operationally would be to estimate the number of minutes the citizen spends per week to cooperate in the enterprise.

To the extent possible, group your criteria in such a way that all the “positive” (benefit) criteria are clustered separately from the “negative” (cost) criteria. In a logical sense, how one does this does not really matter. But it makes for easier reading and discussion. It is a little like arranging your bridge hand by suit and, within suits, by number sequence.

Don’t embrace euphemisms or other dodges as a substitute for words that describe harsh realities. The client for one student project asked for advice on what adult school programs to cut in order to save money in financially desperate circumstances. The students initially put together a brief defending adult school programming in general, at best leaving to inference what elements of the bundle were most deserving of cuts.

## Specify Metrics

Clarity about criteria is greatly helped by specifying metrics. Table I-1 illustrates this point. Qualitative framing of the criterion is stated in the left-hand column, and it is fairly clear. But aiming to specify the metrics that might give it additional meaning helps even more. In some cases, the addition of a metric also adds insight into what one really wants.

Note that the objective here is not to quantify but to clarify. Quantification is desirable when possible, and identifying a metric is a helpful step in doing so. But the real purpose here is to help your thinking.

Nor is the purpose at this stage to canvass data sources or figure out how data are to be collected and reported. That is important, but it is separate from trying to improve your thinking. In fact, for some metrics, you cannot imagine collecting data at all. In Table I-1, although you can imagine how you might collect data about burglary rates (row 3), collecting data about cows doing their maternal duty (row 2) seems a lot harder. And births prevented (row 1) is logically impossible, since the needed counterfactual does not exist. Indeed, this is true of all prevention programs; counterfactuals can at best be estimated, and only by looking at proxy measures and at the results of experiments with well-constructed control groups.

**TABLE I-1 ■ An Example of Metric Specification**

Qualitative Description	Description with Quantitative Metrics
<ul style="list-style-type: none"> <li>● Maximize postpartum accessibility to family planning</li> </ul>	<ul style="list-style-type: none"> <li>● Maximize number of prevented unwanted and unplanned pregnancies within the first x months after birth</li> </ul>
<ul style="list-style-type: none"> <li>● Maximize humane treatment of dairy cattle</li> </ul>	<ul style="list-style-type: none"> <li>● Maximize % of cows allowed to tend and raise offspring for at least x months after birth</li> </ul>
<ul style="list-style-type: none"> <li>● Minimize North Brookwood crime</li> </ul>	<ul style="list-style-type: none"> <li>● Bring burglary rate in North Brookwood down by 10% in next 12 months</li> </ul>
<ul style="list-style-type: none"> <li>● Minimize post-inspection carbon monoxide (CO) emissions from autos</li> </ul>	<ul style="list-style-type: none"> <li>● Minimize average tons of CO emissions from autos in 12 months post-inspection</li> </ul>
<ul style="list-style-type: none"> <li>● Satisfy political feasibility requirement</li> </ul>	<ul style="list-style-type: none"> <li>● Estimate odds &gt; 75% that governor will sign this executive order</li> </ul>
<ul style="list-style-type: none"> <li>● Lower implementation hassle</li> </ul>	<ul style="list-style-type: none"> <li>● Spend fewer hours persuading reluctant implementers</li> </ul>

## Avoid Confusing Alternatives and Criteria

**Semantic Tip** *Alternatives* are courses of action, whereas *criteria* are mental standards for evaluating the results of action. How could you ever mistake an alternative for a criterion, or vice versa? As with many instances of confusion in policy analysis, the source of such a mistake is likely to be semantic. Consider, for example, a senior manager in a state regulatory agency dealing with worker safety. She wishes to incorporate worker complaints into the agency’s strategy for targeting inspections across work sites in the state. Her assistant presents her with a number of alternatives for doing so, one of which is called “rapid-response (twenty-four-hour maximum) hotline.” Not surprisingly, one of the criteria for assessing outcomes is “responsiveness.” The alternative therefore seems a lot like the criterion. But this is an illusion. The alternative (course of action) is really the hotline. The main reason it looks like a criterion is that the *intention* of rapid response has crept into the definition of the alternative.<sup>33</sup> This is a dangerous mistake, because one should not assume through definition that an intention, as expressed in the verbal characterization of the alternative, will actually be realized.

This sort of confusion is most likely to arise when the internal activities of an organization are under discussion, since proposals to create or modify organizational units resonate with intentionality. Consider a proposal to create a performance measurement office, a strategic planning team, and a customer service department. The performance measurement office may end up, for whatever reasons, using meaningless measures collected by unreliable agents; the strategic planning team may be deliberately ignored

by savvy or possibly unsavvy managers; and the customer service department may unintentionally end up as an instrument of customer alienation. We once questioned a student's proposal to create a "drug counseling service" for employees within an organization. The proposal seemed too weak to make a dent in the organization's problem. The student countered, "No, I'm talking about not just any old counseling service that might attack this problem, but an 'effective' one." Nothing in the student's account of how the service was to work increased the odds that it might really be effective. Effectiveness was assumed simply because the student wished to assume it.

### Criteria for the Last Mile

Conceptually, criteria to help choose among alternatives for the Last Mile are simple to state but hard to detail. In "Step Three: Construct the Alternatives," we focused on safeguards. Two main criteria apply to these, effectiveness and lightness of touch. The latter is important to avoid impeding effective and efficient action, such as with endless hearings, writing of unread reports, filling out of forms with data that will never be used, or diverting program action to activities that count for bureaucratic purposes but serve customers poorly or even cause harm. Because of the danger of impeding action with a piling on of formal safeguards, consider informal safeguards like trust and jawboning as alternative, if perhaps less effective, safeguards.

Avoiding the danger of too heavy a touch is not usually a criterion in the forefront of policy designers' minds, but it ought to be.

## STEP FIVE: PROJECT THE OUTCOMES

---

For each of the alternatives on your current list, project all the outcomes (or impacts) that you or other interested parties might reasonably care about. This is the hardest step in the Eightfold Path. Even veteran policy analysts do not usually do it very well. Not surprisingly, analysts often duck it entirely, disguising their omission by a variety of subterfuges. Hence, the most important advice about this step is simple: do it.

At least three great practical as well as psychological difficulties must be confronted here. First, "policy" is about the future, not about the past or the present, but we can never be certain about how the future will unfold, even if we engage it with the best of intentions and the most thoughtful of policy designs.

Second, "project the outcomes" is another way of saying, "be realistic." Yet realism is often uncomfortable. Most people prefer optimism. Policy can affect people's lives, fortunes, and sacred honor, for better or for worse. Making policy, therefore, imposes a moral burden heavier than many people care to acknowledge. Understandably, we would rather believe that our preferred or recommended policy alternative will accomplish what we hope and that it will impose fewer costs than we might realistically fear.

Third, there is what is sometimes called “the 51–49 principle.” That is, in the thick of the policy fray, we are driven out of pure self-defense to treat 51 percent confidence in our projection as though it deserved 100 percent confidence, so that we sometimes mislead not only others but ourselves as well. The first difficulty—namely, that we can never have wholly convincing evidence about the future—compounds the second and third, inasmuch as our wishful thinking is not readily disciplined by reference to empirical demonstrations and proofs.

These psychological difficulties notwithstanding, systematic efforts to project outcomes are essential. For policymakers in a modern democracy, neither following gut instincts nor reading pigeon entrails is a responsible alternative.

### Extend the Logic of Common Sense

In this section we discuss, in a very general way, the logic of combining models and evidence to produce usable projections of policy outcomes for the various alternatives being considered. The logic is largely that of common sense supported by social science methodology but with some important additions and subtractions.

First, policy analysis uses social science to the degree that it can. A great deal of social science is directed toward answering the question, “Is Model X of this piece of the world realistic?” Social scientific studies of this type can often be useful for diagnosing the existence of problems, mapping trends, and deciding whether some seemingly “smart” practice (see Part IV) is worth trying to replicate. You should be careful, however, to avoid using the social scientific standard of adequacy for judgments about the realism of a model, for it is quite conservative. In policy analysis, the looser, but more appropriate, standard should be whether reliance on a model can lead to better results and avoid worse results than less disciplined guesswork.

Second, policy analysis, as we have seen, uses multiple models. Most social science, in imitation of the hard sciences, looks for the “best” model (or, in the case of some practitioners, the “true” model). Because all models abstract from reality, however, even the best models are never complete. Although such abstraction may advance the progress of science, in the world of policy, where real consequences of policy choices are to be experienced by real people, no facet of a problem or the possible alternatives to be adopted can be exempted from analysis. Whatever models *can* be employed to illuminate some important facet of the problem or of the possible outcomes *should* be employed—even if doing so results in an inelegant and ad hoc multiplication of subanalyses.

Third, even when you have adopted adequately realistic models of sufficient number and variety, these models still need to be used in conjunction with evidence about “initial conditions,” or the facts on the ground as they currently exist. For instance, “Deputy Director Smith is as incompetent as they come. The need to work around the deputy director will raise the risks of failure by at least 25 percent.” Or, “The

community appears very angry about the drug scene right now, and residents will almost certainly help the police in the planned crackdown.” Although the projections of many models are not particularly sensitive to initial conditions, some are. These are the models that bear on projections of political acceptability and on the robustness of an alternative to the stresses of the implementation process.

Finally, policy analysis, as we have seen, makes use of the metaphors behind the models—metaphors such as “bureaucracy as automaton” and “politics as theater” and “this piece of the world as production system”—to yield qualitative insights about important causal relationships. The especially important relationships are those that may afford useful intervention points in complex systems or that present potential pitfalls in the policy adoption or implementation process.

### Choose a Base Case

For the next step, “Confront the Trade-Offs,” we counsel comparing the projected outcomes—the work of this present step—so that you can see clearly in what ways the various pairwise comparisons for which trade-offs exist differ from each other. This step, then, prepares the raw material for that next step. (See also the discussion of “setup for the next step,” page 68.)

To do that, your projections should all be defined against a common reference mark, the *base case*. If the base case is whatever condition exists today, and that condition is not expected to change, then each outcome should be described in terms of the difference between what would (probably) exist tomorrow and what (arguably) exists today. For example, if poverty in Rivertown is 15 percent today, and Alternative A is expected to decrease it by, say, 2 percent, then the projected outcome is –2 percent or, in absolute numbers, 1,000 fewer individuals in poverty. If the comparable projection for Alternative B is –3 percent, or 1,500 fewer individuals in poverty, then it will be easy to see, when you come to confronting trade-offs, that B is better than A on this dimension by 500 fewer individuals in poverty. (It may also be more costly or less desirable in other ways, but those considerations can wait.)

If the base case is whatever condition exists today, and that condition *is* expected to change, then a comprehensive investigation could potentially be done of how the world will evolve in the absence of the adoption of each alternative under a particular condition. This may require an investigation of a wide range of potential factors that are expected to influence conditions in the future. For example, the Biden administration proposed that in developing a base case, regulators consider factors such as “technological advances, demographic changes, changes in the economy, or alterations to the climate.”<sup>34</sup>

In sum, if a problem is expected to become more (less) severe over time, the base case should reflect that forecast.

What is a good base case? Between the polar extremes of “whatever conditions exist today” and “how (multiple) present trends would unfold without the policy under

consideration” are many other possibilities. Here is a list of some of them, along with a brief commentary:

- *Future conditions provided that business were to continue as usual.* The analysts whose conclusions on greenhouse gas abatement appear in Table I-2 (see p. 53-54) chose this as their base case. They assumed no new regulations and no changes in fossil fuel consumption other than those caused by demographic changes. They did not include possible technological changes, for instance.
- *Changes from the present that would occur if some policy were to be adopted.* Suppose, for instance, that the state was likely to finance and construct a train system connecting major cities, and that this system was expected to reduce automobile usage overall by, say, 5 percent. This is like “business as usual” except that changes caused by a particular policy are in sharper focus.
- *Projections of the results of one particular policy option.* In 1996 the RAND Corporation, referred to earlier, published a study comparing the cost-effectiveness of crime reduction strategy programs to the base case of the “three-strikes” mandatory incarceration policy that California had recently adopted. Three out of the four programs were clearly more cost-effective, and the fourth possibly so. The objective was to show that on narrow crime-prevention grounds alone, and leaving aside humanistic considerations, three-strikes was wasting taxpayer money, since other options were cheaper for achieving the same objective.

**TABLE I-2 ■ Comparative Analysis: Anytown, USA (2050 baseline: 5.5 million metric tons CO<sub>2</sub>e)**

		{% reduction from 2050 CO <sub>2</sub> e baseline}	{Cost per ton CO <sub>2</sub> e abated}	O (Operational) E (Economic) P (Political)
	Policy Scenario	Efficacy	Cost-Effectiveness	Viability
EXISTING BUILDINGS	Mandate efficiency retrofits for homes	6.9% to 8.8%	-\$130 to \$5	O: High E: Medium P: High
	Mandate efficiency retrofits for commercial buildings	7.9% to 10.5%	-\$132 to -\$30	O: High E: Medium P: High

(Continued)

**TABLE I-2 ■ Comparative Analysis: Anytown, USA (2050 baseline: 5.5 million metric tons CO<sub>2</sub>e) (Continued)**

		{% reduction from 2050 CO <sub>2</sub> e baseline)	(Cost per ton CO <sub>2</sub> e abated)	O (Operational) E (Economic) P (Political)
	Policy Scenario	Efficacy	Cost-Effectiveness	Viability
NEW BUILDINGS	Require zero-energy capable homes	4.1% to 5.6%	–\$132 to –\$25	O: High E: High P: High
	Require zero-energy capable commercial buildings	6.5% to 8.9%	–\$120 to –\$48	O: High E: High P: High
URBAN PLANNING	High-density residential development	2.4%	–\$1,333 to –\$702	O: High E: High P: Medium
ENERGY SUPPLY	Incentives for distributed PV	3.9%	\$15 to \$139	O: High E: Medium P: High
FINANCIAL MECHANISMS	\$20 carbon tax	11.3%	\$20	O: High E: Medium P: Low
	\$50 carbon tax	20.6%	\$50	O: High E: Medium P: Low

Source: Phillip Cryan, Nina Horne, Jessica Shipley, and Benjamin Thornley, *Building a Path to a Greener Future: City Climate Change Policies for 2050* (Berkeley: Goldman School of Public Policy, University of California, 2008).

CO<sub>2</sub>e = carbon dioxide emissions; PV = photovoltaic

It is worth noting that if the base case contains errors (from projecting the future or from misunderstanding the present), these will not matter if the errors do not affect the comparisons of the projected outcomes differently. In the earlier example, if the poverty rate in Rivertown is actually 16 percent rather than 15 percent, this mistake makes no difference if the absolute numbers of individuals projected to be helped are still 1,000 and 1,500.

## Dare to Make Magnitude Estimates

Projecting outcomes often requires you to think not just about the general direction of an outcome but about its magnitude, as well. Typically it's not enough to say, "We expect this program to have a very positive effect on reducing unwanted teenage pregnancies." Instead, you would want to say, "We expect this program to reduce by one hundred to three hundred the number of unwanted teenage pregnancies per year in this community over the next five years." Developing magnitude estimates can help reduce the likelihood that an analysis will be misinterpreted. (If no numerical estimate is provided, policymakers may mistakenly assume the projected impact is zero. And if the projection really is zero, find a way to make that clear.)

Here is an example of a Congressional Budget Office (CBO) magnitude estimate of the impact of a bill (the Nutrition Reform and Work Opportunity Act of 2013) on participation in the Supplemental Nutrition Assistance Program (SNAP):

- Section 109 would reduce the number of waivers available for certain childless adults who would otherwise be subject to work requirements or time limits. CBO estimates that, on average, about 1 million people with higher-than-average benefits would lose eligibility for SNAP benefits under this provision. The number of people losing benefits would decline from 1.7 million in 2014 to 0.5 million in 2023.
- Section 105 would restrict categorical eligibility, a current policy that allows states to determine eligibility for SNAP based on receipt of benefits in other programs for low-income people. CBO estimates that, on average, 1.8 million people with lower-than-average benefits would lose eligibility for the program if this provision were enacted. The number of people losing benefits would decline from 2.1 million in 2014 to 1.5 million in 2023.<sup>35</sup>

Sometimes a single point estimate of your best guess about the degree of magnitude will suffice. But in most cases, you should provide a range.

## Trends Might Be the Basis of Projections

Projecting outcomes is about the future, but this does not mean that one can never glean useful information from trend data. Past trends will only provide a reasonable basis for making projections, however, if the implicit assumption holds that whatever factors influenced changes in outcomes in the past will continue to operate the same way going forward. For example, if the analyst seeks to use past trends in local public school district expenditures as the basis for a projection of how much the school district will spend over

the next decade, the analyst will need to determine if changes in student enrollment, among many other factors, will have approximately the same influence on district spending levels in the future as they did previously. But not all trends are stable. Maybe, to continue with this example, per-student education costs have been changing due to student demography, a shifting mix of state and federal education mandates, and other reasons so that an X percent increase in the student enrollment level may not result in a Y percent increase in district expenditures that was previously associated with an enrollment increment of the same size. Data series can also be subject to seasonal (e.g., traffic congestion levels) or cyclical (e.g., economic fluctuations associated with the business cycle, such as unemployment trends) trends, which also need to be taken into account.

### Break-Even Estimates Can Shrink Uncertainty

“You have no evidence this will work,” carp your critics. You—quite correctly—respond, “You have no evidence it won’t.” You are both right, because “evidence” about events that have not yet occurred is a contradiction in terms. Nevertheless, your critics make the valid point that you probably can’t be very certain that your recommended policy option will work and that the *burden of justification* (not, of course, a burden of literal “proof”) falls on you.<sup>36</sup> You will want to take up this burden using whatever strategic leverage you can muster.

This means that you will set the bar of justification as low as is reasonable. Typically, you will want to claim only that the recommended course of action is “sufficiently likely” to produce results that are good enough to justify the known costs and risks. This approach is known as “break-even” or “threshold” analysis. It is an astonishingly powerful—yet simple, intuitive, and commonsensical—conceptual lens. It builds a decision framework out of what is known or reasonably assumed and handles the residual uncertainties by comparing them to elements in this more secure frame.

Suppose, for instance, that some youth-guidance-oriented policy meant to reduce incarceration of juveniles is under consideration and has known costs of \$1 million, but the level of effectiveness is speculative. You build a decision frame out of (1) a decision rule that says, “If the benefits exceed the costs, do it,” and (2) a known fact about the costs, \$1 million. You then evaluate the remaining uncertainty in these four steps:

1. Locate the point of minimum acceptable effectiveness given the costs. Ask: “What is the minimum level of effectiveness this policy would have to achieve in order to justify our spending \$1 million?” Your answer: “Different observers have different opinions about how much avoiding an incarceration is worth, but leaving that aside and going with my own values, I’d say that a 15 percent reduction is the minimum I would accept given the expenditure of \$1 million.”<sup>37</sup>
2. Referring back to your model of the processes that create the problem and hold it in place, ask: “What new processes, or changes in old ones, could

conceivably produce this level of effectiveness?” This is largely a qualitative analysis. The answer might be this: “Based on previous documentation of how the guidance process works, we can safely say that it works in different ways with different sorts of kids—when it works at all, that is. It can provide about half the kids more constructive life choices; in about a quarter of the cases, it works through heightening the (realistic) perception of punishment; and in about a quarter of the cases, we are just crossing our fingers.”

3. Assess how likely (or unlikely) it is that the processes for improvement thus identified will actually produce the required—that is, the break-even—level of effectiveness. It is particularly helpful to ask whether the break-even level (15 percent, in this case) looks like a plausible number given what is known or assumed about the effectiveness in similar circumstances of similar sorts of interventions. If the number is implausibly high, you might then go on to ask whether special circumstances of some sort might be at work in this case to help achieve it. Note that in this and the previous step you must rely on what we might think of as “theory,” or self-conscious and evidence-based reasoning about the way causal processes work. Typically, these are the weakest links in the chain of policy-analytic reasoning. That is why it is particularly important—and particularly difficult—to take this step as thoughtfully, self-critically, and responsibly as possible.
4. Estimate the probability of failure and the political and other costs of having to accept failure—asking yourself whether these costs would be tolerable should they be incurred.

The federal government frequently uses break-even analysis to assess the merits of proposed regulations. For example, the federal government considered whether to issue a regulation designed to reduce the incidence of prison rape. The annual cost of implementation of the regulation was estimated at \$470 million. The agency was not able to project the number of prison rapes the regulation would prevent. In addition, it had difficulty monetizing the cost of prison rapes to both victims and society. Under break-even analysis, however, the agency decided to go forward with the regulation. It found that at least 160,000 prison rapes occur each year, and it concluded that if a single rape prevention is valued at \$500,000, the rule would be justified if it prevented only 1,600 rapes, or 1 percent of the total. The agency was confident that the new rule would achieve at least this minimum level of effectiveness.<sup>38</sup>

In the hope that it may be helpful to encourage readers to use break-even analysis (when appropriate, of course), we offer two more examples:

- Policy X for establishing a chain of wildlife refuges looks like an excellent choice to implement a broader conservation agenda, provided that the funding comes through as planned. But it might not, because federal grant-in-aid

resources may not be forthcoming, or the governor may give the policy lower priority than they now promise, or some development interests that have their eye on two of the designated sites may find a way to block it. You interview your client, a state environmental agency director, and determine that they like the program so much that they are willing to go for it if it has at least a fifty-fifty chance of working out. Your analysis can then focus the client's attention on why, after considerable research, you have concluded that it has a somewhat better (or somewhat worse) chance than fifty-fifty, even though you may find it impossible to specify exactly how much better (or worse).

- Building a new stadium for the Hometown Heroes looks like a good idea, given the nature of the costs and benefits, if average daily attendance turns out to be no fewer than ten thousand. That's the break-even attendance figure for you and the relevant decision-makers. It's up to them to decide, first, how confident they are that this break-even level will be reached and, then, whether that degree of confidence is enough to warrant making an affirmative decision. You can thus organize your presentation of facts and opinions to focus on these two key issues.<sup>39</sup>

**Semantic Tip** Assuming for the moment that benefits are uncertain while costs are not, ask yourself these two questions: (1) "Given what I know for sure about the costs of this alternative, what is the minimum help we need to get from Condition X to ensure adequately offsetting benefits?" and (2) "How reasonable is it to believe that Condition X will actually produce that minimum?"

See Box I-6 for an excerpt from a policy analysis report that confronts uncertainty.

## BOX I-6: AN ILLUSTRATIVE DISCUSSION OF CONFRONTING UNCERTAINTY FROM A POLICY ANALYSIS REPORT

### OPGEE Challenges

*The largest source of uncertainty in OPGEE [Oil Production Greenhouse Gas Emissions Estimator] is the lack of information on global oil fields. Many operators and many regions of the world have few formal data publication requirements. Data quality is also an ongoing issue in modeling upstream emissions. . . .*

*OPGEE utilizes about 50 data inputs, from simple entries like the name of the country where an oil field is located to challenging-to-obtain information such as an oil field's productivity index (expressed in daily production per unit pressure). Substantial research is involved in gathering OPGEE modeling data, which can be obtained from agencies, reports, scientific literature, and industry references.*

*OPGEE can function with limited data. The model has a comprehensive set of defaults and smart defaults that can fill in missing data. The more data found for a particular field, the more specific and less generic the emissions estimate becomes. All data are used to determine smarter default values over time.*

*As with all life-cycle assessment (LCA) models, boundaries must be drawn around the analysis. The handling of co-products that cross boundaries along the oil supply chain, from extraction to refining to end use, presents methodological challenges. For example, resulting GHG [greenhouse gas] emissions from condensates of light liquids, like ethane, that can be stripped off and sold before oil is transported to a refinery are not expressly included in OPGEE. Emissions associated with exploration occur at the beginning of an oil field development project and are spread over the life of the field. Extraction emissions that occur routinely are estimated at a point in time and assumed to recur over the lifetime of the oil field.*

*OPGEE treats liquid petroleum as the principal product of upstream processes. Emissions associated with electricity generated on-site or natural gas produced that is gathered, sold, and not flared is credited back or deducted from total emissions in OPGEE accounting. Any emissions from co-products like petcoke that are associated with upgrading heavy oils upstream of the refinery—as can be the case with Canadian bitumen and Venezuelan heavy oils—are not included in OPGEE unless the production process directly consumes petcoke [as in some oil-sands-based integrated mining and upgrading operations]. Emissions from net production of petcoke have been included in the OPEM [Oil Products Emissions Module] downstream combustion module.*

*Recent studies have found that uncertainty in OPGEE's results is reduced after learning three to four key pieces of data about an oil field. After learning the ten most important pieces of information about an oil field, there is typically little benefit to learning the remaining data.*

*Imprecise data reporting introduces additional uncertainty. Errors in applying the model can lead to further uncertainty.*

*The key variables to enhance model precision include: steam-to-oil and water-to-oil ratios, flaring rates, and crude density [measured as API [American Petroleum Institute] gravity]. Less important variables in the OPGEE model's ability to analyze GHG emissions include gas-to-oil ratios, oil production rates, and depth [except in extreme cases].*

---

*Source: Deborah Gordon, Adam Brandt, Joule Bergerson, and Jonathan Koomey, *Know Your Oil: Creating a Global Oil-Climate Index* (Washington, DC: Carnegie Endowment for International Peace, 2015), 19–20, [https://carnegieendowment.org/files/know\\_your\\_oil.pdf](https://carnegieendowment.org/files/know_your_oil.pdf).*

*Note: Endnotes, figures and references to figures, and appendix omitted from excerpt.*

## Try Sensitivity Analysis

Which uncertainties are the most important, in the sense that relatively small changes in what you believe would cause you to change your mind about how desirable some alternative might be? By a process known as sensitivity analysis, you can discover these most important uncertainties. The procedures are somewhat technical (Morgan and Henrion 1990, chap. 8), but the intuition behind them is simple. Consider the several assumptions you have made on the way to your conclusion and suppose that each of

them is somewhat mistaken. Now ask yourself this: “How big a mistake can I afford in this assumption before this analysis is in really big trouble?” The smaller the affordable mistake, the more sensitive is your analysis to the particular assumption.

It is not hard to examine these assumptions one at a time. But what if they pile up in such a way that you are “somewhat” wrong on two or three or four assumptions all at once? This situation is typically dealt with by a technique called “Monte Carlo simulation,” which begins by recognizing that each assumption is in itself probabilistic and then combines the probabilities behind the assumptions to create a new set of probabilities about how the combination of assumptions will turn out. You can then say something like this: “Given the many possible scenarios that might occur, there is an 82 percent chance that the actual scenario would exceed our break-even requirement.”<sup>40</sup>

But suppose that projections must be made for a future beset by multiple uncertainties, like climate change or the global configuration of military forces and technologies twenty-five years off, for which probability distributions are not known or are controversial. One promising approach makes use of any of an emerging set of computer-assisted projection techniques, generally known as long-term policy analysis. This approach is similar to Monte Carlo simulation in that it starts with scenarios about alternative futures, but instead it searches for policy choices that would be “robust,” in the sense that they would not necessarily be the best but would satisfy the whole, or nearly the whole, array of minimum policy desiderata. The objective is to minimize the maximum “regret” that relevant parties might experience.<sup>41</sup>

## Confront the Optimism Problem

Great ventures require optimism. Because even small ventures by government can affect so many lives, they are in their own way great. Hence, some realistic optimism is beneficial. But how do you guard against *excessive* optimism?

## Scenario Writing

What scenarios might cause the proposal to fail to produce the desired outcome—that is, solving or sufficiently mitigating the policy problems? Do not create such scenarios from whole cloth; be realistic. And yet, let your imagination run a little so that you have a good chance of thinking of the most dangerous possibilities. In particular, think about the dangers of the implementation process, political and otherwise. Scenario writing also benefits from thinking about possible failures from a vantage point in the future looking backward. Consider the following scenarios:

- In a health or safety regulatory program, the scientific or technical knowledge necessary to produce rational and legally defensible standards may prove to be lacking. As a result, five years from now, symbolic politics, corruption, industry capture, or excessive regulatory zeal will have filled the vacuum.

- Time passes, and budgetary resources and political support that were once available slip away under the impact of electoral change and shifts in the economy. A terrorist-identification program, begun under nurturant leaders and accompanied by editorialists' applause, will have become consolidated with another program then taken over by a different bureaucratic unit and eventually will have disappeared.
- A successful state program designed to furnish technical assistance to extremely poor rural counties will have added a mandate to aid many not-so-poor urban counties, with the result that scarce program resources will have been dissipated and squandered. (We call this scenario "piling on"; see Bardach 1977.)
- A program that subsidizes research and development of "fish protein concentrate," intended as a cheap and nutritious food additive, is launched with great fanfare. Five years from now, it will have been stalled, permanently, by the US Food and Drug Administration, which will not have been able to assimilate this product into its standard operating procedures for regulatory review.

**Semantic Tip** Notice that these scenarios are written in the future perfect tense. Use of this verb tense encourages concreteness, which is a helpful stimulant to the imagination (Weick 1979, 195–200). It often helps your scenario writing to start with a list of adverse implementation outcomes, conjuring up one or more scenarios about how each of them might occur. Remember the list of such outcomes embodied in the scenarios just described: long delays, "capture" of program or policy benefits by a relatively undeserving and unintended constituency, excessive budgetary or administrative costs, scandal arising from fraud and waste, and administrative complexities that leave citizens (and program managers) uncertain as to what benefits are available or what regulations must be complied with.

It is common for actors to perform a "postmortem" analysis when a policy dies or decays. Building on the scenario writing idea, a more proactive approach is to conduct a "premortem analysis."<sup>42</sup> The actors working on a policy are assembled *before* the policy is implemented and told to assume the policy has failed. The challenge of this brainstorming exercise is to come up with a plausible set of reasons why the policy did not achieve its objectives. By encouraging an honest conversation about a policy's vulnerabilities, a premortem analysis may surface ways to increase a policy's robustness and its capacity to survive the pitfalls of the implementation process.

**Semantic Tip** *Undesirable side effects.* Analysts are often cautioned to think about "unanticipated consequences." But this term is not appropriate, for it is often used to refer to perfectly anticipatable, though undesirable, side effects. Here are some common undesirable but foreseeable side effects in public programs:

- *Moral hazard* increases. That is, your policy has the effect of insulating people from the consequences of their actions. For example, increasing the size of

unemployment benefits has the side effect of blunting the incentives to search for a replacement job.

- Reasonable regulation drifts toward *overregulation*, especially if the costs of overregulation are not perceptible to those who bear them. One possible adverse result of setting health or safety standards “too high” and enforcing them “too uniformly” is that you increase private-sector costs beyond some optimal level. For instance, given most people’s preferences for safety, imposing auto bumper standards that cost some \$25 per vehicle but have only trivial effects on improving vehicle crashworthiness would not pass a conventional benefit–cost test.

A second adverse result of overregulation might be that you inadvertently cause a shift away from the regulated activity into some other activity that—perversely—is less safe, less healthful, or more harmful. For instance, some observers argue that overregulating the safety features of nuclear power production has caused a shift toward coal, which they argue is much more hazardous than nuclear power.

- *Rent-seekers*—that is, interests looking out for profitable niches protected from full competition—distort the program to serve their own interests. It is not inevitable that suppliers of goods and services to the government, including civil servants, will find ways to capture “rents,” but it often happens (e.g., with many defense contractors). Rent-seeking also occurs in less obvious ways—as when some regulated firms successfully lobby for regulations that impose much higher compliance costs on their competitors than on themselves.
- *The outcomes produced by one part of a complex policy design undermine the performance of another.* Policies sometimes contain multiple parts, such as both expenditure (or regulatory) and revenue-generation components. Undesirable consequences can arise when one part of the design produces outcomes that counteract the performance of another. For example, an education program paid for by earmarked revenues from a “sin” tax on cigarettes can lead to funding shortfalls for schools if the tax causes many smokers to quit. The same dynamic can occur if a regulatory inspectorate is financed by fines on violators. As the regulations take hold, fewer violations will occur, and the revenue to pay for inspections will dry up. To avoid this problem, analysts should seek to design policies whose constituent parts produce mutually reinforcing outcomes.

### The Ethical Costs of Optimism

It is hard to overstate the importance of worrying about the possible adverse side effects of otherwise “good” policies, not to mention the possibility that even intended main

benefits may fail to materialize under many circumstances (see the chapter on “assessing your ignorance” in Behn and Vaupel 1982). The ethical policy analyst always poses the question, “If people actually were to follow my advice, what might be the costs of my having been wrong, and who would have to bear them?” Keep in mind that the analyst typically is *not* one of the parties who have to bear the costs of the analyst’s mistakes.

*To minimize the risk of undesirable side effects, take into account the incentives of the “targets” of behavioral change.* A fundamental reason why a “solution” to a troubling situation may generate adverse consequences is that the analyst has failed to think about the incentives of the actors whose behavior a policy intervention is meant to alter. The point is not necessarily to condone the goals of such actors but rather to understand what really drives their behavior—so that a more effective, incentive-compatible intervention can be fashioned.

Consider “Ban the Box,” a well-intended policy to prevent employers from inquiring about a person’s criminal record in an initial job application. (Employers would still be allowed to ask about criminal records later in the hiring process.) The policy’s aim is to give people with a criminal record a greater chance to interview for jobs and demonstrate their skills and qualifications before final hiring decisions are made. Advocates hope that “Ban the Box” will boost the low employment rate among ex-offenders, lessen racial disparities in employment, and reduce recidivism.

Unfortunately, rigorous studies have discovered that “Ban the Box” has the undesirable effect of reducing employment opportunities for young, low-skilled Black men *without* criminal convictions.<sup>43</sup> What explains this outcome? Employers want to hire reliable, productive workers. Many have a preference against hiring people with criminal records because ex-offenders are more likely than nonoffenders to have a history of violence or other antisocial behavior. If employers are prevented from knowing which applicants have a criminal record, they may respond by not interviewing young, low-skilled Black men—the group most likely to have recent convictions.<sup>44</sup> In other words, employers engage in what economists call “statistical discrimination”—that is, they make assumptions about individuals based on averages among a group. To be sure, “Ban the Box” does not prevent employers from rejecting applicants with criminal records during the later stages of the hiring process. But screening and interviewing individual candidates is time-consuming, so busy employers may just avoid entire demographic groups altogether.

One plausible response to this adaptation by employers is to rely on civil rights laws banning racial discrimination in employment. However, such laws are difficult to enforce, especially with respect to discrimination that occurs early in the applicant review process.

An alternative approach to promoting the goal of increasing employment among ex-offenders (and reducing racial disparities in hiring) would take into account the desire of employers to have information about the productivity of applicants. The employers’ desire for information could be satisfied by allowing individuals with

criminal records to obtain “employability certificates” from courts.<sup>45</sup> Such certificates (which could be based on a judge’s review of an ex-offender’s completion of a training program) could help convince employers that an applicant with a criminal record has been rehabilitated and is “work-ready.”

The “Ban the Box” case offers several broader lessons for policy analysis and design. First, adaptation is a general phenomenon and a frequent source of failure and backfire. A key challenge is thinking about how actors will respond to changes in the policy environment. It is often a good idea to consider within the menu of policy options alternatives that work with rather than against the incentives of policy targets.<sup>46</sup> Second, seemingly small changes (such as changing the timing of when an employer can learn about a job candidate’s criminal record during the hiring process) can affect big movements in a system. Finally, information flows are of critical importance. The quantity and quality of information available to actors influences the level of uncertainty and thus how actors respond to policy measures and pursue their objectives.

## The Emergent-Features Problem

Policy often intervenes in systems of some complexity, systems populated by actors who adapt to your interventions in surprising ways and whose adaptations lead other actors to create still further adaptations. Surprising behavior may emerge from such dynamics. How can you take such possibilities into account when you make your projections?

In many cases, you cannot, for the systems are too complex and too little understood. The macro-economy is an extreme case—the hypothetical responses of producer interests to “supply-side” tax cuts are a major source of contention between those who think the taxes generated by a growing economy will substantially offset the direct effects of the cuts and those who are deeply skeptical of this scenario. Few cases are that extreme, however, and you might make some progress with what might be called “the other-guy’s-shoes” heuristic.

Imagine yourself in the other guy’s shoes. Say to yourself, “If I were X, how would I act?” And then proceed to crawl into X’s mind and play out, in your own mind, what X might do. Do this systematically for each of the important stakeholders or other affected parties. The value of this exercise is that you will discover them to be adapting in surprising ways to the new policy situation you may be creating.

For example, under chemical right-to-know laws, workers must be told what substances they have been exposed to, and they may examine health records maintained by employers. If you were a worker, how might you use this law? Might you use the information to quit your present job? To demand a higher wage or more protective equipment? To sue your employer or put pressure on your union representative?

And how would your union representative react to such pressure? Might this pressure make the representative’s job harder—or perhaps easier in some way?

Now, suppose that you were an employer. Given what you expected your workers to do, you would face incentives to make adaptations or countermoves. Might you stop

keeping all health records not explicitly required by law? Or continue keeping records but permit doctors to perform only selected lab tests? And if you were a worker and saw your employer doing these things, what countermoves would you make?

Not all the moves and countermoves of players wearing the other guy's shoes will necessarily lead to trouble for the policy alternative you are evaluating. Many such adaptation sequences may prove to be helpful, in the sense that they may help society to adjust to the changes set in motion by the new policy. At some point in the 1970s, the Federal Trade Commission (FTC) attacked the problem of retailers evading implied warranty obligations for defective products by selling installment debts to banks and other collectors that had no duty, under the so-called holder-in-due-course doctrine, to fix the product or to refrain from collecting on the installment debt. The FTC solution was, in effect, to abolish the protections of the holder-in-due-course doctrine. Banks complained that they did not want to go into the toaster repair business. But if you put yourself in the shoes of a bank manager suddenly obliged to become a toaster repairer, might you not have thought of contracting out your repair obligations to repair specialists, or perhaps arranging not to buy installment debts from retailers who you believed could not be relied upon to make good on their implied warranties?

### Construct an Outcomes Matrix

The step of projecting outcomes leads you into a dense thicket of information. At some point along the way, you will probably need to stand back and assess complex and uncertain scenarios for perhaps two to five basic alternatives, combined with their principal variants. A convenient way to get an overview of all this information is to display it in an *outcomes matrix*. The typical outcomes matrix format arrays your policy alternatives down the rows and your evaluative criteria across the columns. Each cell contains the projected outcome of the row alternative as assessed by reference to the column criterion.

Table I-2 is an example. It appeared in a report by four Berkeley students in 2008 that had been requested by the international environmental group ICLEI—Local Governments for Sustainability.<sup>47</sup> They projected outcomes for eight alternatives (“scenarios,” in their usage) across five criteria (in three clusters).<sup>48</sup> We do not vouch for the accuracy of their projections, though they tried the best they could to synthesize the diverse and sometimes contradictory research literature as it existed at the time. Of greater interest is their attempt to fill in the cells in a canonical matrix form. The alternatives are listed down the rows and the criteria across the columns. Three criteria are lumped together under the heading “Viability,” though if space had permitted, the students might have made a separate column for each. The analysis applies to a representative US city called Anytown. Note that the matrix is labeled as a “comparative analysis.” Each projection is compared to a baseline projection for the year 2050, showing only the difference between the baseline projection and the estimated projection for the indicated alternative. Most cells contain a projected range rather than a single

point estimate. In Step Six, we come back to Table I-2 and discuss how this comparative setup facilitates confronting the trade-offs.

An outcomes matrix at this stage of your work is a scratch-pad affair, useful for you and your team members and perhaps a friendly outsider or two. Its main function is to help you see what you have in hand and what you still need to learn about. A secondary function is to prepare to confront the trade-offs (see Step Six). If the matrix looks to you large and complicated, you may be encouraged to shrink it: conceptualize some alternatives as mere variants of more or less the same thing, get rid of alternatives that are obvious losers, and omit criteria that don't differentiate among alternatives (i.e., all the alternatives appear to do about as well or as poorly with respect to these criteria). The students who produced Table I-2 excluded three alternatives that they had originally considered: a local cap-and-trade program, leveraging collective purchasing power in energy markets, and urban forestry.

You may find it useful to go through this exercise more than once, as your analysis evolves over time. (Table I-2 is the final version of several matrices that the student group made.)

A later version of such a matrix may also prove useful when you tell your story (see Step Eight). However, unless the matrix is very well designed and explained, it can impede the flow of your story rather than assist it.

**Semantic Tip** Here is a tip with a graphic dimension. Take advantage of the fact that being listed earlier (more leftward) in the matrix is usually taken to signify greater importance. Even if you are unsure how to weight criteria on some cardinal scale, with equal intervals assumed between all points, you might feel better about an ordinal scale, requiring judgments only of more than and less than. Put what you think should be the weightier criteria in the more leftward columns. A common error that occurs in labeling the criteria columns in such a matrix is to fail to indicate what value is at stake and in what dimensions the measurement is being done. For instance, if you are assessing a rental subsidy program and you enter a plus sign in a column labeled "Landlord/Tenant Relations," the reader may not know whether you think relations will become more harmonious, more confrontational, less dominated by landlords, less dominated by tenants, or something else. It is not sufficient that your surrounding text makes your intention clear; the matrix label itself must be informative. In many cases, it helps to insert the term *maximize* or *minimize* in the criterion label. Table I-2 is exemplary in almost all respects, except that the column labels do not include such words. It happens that the meaning is quite clear from the context, of course, but in the interests of "analytic hygiene" it would have been better to include them.

If you cannot fill in the cell with a quantitatively expressed description of the projected outcome, you may have to settle for a verbal descriptor such as "very good" or a symbolic descriptor such as + or -. The operative word here, though, is *cannot*. Quantification goes a long way toward making an analysis useful, and rough yet adequate quantification is easier than you might suppose. Remember, also, the heuristic of increasing or decreasing "the odds," mentioned in "Step One: Define the Problem."

In listing or stating criteria, speak in the declarative, not the interrogative. “How equitable is the final budget outcome?” is not a criterion; it is a question. “Maximize equity” is a criterion.

## But Policy Contexts Differ

Suppose the policy in question applies to heterogeneous policy contexts, such as different states or different counties within a state or different neighborhoods within a city. Suppose, further, that we should expect policy context to matter. Perhaps the policy would be suited to a well-off urbanized state but not to a low-wealth rural state or to a state with a strong populist tradition. Although the authors of the aforementioned greenhouse gas study may have been justified in using a hypothetical average “Anytown” as their base case, this is not always sensible. Demographic and other differences in context may imply that, while Anytown represents the broad mainstream well enough, “outlier” cases exist that could fare very differently.

One way to handle this is to break the analysis into as many different chunks as you need to handle the variety of important policy contexts. This should help you conceptualize variants that “tweak” a basic policy strategy so as to fit better the variety of contexts in which it will have to work. A citywide policy to encourage fire-resistant roofing materials on new and replacement roofs, for instance, could also incorporate loan assistance to building owners differentially targeted to high-risk areas and/or to higher- and lower-income property owners.

But what is a “policy context” anyway? “Context” does not have a stable, nicely circumscribed meaning. Just about anything that isn’t “the alternative being considered” can be called “context.” Income, race, residential density, and other such demographic features are often taken to be important to policy context. But this is not always true. And sometimes features that are not “obvious” or commonsensical are indeed important, like the degree of prior experience a community has had absorbing new immigrants or implementing novel central-government initiatives. In an abstract sense, policy-relevant context features are those that you cannot control but that probably make a difference to the eventual worth of the chosen policy.

The features of a relevant context are sometimes numerous and interconnected, so much so that one wants to handle them as a stylized bundle. For example, a fairly large jurisdiction, like an American state, often contains a number of smaller jurisdictions to which a statewide policy is applied. It would be unwieldy and unhelpful to analyze the context of each of the hundreds or thousands of localities with a given state separately. If, however, you can divide the whole field on which policy is to be applied into just a few such bundles, you can do subanalyses of three or four of them and feel some confidence that you have a sample that, if not “representative,” nevertheless captures the mainstream and the principal outliers.

So how do we fit this into an analysis done for a statewide policy in a state with many different-sized cities? The first task is to reduce the number of contexts to

something manageable. Three is a nice number. The second is to create abstractions—“ideal types” in social science jargon—that effectively stand in for real places. The ideal types can even be given names for easy reference—for instance, “Gotham” for very large cities, “Middletown” for medium-sized cities, and “Fernville” for small localities. Each of these ideal types can be attached to the problem definition. At this point, doing so expands the number of problems from one to three. We are creating a “separate problem” for each ideal-type-characterized context.

## Setup for the Next Step

A useful test of whether the projecting outcomes step has been done well is that the outcomes should be characterized in such a way that is easy for the analyst (or anyone) to calibrate the trade-offs. (Calibration, of course, is not the only process involved in confronting trade-offs, as will be discussed in the coming section; values matter too.)

Begin by making a table like Table I-3, which concerns a policy problem faced by a municipal library system in reaching out to the poorer, and less inclined toward reading, areas of the city. This table projects the whole outcome, with respect to each of the important criteria, of each alternative under consideration, including (of course) “Let present trends continue.” It provides the big picture of what is at stake in the world for the choices at hand.

Box I-7 displays a specimen from the outcomes projection section of a policy analysis report. Note the use of a base case, the acknowledgment of data limitations, and that the anticipated increases in primary care physicians are discussed in terms of the amount per hundred-thousand population to facilitate comparisons of projected outcomes across alternatives.

**TABLE I-3 ■ Outreach to Poorer Areas by Municipal Library**

Alternatives/ outcomes	Books borrowed in outreach areas (per year)	Borrowers in outreach areas (per year)	Annual cost (\$ per year)	Net political payoff to council member Y*
Base case: continue present trends	60,000	15,000	2,000,000	0
Bookmobile biweekly each area	80,000	20,000	2,500,000	200; 150
Bookmobile once weekly only the poorer areas	75,000	18,000	2,350,000	100; 50
Book fairs twice yearly in 25 public schools	85,000	21,000	2,100,000	0; 0

\*Additional votes in next two elections minus votes lost.

## Confidence Intervals and Regrets

Given the various and numerous uncertainties involved in any projection, it is a good idea to estimate not only a mean for the projection but a confidence interval as well. You might also state the odds that a negative outcome will occur such that a decision-maker might regret the choice of any particular outcome. And don't forget to include in these estimates your analysis about what might happen in the Last Mile (see the next section).

### BOX I-7: AN ILLUSTRATIVE EXAMPLE OF “PROJECT THE OUTCOMES” FROM A POLICY ANALYSIS REPORT

#### Base-Case Projections of the Rural and Urban Primary Care Workforce

*Our predictive models estimated declines in the number of primary care physicians per 100,000 population in both rural and urban areas from 2013 to 2025: 3.66 fewer primary care physicians per 100,000 population in rural counties by 2025, 4.14 fewer in urban counties, 5.07 fewer outside Seattle, and 3.22 fewer within Seattle. These estimated declines were driven largely by recent increases in the percentage of primary care physicians ages 55 and older, many of whom are likely to retire by 2025. In contrast, we projected increases from 2013 to 2025 of 5.38 to 7.79 nurse practitioners (NPs) and 1.84 to 3.08 physician assistants (PAs) per 100,000 population in Washington State. . . .*

#### Open the Elson S. Floyd College of Medicine at Washington State University [WSU]

*We estimated that opening the new medical school in 2017, beginning with 60 students and reaching a steady-state enrollment of 320 students in 2022, would be associated with increases in 2025 of 0.39 primary care physicians per 100,000 population in rural Washington counties, 0.59 in urban counties, 0.76 in Seattle, and 0.39 in Washington counties outside Seattle. These estimated effects of the new medical school offset approximately 11 percent of the projected decrease in rural per capita primary care physician supply by 2025, 14 percent of the projected decrease in urban counties, 12 percent of the projected decrease within Seattle, and 15 percent of the projected decrease outside Seattle.*

#### Increase the Number of Primary Care Residency Positions in Washington State

*We modeled residency policy options ranging up to a 100-percent expansion (i.e., a doubling of primary care residency sizes outside Seattle). The estimated effects of 100-percent primary care residency expansion (adding 36 primary care residents) were*

larger than the estimated effects of opening the new medical school at WSU, without residency program expansion. However, none of the modeled residency scenarios had an estimated effect sufficient to offset the predicted decline in the number of rural primary care physicians (or primary care physicians outside Seattle) per 100,000 population. For the 100-percent residency size expansions, estimated effects ranged from 1.11 primary care physicians per 100,000 population (27 percent of the projected decrease) in urban counties to 2.00 primary care physicians per 100,000 population (55 percent of the projected decrease) in rural counties by 2025.

## Increase the Availability of Educational Loan-Repayment Incentives

To estimate the effect of expanding state-funded loan-repayment incentives in rural areas, we analyzed relationships between the number of National Health Service Corps (NHSC) primary care positions and primary care supply in rural counties. We found that, for each new primary care NHSC position opened per 100,000 county population, the estimated increase was 0.24 primary care physicians per 100,000 county population. Therefore, we estimated that doubling the number of primary care NHSC positions in rural Washington State (by adding 30 more such positions to rural counties, with approximate cumulative population 700,000) would produce an increase of 1.03 primary care physicians per 100,000 population.

## Improve the Quality of High School Education in Rural Washington State

Because we lacked longitudinal data on high school quality (measured as proficiency rates on standardized tests of mathematics and of reading and language arts), we fit cross-sectional models that estimated the effect of increasing proficiency rates on these standardized tests by 0.2 standard deviations among high schools in rural Washington counties. We estimated that this improvement in high school quality would be associated with an increase of 0.80 primary care physicians per 100,000 population in rural Washington, or approximately 22 percent of the projected decline in per capita rural primary care physicians expected by 2025. However, because these models were cross-sectional and the time required to improve school performance is unclear, we cannot estimate the number of years required to achieve this estimated effect.

---

Source: Mark W. Friedberg, Grant R. Martsolf, Chapin White, David I. Auerbach, Ryan Kandrack, Rachel O. Reid, Emily Butcher, Hao Yu, Simon Hollands, and Xiaoyu Nie, *Evaluation of Policy Options for Increasing the Availability of Primary Care Services in Rural Washington State* (Santa Monica, CA: RAND Corporation, 2016), xii–xiv, [https://www.rand.org/pubs/research\\_reports/RR1620.html](https://www.rand.org/pubs/research_reports/RR1620.html).

Note: The report also projected outcomes from several other alternatives, including preserving rural hospitals in Washington State and increasing Medicaid payment rates for primary care physicians in rural Washington State, among others; references to tables omitted from excerpt.

## The Last Mile

Try to be honest with yourself and others: the policy plan you have developed during the policy adoption phase might or might not look much like what happens once the policy is up and running. There could be many causes, but obstacles discovered in The Last Mile are sure to be among them.

One source of this uncertainty is how the interplay of safeguard costs and effectiveness plays out. Surprises are sure to come. For instance: personnel turnover might deprive an agency of needed institutional memory and customized expertise; cross-cutting regulations designed to curb abusive agency discretion might also limit constructive, welfare-enhancing, discretion; an advisory group set up to improve agency decision-making might turn out to be a Trojan horse that allows committed adversaries to enter the agency's inner sancta.

A second source of uncertainty is all that happens in the implementation phase, that is, setting up the whole program delivery apparatus. In the modern era, this apparatus is likely to be technically complex. It is also likely to involve several agencies; link public and private (both nonprofit and market) program delivery organizations, involve more than one level (federal, state, local) of government; rely on public willingness to cooperate in doing certain tasks (fill out tax forms, comply with water-saving guidelines, etc.); and satisfy the rules of all the "overhead agencies" of government that keep it functioning (e.g., budget review, space allocation, fleet management, civil service and personnel, contracting). Each connection point presents an opportunity for delays attributable to review and negotiation time, unexpected technical problems, political interests to forward their agendas, and so on. To make projection even harder, these many connections constantly change the implementation landscape, creating emergent legal and political constraints and drawing in new players from the sidelines.

Unfortunately, there is not much academic or practitioner wisdom about how to cope with the last-mile uncertainties of outcome projection. Of course, a simple design helps. So does the ability and willingness to look around and see what has happened when other jurisdictions have set up similar programs.

Another problem is focusing too much on others' successes ("let's thoughtlessly imitate them . . .") and not enough on their shortcomings and failures ("let's see how we can avoid the traps they seem to have fallen into . . .").

But we do have one prescription: unless there are compelling reasons not to, share your uncertainties. If your projections are expressed quantitatively, report not only the expected mean but also the standard deviation around the mean. Analogously, for qualitative description, you could pick a real-life criterion of interest, like a cost below which some policy would need to fall in order for it to be an acceptable choice, and then describe the likelihood with which that might be achieved. For instance, "Ridership would have to be at least 5,500 trips a day for this transit project to break even, and this is very likely provided no new highway capacity is added." This, of course, is exactly the sort of break-even analysis we have, earlier, recommended in general.

## STEP SIX: CONFRONT THE TRADE-OFFS

---

It sometimes happens that one of the policy alternatives under consideration is expected to produce a better outcome than any of the other alternatives with regard to every single evaluative criterion. In that case—called “dominance”—there are no trade-offs among the alternatives.<sup>49</sup> Usually, though, you are less fortunate, and you must clarify the trade-offs between outcomes associated with different policy options for the sake of your client or audience.<sup>50</sup>

### Focus on Outcomes

A common pitfall in confronting trade-offs is to think and speak of the trade-offs as being across *alternatives* rather than across projected *outcomes*—for example, “trading off twenty foot-patrol police officers in the late-night hours against a lower-maintenance-cost fleet of police vehicles.” Although such a trade-off exists, with a second’s thought you’ll see that you can’t do anything at all with it. Both alternatives must first be converted into outcomes before genuine trade-offs can be confronted. Thus, the competing outcomes might be fifty (plus or minus) burglaries per year prevented by the foot-patrol officers versus a savings of \$300,000 in fleet maintenance.

The most common trade-off is between money and a good or service received by some proportion of the citizenry, such as extending library hours from 8 p.m. till 10 p.m., weighed against a cost of \$200,000 annually. Another common trade-off, especially in regulatory policies, involves weighing privately borne costs (a company’s installing pollution abatement equipment) against social benefits (improved health of the affected population and the protection of forests). If the projected outcomes can be monetized—that is, expressed in dollar terms—it is sometimes simple to evaluate the trade-offs. Just choose the option that yields the largest net value, once costs have been subtracted from benefits. This procedure applies nicely if budgets, and therefore the scope of the activity, are not limited. But it can run afoul of another monetary consideration, cost-effectiveness per unit of activity, if budgets or other inputs are limited. In Table I-2 (p. 53-54), note that the high-density residential development option—presumably limited in extent because of the limited likely scope of new development—is more cost-effective than any other activity but is less efficacious than all the others, too.

In Table I-2, we see that there is no dominant outcome. The really efficacious options, involving a carbon tax, are not viable politically. Retrofitting existing buildings is apparently more efficacious than meeting green standards in new buildings, but it is a strategy deemed by the student-authors to be less viable than the latter. And, as we said earlier, making new developments denser, although cost-effective, is not very efficacious.

The student-authors did not recommend choosing among these alternatives, however, but recommended doing as many of them at one time as was feasible (“viable”). The trade-offs analysis would nonetheless permit decision-makers, and the public, to

prioritize which alternatives to emphasize in the likely case that priorities needed to be set. In their view, setting standards for energy efficiency in buildings was the first order of business.

## Establish Commensurability

Suppose some Alternative  $A_1$  stacks up very well on Criterion  $C_1$ , moderately well on  $C_2$ , and poorly on  $C_3$ . And suppose that  $A_2$  stacks up in the opposite way. We can choose between the two alternatives only if we can weight the importance of the criteria and if we can express their relative weights in units that are commensurable across the criteria. As you may have heard, money is everybody's favorite candidate for the commensurable metric. Using money as the metric is a very good idea, and it often works much better than you might imagine. For instance, even the "value" of life can sometimes be described reasonably well in the metric "willingness to pay X dollars for a reduction in the risk of death by Y percent a year," or something like it.

It is sometimes even possible, using money as a common metric, to compare apples and oranges, through the use of "willingness to pay" for hard-to-quantify outcomes like "better privacy protection" or "less-noisy motor scooters." This is the standard approach of benefit–cost analysis. How one does this is a very technical matter, and occasionally very controversial as well.<sup>51</sup>

In any case, a willingness-to-pay approach eventually runs into limits. To reach a summary judgment as to how much political equality to give up in a political redistricting case, for instance, in exchange for more African American voter power, it seems impossible even to state the trade-off in meaningful terms. In general, this problem is known as the "multiattribute problem." In some deep sense, the problem is logically insoluble, although some heuristics are available to help trim it down to its irreducible size.<sup>52</sup>

## Break-Even Analysis Revisited

We have seen how break-even analysis can help you both to focus on which residual uncertainties you will have to estimate and to frame the terms in which those estimates must be given (e.g., "We have to believe Alternative  $A_1$  will produce at least X results in order to justify choosing it"). We turn now to how break-even analysis can also help to solve commensurability problems.

Consider those policy areas, such as safety regulation, where we are often implicitly trading off dollars against risks to life. It might be supposed that in order to assess these proposals, you would have to decide what a human life is *really worth*—a task many of us, quite understandably, are unwilling to perform. The task is made somewhat more tractable, however, if you work with quantitative estimates and apply break-even analysis. Suppose, for instance, that you are considering whether or not to impose on the auto industry a new design standard that will improve safety and save an estimated twenty-five lives every year into the indefinite future. The cost of meeting the standard

is estimated at \$50 million per year indefinitely. The trade-off at the margin appears to be, therefore, “\$2 million per life.” But you don’t have to answer the question, “What’s a human life really worth?” in order to make at least some sense of this decision. You do have to answer the question, “Is a statistical life (that is, the life of an unknown individual ‘drawn’ in a random manner from some population, rather than a named person’s life) worth at least \$2 million?” That is a break-even analysis sort of question. For reasons best known to yourself, it may be obvious to you that a statistical life surely is—or isn’t—worth that much. And although it’s very difficult to decide whether the worth of a statistical life falls on one or the other side of some monetary boundary, it’s a lot less difficult than coming up with a point value.

Even this sort of trade-off calculation is troubling to many people, and some find it morally repugnant. Unfortunately, repugnant or not, it is in a sense inevitable. Whatever position you take on the auto safety design standard described, you are by implication also taking a position on the dollars/risk-to-life trade-off: If you favor the standard, you implicitly believe the trade-off is worthwhile, whereas if you oppose it, you don’t. Fortunately, this logical implication has its uses. You may in many circumstances quite sensibly prefer to rely on your intuition rather than on some complicated systematic method. Once you have reached your conclusion on that basis, though, you should check your intuition by asking yourself, “Since the implication of my policy choice is that I value X as being worth at least (or at most) thus-and-such, do I really believe that?”

### Frame Trade-Offs Crisply

That is one semantic strategy for thinking about trade-offs. We encountered another one on page X in discussing the weighting of criteria. When choosing between two railroad routes, we asked how heavily the decision-maker wished to weight the welfare of ten households forced out of their homes versus saving \$20 million on construction of the more accommodating route. We can now suggest another strategy for framing trade-offs crisply involving long division. It might help to think about the trade-off here in terms of an “average” individual family rather than the aggregate of ten households. Is it worth spending \$2 million to avoid removing a single family from their home?

But magnitudes are important as well. And once you have projected outcomes, you are in a better position to bring them into the thinking about trade-offs too. In the railroad routing example, would it make sense to approach the trade-off challenge by doing some long division? It might be helpful to think about a single “average” homeowner instead of the ten and ask whether it’s worth spending \$2 million (“on average”) to avoid imposing grief on this family. We should note, by the way, that, even though this is in one important respect a complex moral question, getting some numbers on the table is very helpful, even essential. Just as it is important to know the numbers twenty million, ten, and two million, we also need to remember the number three hundred thousand, the amount of compensation offered to offset at least some of the

average family's distress. It is furthermore necessary to remember that we are dealing with averages here for analytical convenience only but that in the actual situation some families will experience distress much greater than the average and some much less.

## Trade-Offs Are about Increments

The key to confronting trade-offs is to compare increments. "If we spend an extra X dollars for an extra unit of Service Y, we can get an extra Z units of good outcome." This kind of analysis puts the decision-maker in the position to answer the question, "Does society (or do you) value Z more or less than outcome X?" and then to follow the obvious implication of the answer: if yes, decide for another unit of Y; if no, don't.

The outcomes projections you have already done—the cells in your matrix—set you up to make these comparisons. All outcomes are expressed as increments or decrements with respect to some base case outcomes. From Table I-2 we learn that if we choose incentives for photovoltaic distribution (PV) over a policy of high-density residential development, we might expect about a 1.5 percent greater decrease in carbon emissions, but the cost per ton of those decreases will be higher by somewhere between \$841 and \$1,318. We also learn from Table I-2 that there are no trade-offs between the two policies in terms of the three "viability" dimensions, since they are rated the same on all.

The comparisons among increments are done essentially by subtraction. The values for high-density residential development have been subtracted from those for PV incentives. The residuals, following the subtraction, characterize the ways in which the two policies differ in their outcomes. One gets us more abatement, while the other is more cost-effective. These, of course, are only some of many trade-offs worth considering, others of which do not show up in Table I-2. Total cost, for instance, would probably be of interest, but one needs to go outside the table to get this information. Note what has dropped out of the discussion: the ways in which these two policies are similar—that is, accomplish the same goals. Presumably both are able to curb a substantial amount of carbon emissions, but this fact has disappeared from Table I-2 and therefore from the discussion of that table in this section. This is because Table I-2 has entered only information about how the alternatives considered improve upon (or fall short of) what would happen if the base case (called baseline by the student authors) were chosen. Is this disappearance helpful or not? It is not helpful if we want to be reminded of the overall stakes in the policy choice. But it is helpful if we want to focus, without distracting clutter, on the consequences of choice, which are the incremental differences between the outcomes projected from various alternatives.

Returning to our municipal library example, Table I-3 draws out what is at stake in the world for the alternatives under consideration. But it does not, by itself, help us to focus on the small pictures that show the incremental differences and, therefore, the trade-offs between the outcomes of the various alternatives.

To do that, Table I-4 needs to be created out of Table I-3. This can be done in the mind alone or, more reliably, with pencil and paper or on a spreadsheet program. Table I-4 focuses on the incremental differences between each of the alternatives and the base case. It focuses on how much the world would differ were we to choose a given alternative rather than to let the base case unfold. Essentially, it is created out of Table I-3 by subtracting the base case outcomes from the outcomes of each of the projected alternatives. The result tells you what you gain or lose on the several criterion dimensions by choosing a particular alternative over the base case.

But it does not tell you what you get by choosing one alternative over another alternative that is not the base case. In the example at hand, three such comparisons are possible. In a problem with, say, five alternatives other than the base case, the three become nine. You could make a separate table for each such comparison. But in most cases this would be wasteful. Instead, you could pick the most plausible comparison and discard the alternative that looks less good—and keep on doing this until you are confident that the best survives.

This procedure would be suitable, however, if we were to believe that the analyst's choices along the way would be the same as those of the client. However, the analyst cannot be sure this is true. The analyst's duty to the client would be to present a few (three or four?) alternatives to the client with the pros and cons of each, and some suggestions as to where the client should look for some key trade-offs of likely interest.

**TABLE I-4 ■ Incremental Comparisons between Each Alternative and the Base Case**

Alternatives/ outcomes	Additional books borrowed in outreach areas (per year)	Additional borrowers in outreach areas (per year)	Annual cost (\$ per year)	Net political payoff to council member Y*
Base case: continue present trends	0	0	0	0
Bookmobile biweekly each area	20,000	5,000	500,000	200; 150
Bookmobile once weekly only the poorer areas	15,000	3,000	350,000	100; 50
Book fairs twice yearly in 25 public schools	25,000	6,000	100,000	0; 0

\* Additional votes in next two elections minus votes lost

**Semantic Tip** A linguistic device to help you stay focused on the margin is frequent use of the word *extra*. Note that this word appears three times in the example analysis in the first paragraph of this section.

Some units of Service Y can be purchased only in “lumps” larger than one—sometimes much larger. Consider transportation services provided by highways and bridges. T might be one passenger trip from A to B, but most transportation construction projects (highway enlargements, new bridge crossings) can be undertaken only for minimum bundles of T that run into the thousands of trips. Or suppose that a police chief must choose one of two “lumpy” alternatives such as \$1 million per year for more overtime on the night shift or \$250,000 per year for more rapid replacement of police cars. The first alternative is lumpy because the police union insists on a minimum overtime rate for all 150 officers on the shift, and the second is lumpy because the auto supplier charges much less per vehicle after some threshold number of vehicles. If, say, the projected decrease in burglaries from increased overtime is 200 per year and that from newer vehicles is 50, the trade-off confronting the decision-maker at the margin is an extra \$5,000 per extra burglary prevented. In this case, the margin is a lumpy 150 burglaries and \$750,000. (Criteria other than burglary prevention and cost efficiency would, of course, be relevant to this problem.)

### The Better and the Worse

Trade-offs that are quantified are more useful than trade-offs that are not quantified. But quantification is frequently not possible. Suppose, for instance, that the board of a local community foundation wants advice about how to evaluate grant applications for a social services activity, and you envision these alternatives: invest more in getting to know the human capacities in the applicant organization, collect better data about supposed outcomes of the activity, and seek the advice of consultants who are expert in the activity in question. The three alternatives differ in their strengths and weaknesses. Hence there are trade-offs among them. Although it is hard or impossible to quantify these trade-offs, it is still possible to rank-order the three alternatives. The trade-offs then become implicit rather than explicit. But that is better than not confronting them at all.

Rank-ordering is especially useful when you face an uncertain “budget” constraint. This budget can be in money or in personnel time or in expected administrative hassle or in political favors that need to be called in—practically anything, that is, that is important to policy execution and is in limited supply. The uncertainty is about what exactly the limits are. Does the board of the community foundation wish to allocate more money or more time, or more of each, to evaluating grant applications? The policy analyst does not know in advance, and probably the board does not know either. Rank-ordering of alternatives tells decision-makers, in effect, “Start with the ones at the top, and keep going down until you run out of whatever you think is the relevant budget.”

## Trade-Offs and the Theory of the “Second Best”

A special type of trade-off arises in situations in which there are multiple failures in a market.<sup>53</sup> For example, suppose there is only a single firm that produces widgets in a region. As a monopolist, the widget company has an incentive to hold down widget production to raise prices and increase profits. A conventional solution would be to break up the monopolist to compel a higher level of production, thereby lowering widget prices and increasing consumer welfare. But now suppose the company is creating negative externalities by discharging the waste it generates from widget production into a river. If the monopolist is broken up, widget production will rise, but so too will the level of water pollution. In sum, there is a trade-off between the costs and benefits of increased widget production. The “second best” policy is the one that “optimally balances” this trade-off (which will involve lower levels of widget production than would occur in a perfectly competitive market).<sup>54</sup>

It is regrettable that economists use the term *second best* to describe the policy that produces the best achievable outcome for society in situations such as this. Policy analysts should view second-best policies in a positive light. Second-best policies are in fact the policies that maximize overall social welfare, taking into account the complexities of the situation at hand.

## Trade-Offs for the Last Mile

The ideal balance for efficacy-versus-accountability trade-offs is having the last unit of each value equal the last unit of the other. (If the values are not equal, this means that overall value could be increased by shifting the balance a bit.) This is a useless abstraction if taken too seriously, but it does capture the essential spirit of the process. In this spirit we estimate that accountability, roughly and on the whole, gets overweighted relative to efficacy. To oversimplify somewhat, this occurs because psychological and career consequences for stepping out of bounds or committing a visible mistake are disproportionately high. Enemies wait in ambush and are always ready to cry “Gotcha.”

During the implementation process, more trade-offs occur. They typically involve goals, delay, cost overruns, perversion of the program by rent-seekers, and risks of waste, fraud, and abuse. These trade-offs emerge in large part from bargaining and adaptations by important actors trying to obtain (or save) what they could not achieve during the policy adoption process.

Policy designers should mainly worry that goals will be sacrificed in the interests of shortening delay, saving money, and reducing career-related risks of being held responsible for real or perceived inefficiencies and blunders. They should have such worries even though they cannot foresee exactly how they might materialize. The solution is to try to design the policy so that its goals will stand up robustly against the many forces that might diminish them in favor of reducing delay, avoiding cost overruns, and creating rents.

## STEP SEVEN: STOP, FOCUS, NARROW, DEEPEN, DECIDE!

Up to this point, progress on the Eightfold Path has mainly bred expansion: of problem elements, alternatives, and criteria. It may also have bred an undesirable formalism, such that lists of these items may have come to have a life of their own. The outcomes matrix, which ideally would have served as a sort of “rough draft with attitude,” may have displaced the problem with which the project began. But the object of all your analytic effort should not be merely to present the client with a list of well-worked-out options. It should be to ensure that at least one of them—and more than one, if possible—would be an excellent choice to take aim at solving, or mitigating, the problem.

At a minimum, this need to focus, narrow, and deepen your analysis of the most promising alternative(s) means that you must think very seriously about (1) the politics of getting this alternative legitimated and adopted and (2) the design of the ongoing institutional features that will have the power and resources to implement the policy or program in the long run.<sup>55</sup>

At Step Seven, it is useful to remind yourself that the Eightfold Path is an *iterative* process. Before finalizing your analysis, pause, take stock, look at the big picture, review what you’ve done, and make any changes. The pitfalls of working through a problem are numerous, and even experienced policy analysts can get it wrong. Box I-8 presents a selective list of pitfalls for each step of the Eightfold Path (along with semantic remedies). These are some of the most common and treacherous pitfalls in our experience, but we invite you to add your own to this list—and let us know if you have better candidates for the next edition.

As another check on whether you have done your job well to this point, even though you personally may not be the decision-maker, you should now pretend that you are. Then, decide what to do, based on your own analysis. If you find this decision difficult or troublesome, the reason may be that you have not clarified the trade-offs sufficiently, or that you have not thought quite enough about the political barriers to adoption or probability of serious implementation problems emerging (or not emerging), or that a crucial cost estimate is still too fuzzy and uncertain, or that you have not approximated carefully enough the elasticity of some important demand curve, and so on.

Think of it this way: unless you can convince *yourself* of the plausibility of some course of action, you probably won’t be able to convince your client—and rightly so.

Of course, when you tell your story to your client or any other audience, you may not think it appropriate to make reference to your own decision. You may choose, instead, to simply limit your story to a clarification of the relevant trade-offs and leave the decision completely up to the audience.

Box I-9 displays a portion of a report from the Legislative Analyst’s Office (LAO) in California that encapsulates many of the key elements of this step. The excerpted section organizes and clarifies the good and the bad, the advantages and the disadvantages of the options, and also explains what alternatives are not recommended. Doing Step Seven well sets up the analyst to move easily to the final step, telling your story.

## BOX I-8: THE EIGHTFOLD PATH: PITFALLS AND SEMANTIC REMEDIES

**Define the Problem.** *Pitfall:* unwittingly smuggling a solution into the problem definition. *(Semantic) remedy:* “Our problem is there is too little [too much] X.” Or: “Our problem is that X is growing too fast [too slowly].”

**Assemble Some Evidence.** *Pitfall:* data for their own sake—that is, spending time collecting a lot of data without sufficient attention as to whether the data can be transformed into information and the information into evidence. *Corollary pitfall:* ending up without the evidence you really want. *Remedy:* “This is evidence for the important idea that . . .”

**Construct the Alternatives.** *Pitfall:* too vague specification of what the alternative really is; not behavioral or concrete enough. *Remedy:* “If we do this, next Monday morning Josephine and Roger should . . .”

**Select the Criteria.** *Pitfall:* they apply to the outcome, not to the alternatives. *Remedy:* use “maximize” or “minimize” or “Do enough to . . .”

**Project the Outcomes.** *Pitfall:* ignoring uncertainty. *Remedy:* “The odds this will happen are . . .” Or: “Here is the likely range of possibilities, with the most likely being in the middle.” (The latter framing is good if you want to rid yourself of thinking about low probabilities, as the odds of the whole range encompassing the outcomes would be very high, even if any particular outcome’s odds would be low.)

**Confront the Trade-Offs.** *Pitfall:* not focusing on the margin. *Remedy:* “The gain of extra X is worth giving up extra Y.”

**Stop, Focus, Narrow, Deepen, Decide.** *Pitfall:* shifting the burdens of your own vague analysis to the client or someone else. *Remedy:* “If I were making this decision by myself, here is exactly what I would do . . .”

**Tell Your Story.** *Pitfall:* vagueness caused by pulling punches, and using euphemisms and circumlocutions. *Remedy:* “We have this problem, and here is what I think we should do about it, though admittedly there are many challenges along the way . . .”

## BOX I-9: AN ILLUSTRATIVE EXAMPLE OF “STOP, FOCUS, NARROW, DEEPEN, DECIDE!” FROM A POLICY ANALYSIS REPORT

### Healthy Families Program

#### Background

**Program Draws Down Federal Matching Funds.** *The federal Balanced Budget Act of 1997 (BBA) made available approximately \$40 billion in federal funds over ten years to states to expand health care coverage for children under the State Children’s Health Insurance Program (SCHIP). The BBA also provided states with an enhanced federal match as a financial incentive to cover children in families with incomes above the previous limits of their Medicaid programs. Under SCHIP, the federal government provides states with flexibility in designing a program.*

California decided in 1997 to use its approximately \$4.5 billion share of SCHIP funding to implement the state's Healthy Families Program. Funding for the program generally is on a 2-to-1 federal/state matching basis. Families pay a relatively low monthly premium and can choose from a selection of managed care plans for their children. Coverage is similar to that offered to state employees and includes dental, vision, and basic mental health care benefits. The Healthy Families Program also covers more intensive mental health services for children with serious emotional disturbances, which are directly provided through county mental health systems and supported primarily with county and federal funding. . . .

## Enrollment Cap Proposal Raises Policy Concerns

The Governor's budget proposal to cap Healthy Families Program enrollment, while feasible and effective in addressing the state's fiscal problems, raises a number of issues. We recommend against this approach because other alternatives are available to the Legislature to hold down the cost of the Healthy Families Program. . . .

## Governor's Proposal Has Some Advantages

**Savings Would Be Realized.** The overall administration proposal to cap health and social services program caseloads is discussed generally in the "Crosscutting Issues" section of this chapter. Policy issues of particular importance to the Healthy Families Program are discussed below.

Our analysis of the Governor's proposal indicates that it is technically feasible and would probably generate program savings of the magnitude estimated by the administration. Assuming the cap were maintained, the amount of savings achieved from a freeze on enrollment would grow significantly over time and contribute to addressing the state's structural imbalance between revenues and expenditures.

The administration's approach would also be less disruptive to the ongoing operation of the program than other possible approaches for achieving savings. No child now receiving coverage through the Healthy Families Program would lose his or her benefits. It is also possible that the prospect of long waiting lists would provide additional incentive for parents of Healthy Families children to become more diligent about submitting annual eligibility documents in a timely fashion, and reduce the high rate of disenrollment of children from the program.

## Several Issues Warrant Consideration

The Governor's proposal to cap program enrollment in Healthy Families (as well as comparable caps on other health and social services programs) raises a number of significant policy issues that the Legislature may wish to consider.

**Waiting Lists Could Create Inequities.** The administration's proposal raises some distinct equity issues. First, children who entered the program before January 1, 2004 would be treated differently than children who applied after that date even though they met the same eligibility criteria. Also, the administration proposal is for a first-come, first-served approach in which the first person on a waiting list would be added to the Healthy Families Program caseload as children were disenrolled and "room" was created for additional children on program rolls. While this approach is equitable—all children on the waiting list would be treated alike—it also raises other questions of

fairness, in that children would be added to program enrollment in the future regardless of a child's medical needs or family income level. . . .

**Time on Waiting List May Be Underestimated.** Another concern is that the waiting time for an applicant to actually receive health coverage could turn out to be longer than the maximum of six months estimated by the administration. That estimate is based on current disenrollment and enrollment trends. To the extent that parents' behavior changed, as discussed above, so that disenrollment rates in the program decreased, the waiting period for coverage could be longer than projected. As noted earlier, the waiting period for enrollees would be likely to exceed one year by June 2006. . . .

**State Would Lose Additional SCHIP Funds.** The proposal to cap enrollment in the Healthy Families Program would result in state savings, but also reduce by about \$55 million the amount of federal SCHIP funds being drawn down for health coverage of the uninsured. Since the inception of the Healthy Families Program, California has struggled to fully utilize its federal allotment of SCHIP funds. To date, the state has reverted \$1.1 billion in unspent funds back to the federal government, which was redistributed to other states that were able to expend their allotment within the specified time period. As of May 2003, California had approximately \$1.9 billion in unspent SCHIP funds remaining. We would acknowledge, however, that some other strategies for containing state costs for Healthy Families coverage would also add to the amount of SCHIP funds that would go unspent.

**Some Children Would Lose Insurance Coverage.** The Healthy Families Program was established to operate in tandem with Medi-Cal to ensure seamless health care coverage for children ages 0 to 19 living in families earning up to 250 percent of the FPL [federal poverty level]. Due to the income and age-based eligibility structure for both programs, the proposed enrollment cap would place certain children who were enrolled in Medi-Cal at risk of losing insurance coverage. Specifically, upon reaching their first and sixth birthday, children who would traditionally transition to the Healthy Families Program because their families' incomes would no longer qualify them for Medi-Cal would instead be placed on a waiting list for coverage.

## Analyst's Recommendation

**Other Alternatives Available . . .** After weighing the advantages of imposing an enrollment cap on Healthy Families against the issues discussed above, we recommend against the Governor's proposal because, in our view, other alternatives are available to the Legislature to hold down the cost of the Healthy Families Program. As we will discuss later in this analysis, we believe there are other strategies that could be adopted to reduce program spending that would be more equitable to beneficiaries, more consistent with other state efforts to assist the uninsured, and that would make more effective use of the available federal SCHIP funds.

**. . . But if Proposal Is Adopted.** Should the Legislature decide to adopt the Governor's proposal, there are several steps it could take to address some of the issues we have outlined. In that event, we would recommend that the Legislature consider the following actions:

- Modify the first-come, first-served approach to prioritize for Healthy Families coverage the poorest eligible children, and-or those with the most significant medical needs. These actions would partly reduce the savings but ensure that state funds are used for those who are most needy.

- Modify the CHIM [County Health Initiative Matching Fund] program to allow coverage of individuals otherwise eligible for Healthy Families but placed on a waiting list. This could address the inequity by which CHIM children in families with higher incomes would receive coverage quickly, while those in families with lower incomes would remain on waiting lists.
- Adopt supplemental report language directing MRMIB [Managed Risk Medical Insurance Board] to provide the Legislature with a quarterly report providing a statistical summary of the number of children placed on waiting lists, the period of time applicants must wait for coverage, and the effect of waiting lists on program enrollment rates. This information would enable the Legislature to assess the impact of the enrollment caps upon their implementation.
- Direct MRMIB to report at budget hearings on how conflicts with the CHDP [Child Health and Disability Prevention Program] gateway, parent expansion of Healthy Families, and SB 2 [Chapter 673, Statutes of 2003] should be addressed.

Source: *Analysis of the 2004–05 Budget Bill* (Sacramento, CA: Legislative Analyst’s Office, February 2004), [https://lao.ca.gov/analysis\\_2004/health\\_ss/hss\\_12\\_4280\\_an104.htm#\\_Toc64277937](https://lao.ca.gov/analysis_2004/health_ss/hss_12_4280_an104.htm#_Toc64277937).

## Apply the Twenty-Dollar-Bill Test

Before finalizing your decision, you should subject your favored policy alternative to the *twenty-dollar-bill test*, a good final check that your idea is indeed solid. The name of this test is based on an old joke that makes fun of economists. Two friends are walking down the street when one stops to pick something up. “What about that—a twenty-dollar bill!” he says. “Can’t be,” says the other, an economist. “If it were, somebody would have picked it up already.” The analogy is this: *If your favorite policy alternative is such a great idea, how come it’s not happening already? Why hasn’t the proposal been enacted?* The most common sources of failure on this test are neglecting to consider the resistance of interest groups, bureaucratic and other stakeholders in the status quo, and the lack of an entrepreneur in the relevant policy environment who has the incentive to pick up what seems like a great idea, win political credit for taking an agreeable stance, and see it through. Failure on this test is not fatal, of course. You might keep fiddling until you invent a variant of your basic idea that will pass.

## A Focus for the Last Mile

It is useful to focus—hopefully, not for the first time—on protecting goals and avoiding excessive delay. Make sure you have built into the implementation process targets and deadlines, numerical if at all possible. For instance: “By [date] we will have reached ten thousand children statewide with at least five of the following ten services, which on average will be of at least X quality.” There are of course political risks to falling short, especially in a process as nebulous as designing for the last mile.

## STEP EIGHT: TELL YOUR STORY

---

After many iterations of some or all of the steps recommended here—principally, redefining your problem, reconceptualizing your alternatives, reconsidering your criteria, reassessing your projections, and reevaluating the trade-offs—you are ready to tell your story to some audience. The audience may be your client, or it may include a broader aggregation of stakeholders and interested parties. It may be hostile, or it may be friendly. Your presentation may be a one-time-only telling, or it may be merely the first effort in a planned long-term campaign to gather support behind a legislative or executive change.

### Apply the Grandma Bessie Test

Before proceeding further, however, you need another little reality check. Suppose your Grandma Bessie, who is intelligent but not very sophisticated politically, asks you about your work. You say you are a “policy analyst working for . . .” She says, “What’s that?” You explain that you’ve been working on “the problem of . . .” She says, “So, what’s the answer?” You have one minute to offer a coherent, down-to-earth explanation before her eyes glaze over. If you feel yourself starting to hem and haw, you haven’t really understood your own conclusions at a deep enough level to make sense to others, and probably not to yourself, either. Back to the drawing board until you get it straight.

Now consider the possibility that someone might actually wish to base a real decision or a policy proposal on your analysis. (It’s been known to happen.) Even if you, as an analyst, would not have to deal directly with such a tough audience as Grandma Bessie and her kin (including, of course, Grandpa Max), it’s likely that someone will have to do so. At the very least, therefore, you’ll have to be able to explain your basic story to someone in sufficiently simple and down-to-earth terms that that someone will be able to carry on with the task of public, democratic education.<sup>56</sup>

### Gauge Your Audience(s)

Assuming that you’ve passed the Grandma Bessie test, identify and assess the likely audience(s) that are more sophisticated and involved than Grandma Bessie.

First comes your client, the person or persons whose approval you need most—your hierarchical superior(s), perhaps, or those who are funding your work. What is the relationship between you and your client? What you say and how you say it should depend a great deal on whether your relationship is long-term and on whether it is carried on face-to-face. In particular, how easy will it be for you to correct any misunderstandings that may arise?

Next, think about the larger political environment. Who do you think will “use” the analysis and for what purpose(s)? Will anyone pick up your results for use in an advocacy context? Would you regard this use of your results as desirable? Or desirable if certain advocates use your work and undesirable if others do so? Do you want to do

anything to segregate the elements of your analysis by the type of audience you might want it to reach—or not reach? Are you, perhaps inadvertently, using scare words that will alienate certain audiences who might otherwise be open to your analysis?

If one of your goals is to engage a lay audience, keep in mind that ordinary folk are rarely moved by statistics alone. Indeed, relying on numbers to demonstrate the importance of addressing a problem can actually undermine the psychological processes needed to prompt a response; people may not only fail to grasp the statistics, but they may be numbed into inaction.<sup>57</sup> Data and statistics are obviously indispensable to analysis, but when it comes to telling your story to a general audience, be sure to put a human face on the problem. And show how your solution could make life better for real people.

If you are making a clear recommendation, make sure that you raise and rebut possible objections to it that might occur to various important audiences. Also, make sure that you compare it to what you or others might regard as the next best course of action, so as to be ready to show why yours is better.

## Consider What Medium to Use

You can tell your story in written or oral form. In either case, communicate simply and clearly. The guiding principle is that, other things being equal, shorter is always better. In written presentations, good subheadings and graphics can make reading and comprehension easier.

Anecdotes carry a lot of weight (indeed, more than they deserve to). Nevertheless, they are much, much more effective at seizing attention and increasing credibility than pure exposition—never mind facts and statistics. They are especially effective when they relate to the personal experiences and friendships of members of your audience. Oral presentations are more effective as vehicles for these all-important anecdotes, although combining the oral and written modes is probably better than either alone.

Oral presentations require practice, self-discipline, and a little knowledge of some basic principles. The most basic of the basic principles are these: Speak *very slowly and distinctly*; speak loudly enough to be heard throughout the room, even over distracting noises; speak in a lower register, which tends to increase perceived trustworthiness and credibility; do not fidget, but don't stand like a stick, either; make lots of eye contact with audience members and, in doing so, don't favor one side of the room over another. Speaking slowly and distinctly is probably harder than you think—and more important, too. Visual aids such as PowerPoint slideshows often help in oral presentations.

An increasingly common medium for telling your story is the issue brief. The best issue briefs are short, crisp, and visually attractive. Issue briefs can be stand-alone documents, or they can be supplements to traditional written or oral presentations. Box I-10 shows an example of an issue brief from Research Improving People's Lives (RIPL). Note that the issue brief proposes the implementation of a randomized

controlled trial, a rigorous scientific method to measure policy impact and effectiveness (see Appendix F). The issue brief is also an example of the “learning by doing” variant of multistage analysis discussed earlier in Part I.

## Give Your Story a Logical Narrative Flow

Your story’s flow should be designed with the reader’s (or listener’s) needs and interests and abilities in mind. In both written and oral presentations, it should be evident to the audience what motivates the entire analysis. Therefore, it is best to open with a statement of the problem your analysis addresses.<sup>58</sup>

It is also important to *motivate* the more detailed steps in the flow of the analysis—that is, the sections, paragraphs, and sentences. Most readers will look for the motivation of any element in what immediately precedes it, which makes it important to avoid lengthy digressions. For these reasons, be wary of sections that you are tempted to label “Background.” Similarly, the phrases “Before turning to . . .” and “It is first necessary to explain/understand the history of . . .” are usually signs of undigested material. Many readers will be alert to these danger signs, so you should be, too. Policy analysis, remember, is about the future. Perhaps surprisingly—it is often not obvious how, or whether—history affects the future, but the burden should be on the writer or speaker to show exactly how this effect will come about.

A common, though not uniformly applicable, organizing framework is to begin with a good problem definition and then to treat each alternative you consider as a major section. Within each such section, you project the probable outcome(s) of implementing the alternative and assess how likely such outcome(s) are in the light of some causal model and associated evidence. Following these discussions, you review and summarize the alternative outcomes and discuss their trade-offs. This framework contains no special discussion of criteria; however, sometimes an explicit discussion of criteria is important. If so, it might appear either just before or just after the presentation of the alternatives and their associated outcomes.

Do not be afraid to start with a recommendation (if you intend to make one) and an assertion that, now that you have put it out, you intend to present all the necessary steps to justify it. You can often help the reader by providing a simplified, stylized account of a topic and then complicating it with additional details. You will find yourself writing (or uttering) sentences like “That is how a subsidy strategy will work if it works perfectly, but now we need to introduce at least two sources of friction. One has to do with the likelihood that 5 to 10 percent of the claimants will probably be ineligible, and a majority of those may be fraudulent as well. The second source involves finding competent and willing partners in the nongovernmental organization (NGO) community to help with the outreach component.” These sentences would prepare the reader for discussions of the ineligibility problem relative to the established scenario for working “perfectly,” the political backlash if fraud were to become a big issue, and the complexities of partnering with NGOs.

# BOX I-10: AN ILLUSTRATIVE EXAMPLE OF “TELLING YOUR STORY” FROM A POLICY ANALYSIS REPORT

**RIPL**  
RESEARCH IMPROVING PEOPLES’ LIVES

INITIATIVE THREE  
CRIMINAL JUSTICE

## Thirty percent of Rhode Island inmates return to prison after one year. Can Supplemental Nutrition Assistance Program (SNAP) help reduce recidivism?

### Strategic Goal

Governor Gina Raimondo wants to reduce the three-year recidivism rate from 52% to 44% by 2020. What are some low-cost ways to meet this goal?

### Assessment

RIPL analysis shows that individuals who enroll in SNAP post-release are 7.4 percentage points less likely to recidivate than their non-SNAP counterparts within the first six months. However, only 40% of individuals enroll in SNAP post-release even though nearly all qualify.

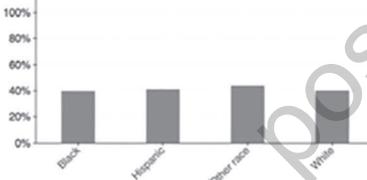


Figure 1: Percent of individuals who enroll in SNAP within three months post-incarceration

### Understanding the Challenges

Currently, inmates may receive help to complete a SNAP application while incarcerated but, once released, they must visit a Rhode Island Department of Human Services (DHS) office to complete their enrollment and receive a benefits card (EBT card). However, the myriad challenges individuals face upon release, lack of accessible transportation and wait times may delay or derail enrollment.

RIPL, in collaboration with the Rhode Island Department of Corrections (DOC) and DHS, developed Connect for Success to ensure all inmates have activated SNAP benefits cards post-release. We began by conducting field work to design a likely-successful program:

- We designed a mock-pilot to mimic SNAP enrollment and benefit distribution.
- We surveyed over 100 inmates right before release, and then tested ways to distribute gift cards as mock EBT cards.
- Our survey results showed that 70 percent of prisoners plan to enroll in SNAP upon release; nearly double the number who end up enrolling.

- We found that individuals are anxious to leave, and may miss cards if they are not distributed with exit materials.
- When offered the opportunity to get a \$5 gift card from a nearby building upon their release, only 4% did so.

### Design and Test Improvements

Based on our field work, Connect for Success follows the process below to maximize connecting releasees with benefits.

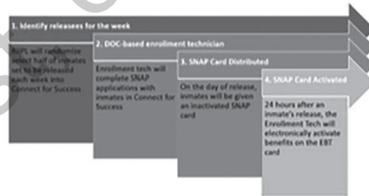


Figure 2: Connect for Success Process Map

DOC will run a RIPL-designed Randomized Controlled Trial (RCT) of Connect for Success in fall 2017. Why an RCT? First, an RCT allows us to prove impact and measure success. Second, it allows us to rule out any negative, unintended consequences before taking a program to full-scale. Third, it allows others to learn from our trial and scale a proven-successful program in their own state.

### Conclusion

Through Connect for Success, we hope to lower recidivism rates by three percentage points – or 30% of the Governor’s goal, at almost no cost. Fact-based policy is a powerful tool for alleviating poverty and improving policy effectiveness efficiently.

WWW.RIPL.ORG

Source: Research Improving People’s Lives, “Reducing Recidivism by Connecting Releasees with Social Services,” 2018, [https://www.ripl.org/wp-content/uploads/2018/08/RIPL\\_i3\\_Reducing-Recidivism-by-Connecting-Releasees-with-Social-Services.pdf](https://www.ripl.org/wp-content/uploads/2018/08/RIPL_i3_Reducing-Recidivism-by-Connecting-Releasees-with-Social-Services.pdf). Reproduced by permission.

## Some Common Pitfalls

### Following the Eightfold Path Too Closely

Occasionally, the Eightfold Path helps to structure your narrative flow as though you were leading the reader by the hand down its course. But this approach is almost always a really bad idea. It leads to a wordy, mechanistic product that repels rather than attracts the reader's attention. The purpose of the Eightfold Path, remember, is to help you *think through* a complicated problem. It is not necessary to use it in telling the story. Don't tell the reader about all the alternatives not taken and the many reasons why they were not. That's *much more* than the reader wants or needs to know. The reader should or would probably be interested in the weightiest reasons why one or two of the next-best alternatives are not recommended in preference to the one(s) that is (are)—but no more than that.

### Compulsive Qualifying

Don't interrupt the flow of an argument in order to display all the qualifications and uncertainties about some particular element in the argument. A linguistic way around this pitfall is to use adjectives or adjective phrases, such as *most*, *on average*, and *more often than not*, to state the generality and then to return to the exceptions in the next section. (Or, if the exceptions and qualifications really can't wait, try a parenthetical sentence or a footnote.)

### Showing Off All Your Work

Don't include every fact you ever learned in the course of your research. Even if you've done a good and thorough job of research and analysis, most of what you have learned will prove to be irrelevant by the time you're finished. That is, you will have succeeded in focusing your own attention on what is really important and in downplaying what only appeared important at the beginning. You don't usually need to take your reader on the same wandering course you were obliged to follow.

### Listing Without Explaining

Should you list every alternative policy that you intend to analyze in the report before you actually get around to providing the analysis? Such a list is a good thing when the alternatives are not numerous, when they are all taken seriously either by you or by your audience, and when they will prepare the reader's mind for the detailed assessment that will follow. However, if you have many alternatives to consider, the reader will forget what's on the list, and if some of the alternatives turn out to be easily dismissed upon closer scrutiny, you'll simply have been setting up straw men and wasting the reader's mental energy.

Similarly, be cautious about listing every evaluative criterion of interest before coming to the assessment of the alternatives being considered. Usually—though not

always—not much can be said in a separate section about criteria that can't be better said when you're actually writing the assessment sections.

### **Spinning a Mystery Yarn**

Start with the conclusion, the bottom line, the absolutely most interesting point you intend to make. Then present all the reasoning and evidence that you have to make your audience reach the same conclusions you have reached. In short, follow the opposite strategy from that which a novelist would follow.

### **Inflating the Style**

Avoid the pomposity and circumlocutions of the bureaucratic and the academic styles. (Essential reading: George Orwell, “Politics and the English Language.”) Also avoid a chatty, insider's style—such as “We all understand what fools our opponents are, don't we?”

### **Forgetting That Analysis Doesn't Persuade—Analysts Do**

No matter how competent your analysis, it will be only as persuasive and credible as the person or organization who produces and communicates it. For oral presentations, the quality of public speaking matters; if you haven't taken a public speaking class, consider taking one. For presentations that will be televised or webcast, there are (nonobvious) dos and don'ts about how to look professional on camera; learn them.

Finally, remember that the persuasiveness and credibility of an analysis depends in no small part on the personal and organizational reputation of the analysts who produced it. If you or your organization produce a shoddy study, or appear unprofessional when you disseminate its results, the reputational costs may be with you for a long time. When you communicate an analysis, remember that you are not simply sharing your solution with the audience you are presently speaking or writing to—you are also building personal, political, and organizational capital to put yourself in the position to communicate other solutions—maybe even more important solutions—in the future. The LAO report excerpted in Box 1-9 on the Healthy Families Program establishes credibility by acknowledging that a proposal it recommends against has virtues. Such credibility makes it more likely that the audience will be receptive to the analyst's story about why the report's preferred alternative is superior.

### **Structure Your Report**

Unless the report is short, begin with an executive summary.

If your report is over fifteen to twenty pages long, say, a table of contents may well be helpful. If there are many tables and figures, either in the text or in the appendices, a list of these items can be helpful, as well. Detailed technical information or calculations should appear in appendices rather than in the text. However, enough

technical information, and reasoning, should appear in the text itself to persuade the reader that you really do know what you're talking about and that your argument is at least credible.

Use headings and subheadings to keep the reader oriented and to break up large bodies of text; make sure your formatting (capital letters, italics, boldface, indentation) is compatible with, and indeed supports, the logical hierarchy of your argument.

### Table Format

Current professional practice is very poor with respect to the formatting of tables. Do not imitate it but strive to improve it. Every table (or figure) should have a number (Table 1, for instance, or Figure 3-A) and a title. The title should be intelligible; it is often useful to have the title describe the main point to be learned from the table (e.g., "Actual Risks of Drinking and Driving Rise Rapidly with Number of Alcoholic Drinks—but Are Greatly Underestimated by College Students"). Each row and column in a table must be labeled, and the label should be interpretable without too much difficulty.

Normally, a table either is purely descriptive or is designed to demonstrate some causal relationship. In the latter case, it is usually desirable to create a table that makes a single point (or at most two) and that can stand alone without need of much explanation in the surrounding text. It is usually better to use two or three small tables to make two or three points than to construct one massive table and then try to explain its contents by means of the text that surrounds it.

Tables usually require footnotes, and there should almost always be a source note at the bottom. Sometimes these notes refer to data sources used to make the table, and sometimes they attempt to clarify the meaning of the row or column labels, which are necessarily abbreviated.

Please do *not* imitate academic practice, which is to overstuff tables with all kinds of numbers and to mindlessly apply obscure column and row labels. Academic practice presupposes that all the data have been gathered "scientifically" and without serious bias; therefore, the presentation style aims to convey these facts. Unimportant data share space in academic tables with important data so as to permit the reader to see that the complete truth has been told, or that the author has not cherry-picked the data to convey only what is interesting and has conveyed the full story about what is statistically significant as well as what is not. If these issues are important to you and your readers, by all means provide the full story. But do it in appendices. In most cases, though, try to minimize the information provided in a single table.

### Statistics

Your audience probably does not understand statistics as well as you do, so keep your statistics few and simple. Percentages are good, and differences in percentages even better—for example, "The food budget for juvenile facilities serving boys (girls) is

10 percent higher per capita than that for facilities serving girls (boys).” If regression coefficients must be used, make sure the raw coefficients correspond to intelligible real-world phenomena rather than to mere index numbers that researchers have found useful. Intelligibility, moreover, always implies using metrics that are meaningful to your audience; for example, “A ten-cent tax on high-sugar foods would probably reduce per-capita consumption by 6 to 10 ounces per week.” In this case, ounces per week is better than pounds per year or ounces per day, since yearly consumption means nothing to most people and most people probably suppose their daily consumption is highly variable whereas weekly intake smooths out the daily unevenness.<sup>59</sup>

## References and Sources

Include a listing of references and sources at the end of the presentation. Books and articles should be cited in academic style (alphabetical order by author). The main point is to provide bibliographic help to curious or skeptical readers who want to track down references for themselves. There are several acceptable styles, but a good model is the one used in the book review section of the *Journal of Policy Analysis and Management*, which is simple and direct.

The current trend is toward “scientific citation” in lieu of footnote references in the text. That is, cite the author’s last name and year of publication in parentheses in the text; the reader then consults the references section at the end for the full citation. If you follow this practice, the reference section should list the author(s) *before* the title of the work and other publication details, including the year. Sometimes you will want to include a page number in the parenthetical citation, as well.

Legal citation style is quite different. If most of the references are legal, then it is advisable to cite all references in bottom-of-page footnotes. However, you can keep the scientific citation format within the footnotes.

Notes are easier to read if they appear on the same page as the referenced text—that is, if you display them as footnotes rather than as endnotes.

## Using a Memo Format

If your analysis is to be delivered in a memo, you should present it within a standard memo format, as follows:

[Date]

To: [Recipient name(s), official position(s)]

From: [Your name, position. Sign or initial next to or above your name.]

Subject: [Brief and grammatically correct description of the subject]

[The first sentence or two should remind the recipient of the fact that they asked you for a memo on this subject, and why. Alternatively, you could explain why you are submitting this memo on this subject to the recipient at this time.]

[If the memo is long, you might open and close with a summary paragraph or two. If you open with a long summary, the closing summary can be short.]

[If the memo is long, consider breaking it up with subheads.]

## Develop a Press Release

Most policy analyses do not become the subjects of press releases or of radio or television sound bites, but some do. Others become candidates for such treatment, and all can profit, even in their extended form, from the analyst's reflecting on how to condense the essential message. Hence, it will probably serve an analytic purpose—and sometimes a political one—if you sketch out a press release or a few ideas for sound bites. You may also want to think strategically and defensively to see how an opponent might characterize your work in a press release or sound bite.

## PowerPoint

PowerPoint slides frequently supplement oral presentations and indeed sometimes replace written reports altogether as nonverbal means of communication. Following are some brief comments on the use of PowerPoint; plenty of full-scale manuals are available.

- Keep it simple: have each slide present a separate point; use phrases, not sentences; and use only two or at most three colors.
- Choose text color and background color so that the text color is very legible on the background color.
- Avoid cutesy icons and “cool” moving animals.
- Think of the viewer's needs: to see letters and numbers at a fair distance, and to not be bored by having you as presenter simply read what is on the screen.
- Display the slide for long enough so that the viewer can actually read and absorb its contents—especially important for tables and graphs.
- Include slides at suitable intervals that summarize what has been said so far and point the way to what is yet to come.
- Make available to the audience, after the presentation—not during, as it is distracting—hard copies of slides (arranged six per page).

Visual supplements, such as photographs, can nicely support all the words, provided they are carefully chosen and displayed.

## The Last Mile

Telling your story is an opportunity not only to give a succinct elevator speech on the substance but also to add to the political momentum. Convey that “Something is going to happen.” The main point of this is to induce calculation that, since the train is leaving the station, your audience had best get aboard and maybe have some say in where it’s going and how.

### NOTES

1. See Roose (2018).
2. For an analysis of most traditional market failures in transaction cost terms, see Zerbe and McCurdy (1999), which also emphasizes the rich variety of interventions besides those undertaken by government to remedy traditionally conceived “market failures.”
3. For an excellent discussion, see Weimer and Vining (2017, chap. 6).
4. “Social capital” is a fuzzy idea, but a good one, referring to a real phenomenon, which is of course always changing and never the object of firm consensus.
5. However, see Herd and Moynihan (2019) on administrative burdens; on choice architecture and social norms, see Goldstein et al. (2008) and Thaler and Sunstein (2008).
6. See Nyhan et al. (2014).
7. See Oster (2018).
8. This happened to a graduate student group at the Goldman School whose client was the Oakland Police Department. Members of the group struggled hard to escape the initial assumptions held by their client and eventually to refocus their work.
9. Some analysts also claim that it is simply not worthwhile to define as “problems” conditions that cannot be ameliorated: “Problems are better treated as opportunities for improvement; defined problems, as problems of choice between alternative means to realize a given opportunity. The process of problem definition would then be one of search, creation, and initial examination of ideas for solution until a problem of choice is reached.” See Dery (1984, 27).
10. Social scientists today are increasingly using rigorous experimental and quasi-experimental methods to identify causal relationships between variables, including randomized controlled trials, difference-in-differences, instrumental variables, and regression discontinuity. There are growing calls for government to use such research to inform policy analysis and evaluation. See Haskins and Margolis (2014).
11. See Shipan and Volden (2021).
12. Congressional Budget Office, *The Army’s Ground Combat Vehicle and Alternatives* (report), April 2013.
13. These tips for finding creative solutions to problems are from Nalebuff and Ayres (2003). See also the valuable “nudge” approach in Thaler and Sunstein (2008).

14. See Adams et al. (2021).
15. For a good, brief discussion, see Stokey and Zeckhauser (1978) and Victorio (1995); also see the models, particularly that of case management, in Rosenthal (1982).
16. For other ideas and an excellent discussion of the uses of models generally, see Lave and March (1975).
17. Often, though not always, the basic element is something like a smart practice—that is, an intervention strategy that attempts to take advantage of some qualitative opportunity to create valued change at relatively low cost or risk. See Part IV, “‘Smart (Best) Practices’ Research.”
18. Choosing a numerical target can help to focus energies and can force you to think about what effects are too small to be worth seeking. But when all increments are of equal value, choosing a target may be arbitrary and self-defeating.
19. See Aizer et al. (2022).
20. Cost-effectiveness analysis is often used when the benefits of policy alternatives are difficult to monetize. For example, a Resources for the Future/National Energy Policy Institute study employed cost-effectiveness analysis to evaluate a carbon tax and cap-and-trade programs because of the difficulty of monetizing the benefits of reducing oil dependence and carbon dioxide emissions. See Krupnick et al. (2010).
21. For a provocative argument that cost-effectiveness analysis is a better technique for most public purposes than benefit–cost analysis because the former looks away from individual preferences toward collectively established objectives, see Moore (1995, 35–36).
22. See Moore (1995, 36).
23. For an excellent textbook on benefit–cost analysis, see Boardman et al. (2011).
24. See Karoly and Bigelow (2005, xiv).
25. We said earlier that criteria apply to outcomes and not to alternatives. However, this statement needs a slight amendment in the case of practical criteria, which apply not to outcomes but to the prospects an alternative faces as it goes through the policy adoption and implementation processes.
26. For an accessible introduction to this approach to the study of politics, see Shepsle and Bonchek (2010).
27. An analogous procedure was first given prominence by Graham Allison (1971).
28. On penalty defaults, see Mansbridge and Martin (2013).
29. On the implementation challenge of obtaining compliance among program targets, see Weaver (2010).
30. See Campbell (2012, 336).
31. See Patashnik (2008).
32. See Stokey and Zeckhauser (1978, chap. 11).
33. Also, in this case, the stem *respons-* appears in both alternative and criterion.
34. See p. 12, no. 4, at [www.whitehouse.gov/wp-content/uploads/2023/04/DraftCircularA-4.pdf](http://www.whitehouse.gov/wp-content/uploads/2023/04/DraftCircularA-4.pdf).

35. Congressional Budget Office, *Letter to the Honorable Frank D. Lucas*, September 16, 2013, <https://www.cbo.gov/sites/default/files/hr31020.pdf>.
36. This assumes that you do make a recommendation. But even if you only lay out options and attach projected outcomes to them, you still cannot escape justifying the projections.
37. Some people speak of “switchpoint analysis” and would refer to the 15 percent here as the “switchpoint” at which a decision-maker would switch from a favorable view of this policy to an unfavorable view or vice versa. Others refer to “threshold analysis” and would call the 15 percent figure the threshold level of effectiveness we would need to assume in order to justify choosing this alternative.
38. The example is discussed in Sunstein (2014, 75). Sunstein offers an excellent review of the uses of break-even analysis. Appendix D of that book provides a list of selected examples of break-even analysis carried out by the federal government.
39. A special case of break-even estimation is a fortiori estimation. If you hypothesize worst-case estimates of all important parameters that remain uncertain, and the policy alternative still satisfies your decision criterion, the alternative would, a fortiori, prove satisfactory even if more careful estimates were to be more favorable. In that case, the more careful estimates are unnecessary. See MacRae and Whittington (1997) on a fortiori analysis (218–219) and, more generally, on the question of precision versus approximation in projecting outcomes (209–224).
40. For further details, see Morgan and Henrion (1990, chap. 8). You can use the commercially available (and very user-friendly) Crystal Ball program to run Monte Carlo simulations.
41. Most of the work on this type of simulation has been done at the RAND Corporation. See Lempert et al. (2003).
42. On performing a postmortem analysis, see Klein (2007).
43. See Doleac and Hansen (2016).
44. Ibid.
45. See Doleac (2016).
46. See Weaver (2009).
47. See Cryan et al. (2008).
48. They grouped their eight alternatives into five subgroups, however, to simplify the analysis.
49. Even when one policy alternative dominates other options, opportunity costs still must be faced. The implementation of policies nearly always requires the use of some resources that could be used to produce other things of value.
50. Confronting the trade-offs may require the analyst to acknowledge that certain negative outcomes are deemed to be acceptable, and not deserving of much weight, even if some people might be unhappy about them. For example, in the vaccination example discussed in Step One, authorities had to genuinely sacrifice the preferences of anti-vaccination parents who believe (erroneously) that vaccination causes autism. The analyst is thus deciding that the public’s beliefs about vaccination are misguided and should count much less than the recommendations of scientific

experts. Similarly, policy analysts might decide on the basis of science to regard climate change as a serious problem, even if this means ignoring the preferences of citizens who believe that climate change is a hoax.

51. Good discussions can be found in Adler and Posner (2001).
52. See Stokey and Zeckhauser (1978, 117–133) and MacRae and Whittington (1997, 201–203). One potentially misleading heuristic has the analyst creating a score for each alternative with respect to each criterion and then manipulating the scores arithmetically. It is easy to get the arithmetic right, but it is often hard to come up with scoring procedures that are not at some level arbitrary (e.g., anchored against some arbitrarily defined level of excellence or its opposite).
53. This section draws on the excellent discussion in Bueno de Mesquita (2016, 125).
54. *Ibid.*
55. For reasons of space, we do not discuss the first of these matters here, but see Appendix C for a very brief survey of pertinent institutional issues.
56. Sometimes this is referred to as the challenge of giving an “elevator speech.” You and your boss, or some relevant other, find yourselves together in an elevator for too long a time to make do with just “Hi, how are ya?” The boss asks how your project is going. You have maybe a minute to explain what you’re up to and why people should be interested and perhaps persuaded. So have your elevator speech committed to memory and ready to go at a moment’s notice.
57. See Slovic (2007).
58. An unusually fine manual on how to give slide-based oral briefings is published by the RAND Corporation (1994).
59. See Browner (2012). Although this is directed to academic researchers in medicine, much of the advice can easily be extrapolated to policy analysis.