

The Limits of Social Science in Guiding Policy

Mark H. Moore

October, 2002

(A "Reaction Essay" for the Journal on Criminal Justice Policy")

More than a decade ago I published an article entitled "Policy Analysis v. Social Science: Some Fundamental Differences."¹ In that article, I claimed that there was, indeed, an important overlap between social science efforts made to understand and explain how the world works on one hand, and policy analytic efforts made to evaluate alternative courses of action in the world on the other. That connection lay in two closely connected ideas. First, that the methods of social science (both the close reasoning we associate with theoretical efforts, and the careful use of real experience captured in data that we associate with our best empirical work) constituted the best method humans could use to understand their world. Second, that in order to act both responsibly and effectively, policy-makers needed social science methods and findings to be able to make reasonably accurate estimates about the consequences of their action. It was science's ability to supply information about estimated effects, and the policy-maker's need to estimate consequences of proposed actions that formed the basis of a potentially important partnership.

But I also argued that, *in practice*, the relationship had proved to be much less productive than it first seemed. Further, I argued that the reasons for the difficulty were not all that hard to understand once one stopped assuming that these enterprises would naturally be closely aligned, and instead made the assumption that the enterprises were very different. It seemed obvious, on reflection, that while the *methods* of social science (close reasoning, careful use of empirical evidence) seemed crucial to meet the policy-maker's need to know, the fundamental *goal* of the social sciences (to work cumulatively to produce verified theories about how the world works) was a different project than the policy-makers had in mind. The policy makers didn't need to know how the world worked *in general*. And they couldn't wait to act (or to refrain from acting) until everything was known. They needed to know how a *particular piece* of the world for which they were responsible worked, and what they, using resources and capabilities they actually had, could do to deal with the problems or exploit the opportunities in their larger or smaller corner of the world. In short, policy makers were less interested in *general causal description* than they were in *specific policy prescription*.

¹ Mark H. Moore, "Policy Analysis v. Social Science: Some Fundamental Differences" in Daniel Callahan and Bruce Jennings eds., *Ethics, The Social Sciences, and Policy Analysis* (New York, Plenum, 1983, pp. 271-291)

The move from both the *general* to the *particular*, and from *description* to *prescription*, imposed burdens on reasoning that science was either not particularly interested in (the specific rather than the general), or was not particularly good at (prescription rather than description). It was not particularly interested in the particular because the main goal of science was to produce findings that were true regardless of a particular time and place. While having such knowledge would certainly give policy-makers a huge advantage, in most areas of practical interest, this kind of knowledge simply did not exist, and was not likely to exist anytime soon.

The difficulty between science and policy-analysis was even more obvious when the focus shifted from (causal) *description* to (policy) *prescription*. As a logical matter, to decide on an action, one has to have a goal, or a purpose, or a way of evaluating whether the world is better or worse off than it was before one acted. Often, that requires one both to describe the intended effects of a given policy (the *positive* part of the task), and to offer philosophical and political defense of why those effects are valuable, and why they should be pursued with the collectively owned powers and material resources of the state (the *normative* part of the task). It is often also important to be able to imagine a set of *unintended effects* of a given policy that could also occur, and would affect society's overall judgment of whether things had improved or not, all effects considered. In order to get close to such judgments, one has to fire one's moral imagination about what society does or should value, as well as exploit one's scientific capacities to estimate the consequences of a particular line of action for each of the valued effects, whether the effects were intended or not.

In short, policy analysis nominates a particular set of dependent variables that are likely to move as a consequence of a particular course of action, and does so by combining scientific judgments about how the world will be affected by a particular course of action, as well as moral and political judgments about the kinds of effects that would be worth noticing as we tried to make an overall judgment about whether the world had been made better off. Science has always said that it was incapable of making value judgments (other than the idea that it was always better to know than to remain ignorant, and that the methods of science were the most powerful ways of knowing). It was capable only of making estimates of consequences. What constituted an important consequence of a policy from a normative perspective was left to philosophers, to political processes, or to individuals who were free to have their own views about what was valuable.

Importantly, program evaluation emerged as an applied intellectual effort that, in many ways, represented the happiest combination of science on one hand and policy-making concerns on the other. In its best versions, program evaluation relied heavily on the methods of social science. We used experimental designs to test the impact of particular program interventions. When experimental designs were not feasible, we used elaborate statistical methods to exploit the unplanned variation that occurred in the world to help us guess the like impact of particular policy interventions. We got past the normative issue of identifying the set of effects that could be viewed as philosophical, social, or political justifications for the particular policy intervention by finding the effects to be evaluated in the stated intentions of those who authorized the intervention. That is, we consulted the legislative or executive policy making discussion about the policy to learn what effects the policy makers seemed to have intended, and used that list of desired results become the dependent variables to be measured and analyzed in program evaluations. (We then often consulted our own empirical and moral imagination

to nominate and explore a different set of unintended consequences that were both valuable -- or harmful -- enough, and big enough, and foreseeable enough, to cast doubt both on the overall value of what was done, and the quality of the decision-making process that preceded the decision.)

While program evaluation efforts often seemed to exploit the best opportunities for science and policy-making to join together in common cause, it is worth noting how far even this significant success is from a true unification of the two distinct enterprises. From the point of view of pure social science, program evaluation has always seemed like a second-rate, applied activity. After all, the only thing a successful program evaluation showed was whether or not a particular intervention did or did not produce a particular result. It did not necessarily show why it had succeeded or failed. The underlying, general causal mechanisms remained hidden. From the point of view of policy-analysis, program evaluation was useful because it commented on the effectiveness of past actions, and provided reasonable estimates of what would happen if we tried that particular intervention again, but it was silent on the question of what might be the right action to take prospectively from this particular point in time and space. There was also the worry (supported by lots of actual experience) that the results of the most recent evaluation would be overturned by both the criticism of the social scientists, and the next evaluation effort that was made. So, what looked like a strong finding, bearing directly on the point of what should be done, often turned out to offer only weak help to policy-makers confronting the issue of an unknown future they were being asked to shape.

I make these points partly for the not very good reason that I am disappointed that they do not seem to be widely understood, even though they seem undeniably true and pretty elementary to me, and this seems like a good opportunity to repeat them. More to the task at hand, however, I make these points because I want to illustrate their importance to what might be considered the policy interpretation of the results produced by _____ and _____ in "_____."

In most important respects, this piece of work could be viewed as a classic piece of program evaluation. The policy being evaluated is a federal funding program that distributed tax dollars to police departments throughout the country. The program had several different funding streams, each of which could be viewed either as a separate program, or as a part of the larger program. The authors assume that either an important purpose of this program (i.e. intended effect), or an important consequence of this program (i.e. unintended effect), was to raise or lower levels of reported violent and property crime in the country. The task they essay is to estimate the impact of this funding on reported violent and property crimes in ____ cities and towns in the United States, as though knowing this would help us decide whether and how to continue this Federal funding program.

No doubt, one can quibble with their analysis. One can say that their data is too weak, that they have ignored many other factors that might have produced the results that they found, that the methods they used for making the causal inferences are not the right ones, and so on. I would happily join the chorus here. (If given this opportunity, I would focus particularly on the question of whether levels of funding actually produced any changes in the level, style, or focus of policing in the communities that were being studied. Given that most empirical studies of federal funding programs tend to show that even though the money was spent, the programs that were ostensibly funded were often

not implemented, it is not at all clear to me that the receipt of federal grants is the same as either increased or improved or more effectively focused policing. And if we do not have a good measure of the independent variable, how can we make claims about the effect of that variable on crime?) But I will stand aside in awe of the brute empiricism of a sample of _____ years of federal funding for police and crime experience in _____ cities and towns. We may not be sure what happened, but it sure seems like something did. I will focus instead on what is often euphemistically called the "policy implications" of this evaluation of a policy intervention.

Viewed from one perspective (and not a dumb one), the implications are pretty straightforward. The results say that federal funding of police departments seems to be able to reduce crime. If reducing crime is an important social objective, and we want to continue to pursue it, we should continue, maybe even expand this program. It is precisely at this stage that one could begin a different kind of quibble than an empirical quibble -- let's call it a *normative policy quibble* as opposed to a *scientific empirical quibble*. What one begins to worry about is not whether the effects are as they seem to be, but instead whether the normative framework that is wrapped around the facts, and carries the facts from brute, scientific facts to important reasons for taking action in the world is adequate to the task. (As social scientists, we are all well trained – perhaps too well trained! – to quibble on scientific, empirical issues. We are trained not at all – indeed, are urged to stay far away from quibbling on normative grounds.)

Recall, again, that the important empirical finding is that a federal program giving money to local police departments through three sub-programs seems to reduce violent and property crimes. The important policy implication is that if we like the effect and wish to continue to pursue it, we should continue the program. Needless to say, I like that conclusion. But it isn't difficult for me to imagine a variety of ways in which that logic could be picked apart – on normative rather than scientific grounds.

Consider, first, the idea that the program should be continued because it is achieving a desirable objective: namely, reducing violent and property crime. Nobody (well, almost nobody) doubts that reducing crime is a valuable social result. An important question, however, is exactly *how* valuable. This seems important to decide so that we can compare it with the *costs* of producing that effect. To weigh the benefits against the costs, we have empirical analysis gives us this.) But beyond this, we have to decide how much we should *value* an effect of that size. (The empirical analysis can say nothing about this.) It would be particularly nice if we could express the value of that effect in the same currency that we use to describe the costs: namely, dollars of crime reduction benefit compared to dollars expended in the federal program. Given that we have pretty complete and accurate information on costs, that we have a good estimate of the size of the crime reduction effect, and that there have been many different efforts made to estimate the value of avoiding the cost of particular kinds of crime, it seems like it would have been a natural extension of this analysis towards a full cost-benefit analysis of the COPS program. We have the costs, we have the estimated beneficial effects, we have a way of monetizing the value of the beneficial effects. Why not go ahead and calculate the "net public value" of the effort and represent it in financial terms, or as a ratio of benefits to costs?

There are many possible answers to that question, including the claims that the authors simply didn't have the time; or that they were saving this for an additional publication; or that they were counting on one of their colleagues to go ahead and finish

the calculation for them. But one of the reasons I suspect these authors did not go further was that to go further would be to embroil them in a messy normative discussion where the purchase of science is much less strong than in what they did.

They have an estimate of a relationship between an independent variable (funding from the COPS program) and a dependent variable (reductions in violent and property crime.) They have built that estimate from established scientific methods. That alone is enough to meet the minimal demands of program evaluation. They have investigated the question of whether a program accomplished (at least one) desired result, and found that the sign of the effect was probably in the right direction, and of (empirically) significant size. As they report, the COPS program seems to have reduced crime in cities, and done no harm in smaller towns.

To take a step further, and look closely at the costs of what was produced would have moved their study from an analysis of whether the program worked, to an analysis of whether the program was "cost-effective." By a cost-effectiveness study, we mean a study that helps us understand whether this policy worked better than other policies in achieving the same desired result, or whether the effectiveness of the policy was great enough in quantitative terms to justify the resources required to produce the results. That would have required more work. The work would include not only the collection of cost information, but also a beginning exploration of the crime control effectiveness of other policies and programs, and/or an exploration of whether the size of the crime reduction effect seemed large or small relative to the costs of producing it.

At that stage, they could have taken a step further and tried to estimate the value of the particular size crime reduction effect they achieved by monetizing the value of avoiding a certain number of crimes. Each step along the way plunges them into not only more work, but *more controversial* work. The reason is that they are trying to go from physical, material relationships in the world to normative evaluations of a set of effects. After all, to describe something as a cost and another thing as a benefit is to make a moral judgment about the value of that thing. (A cost is something that is negatively valued by the society. A benefit is something that is positively valued.) So, as soon as one starts talking in terms of costs and benefits, one has interjected some important normative content into the discussion.

These concerns alone might have caused me to resist the temptation to carry out a full cost-benefit analysis. But there are additional problems lurking just beneath the surface. On one hand, it is not at all clear to me that this analysis has fully captured all the relevant costs and benefits of the policy. Looking first at the cost side, it is not clear to me that the only costs worth considering are the monetary costs of the policy. We often think (rightly, I believe) that when we expand publicly financed law enforcement efforts, we are spending not only our hard earned and grudgingly surrendered tax dollars that might better be spent on our own welfare or some other more favored government program, but we are also giving up some share of our privacy and our liberty. We have strengthened the capacity of the government to monitor, to observe, to pry into. After all our liberty is protected not only by the fact that we have constitutional rights which we can defend when they are violated, but also by the fact that the government doesn't really have enough money or agents to be looking very closely at everything. For the most part, the government conserves its money (as well as our freedom) by waiting for an urgent situation (like a crime) to occur before it dispatches an agent and begins looking around. An important question, then, is whether we should view the loss of

liberty as an important cost of policing along with the dollars we expend to enlarge the police force.

Note that I am not saying here that our privacy and our liberty are not worth giving up to accomplish a practical goal like keeping ourselves safe from crime. I am perfectly willing as a citizen to give up both money and freedom to accomplish this goal that is important to me both as a potential victim in the society and as citizen who would prefer to live in a secure society. The only point is that it might be worth recognizing how much of my liberty I am expending as well as how much of my money. This might be particularly true if one of the ways that the federal government is helping the police to reduce crime is by encouraging them to spend more of my privacy and liberty without having to account to me for having done so. Insofar as the evaluation given to us is silent on the question of both monetary costs on one hand, and privacy and liberty costs on the other, we are not in a very good position to know whether we should carry on with the existing policy, abandon it, or shift to something else. We know that the program worked to produce a desirable result. (That is good to know. If we had found the opposite we would surely know what to do!) But once we see that we had a good effect, we still can't be sure that the program is worth it, or that it is the best way to accomplish the result -- all things considered.

Once we start talking about the use of force and authority as a cost (financed by abandoned privacy and liberty), we create not only technical and empirical problems in measuring these costs, we also begin to cast doubt on the power of what could be called the "social utilitarian frame" we commonly use to evaluate social programs, including criminal justice programs. By a social utilitarian frame, I mean the basic idea that we should view social policies largely in instrumental terms: as things that cost a certain amount of money, and produce a particular set of social consequences that could be judged to be good or bad relative to our ideas about what makes for a good and prosperous society. To a degree, I am relying on this frame when I try to talk about the use of force and authority as a cost of a particular policy as well as the financial or economic cost of the policy. I am trying to capture something that is actually used in making a policy more or less successful in reducing crime (namely, the use of state authority) in the utilitarian framework in which we look at all negative effects as costs and all positive effects as benefits.

But I suspect that to many, this clumsy effort to treat the use of force and authority as a cost that needs to be reckoned in the calculation of overall societal benefit seems both forced and wrong-headed. To many, the use of force and authority shouldn't be judged by its effectiveness in producing a result; it should be judged by its propriety and fairness. There are some acts and some situations that deserve (or require) the use of force and authority if justice is to be served. This invokes a very different evaluative frame for looking at criminal justice policies; one that could be called the "deontological" or "justice" frame. In this framework of evaluation, criminal justice policies are considered good or bad not in terms of whether they reduce crime, but instead in terms of whether they enhance the quality of justice in the society. The idea of justice, in turn, could focus on the question of whether offenders got what they deserved and were called to account for their crimes. (This tends to be the focus of the right of the political spectrum.) Or, it could focus on the question of whether individuals suspected of crimes had their rights protected in the course of the investigation. (This tends to focus of the left of the political spectrum.) Or, it could even focus on the question of whether citizens in the society were generally treated equally both in terms of having to bear the burden of police scrutiny

when the police are doing their duty of trying to protect us all from crime, or receiving equal benefits in terms of a prompt, polite and resourceful response when they are called. The point is, here, that when we do evaluations of criminal justice policies, it might not be adequate to evaluate them solely in terms of whether they are successful in reducing crime, and whether they do so at a low cost in terms of both the use of money and authority. We might have to address ourselves to the question of whether the justice of the system has been enhanced or diminished by the initiative.

One last point: it seems to me that one of the most important questions we might have to consider in whether we should continue the COPS program, end it, or modify it in some way has something to do not only with the effectiveness of the program in reducing crime, but also with maintaining the "right balance" between the federal government on one hand and state and local governments on the other. Arguably, the foray of the federal government into the more or less permanent, large-scale funding of local police agencies raises many important questions about the political economy of the justice system.

Since our founding, we have valued the idea that local communities should both have the right and bear the responsibility of policing themselves in the ways they thought best. That commitment produces a high degree of responsiveness to the different concerns of local communities. It also produces a kind of fairness in the sense that citizens only get the kind of policing they are willing to pay for and demand; they can't get others to pay for the policing that they want, and they don't have to put up with a kind of policing that doesn't suit them. And it also (potentially) produces a great deal of variety in police practices, which (in principle) allows us innovate and learn about what is both effective and just in policing. These are the commonly asserted benefits of a decentralized, federal system of government.

Presumably, that regime is threatened at least a bit by the federal funding. In the extreme case, we imagine the federal government luring innocent local police departments into a commitment to a certain kind of policing that the federal government has mandated, or into a working relationship with federal agencies, and then backing out of their financial commitment to keep paying for the policing. In short, we see in the background the threat of a "national police force," or of "unfunded federal mandates."

On the other hand, one can argue that the idea of widely decentralized policing is an idea whose time has passed. In a world in which we confront a serious threat from terrorism, in which local police departments constitute a front line of defense it seems important to develop a more centralized, co-ordinated system so that intelligence clues don't fall between the cracks, and where organizations can act in concert in response to threats that vary in scale and location. It may also be true that policing could be more efficient and effective in dealing with ordinary crime if the industry as a whole were organized in a more concentrated way. After all, there may be significant economies of scale to be exploited in policing just as there are in other enterprises. It may be that there are great technological benefits to be seized if we organized the demand side of the market for police technology in ways that lowered the risk to technology developers. And so on. These observations might well lead to the conclusion that the time has come to centralize the system of policing in the United States, and federal financing is a good first step.

I don't think there is a right answer to this institutional design question. And even if there were one, I'm not sure that social science could generate it. These are the kinds of questions about institutional structure that statesmen must face. And they must do it with the kind of long term, broad perspective that is helped more by a sense of history, and a comparative analysis of institutional arrangements in other countries than from the usual kinds of social science studies.

So, the argument I am making is simply this: social science methods, when well deployed, can offer bits of information that are important to policy-makers. As a matter of principle however (to say nothing of practice), social science findings can never fully dictate the right answer to an important policy question. They cannot do this even when the methods are deployed powerfully in program evaluations. And it is not just because the relevant sciences are not yet mature. It is because important normative questions remain entirely beyond the reach of science, and because any important policy choice involves important positive issues that science has not yet, or could not easily ever reach. An important question for those who use scientific methods, theories and findings to inform policy, then, is how they should deploy their commitments to dispassionate reasoning and interpretation of evidence to issues that are currently (and perhaps forever) beyond the reach of science. The answer to that question would show us what the world of policy analysis and design should be about, and I suspect (well, actually I know) that it is a much different world than the world of social science.