

The Elusive Craft of Evaluating Advocacy

By Steven Teles & Mark Schmitt

Stanford Social Innovation Review
Summer 2011

Copyright © 2011 by Leland Stanford Jr. University
All Rights Reserved



The Elusive Craft of Evaluating Advocacy

BY STEVEN TELES &
MARK SCHMITT

*Illustration by
Gordon Studer*

Very few big social changes happen without some form of advocacy. When these efforts succeed, the results can be transformative. Consider the recent expansion of charter schools or health care reform in the United States. Good ideas like these did not catch on widely just because they worked. They happened because of creative investments in public persuasion, legislative action, and political activity.

Most successful foundations and nonprofits understand the importance of advocacy. Over the last decade, foundations have put more resources into advocating for the policies they believe in, with some notable successes. Yet grantmakers have often hesitated to plunge in. Sometimes they worry about appearing too political or partisan. But more often they hesitate because effective advocacy is difficult, and evaluating whether various approaches are working is even harder.

That is not the case when it comes to service delivery programs—such as well-baby clinics or job-training classes—where foundations, universities, and government agencies have developed sophisticated tools for evaluating the effectiveness of these efforts. The tools range from controlled experiments, to extracting from experience best practices that can be adapted from one successful

program to another, to a more malleable form of evaluation based on assessing the theory of change underlying an initiative. The development, refining, and implementation of these tools constitute a growing industry.

Unfortunately, these sophisticated tools are almost wholly unhelpful in evaluating advocacy efforts. That's because advocacy, even when carefully nonpartisan and based in research, is inherently political, and it's the nature of politics that events evolve rapidly and in a nonlinear fashion, so an effort that doesn't seem to be working might suddenly bear fruit, or one that seemed to be on track can suddenly lose momentum. Because of these peculiar features of politics, few if any best practices can be identified through the sophisticated methods that have been developed to evaluate the delivery of services.

Advocacy evaluation should be seen, therefore, as a form of trained judgment—a craft requiring judgment and tacit knowledge—rather

than as a scientific method. To be a skilled advocacy evaluator requires a deep knowledge of and feel for the politics of the issues, strong networks of trust among the key players, an ability to assess organizational quality, and a sense for the right time horizon against which to measure accomplishments. In particular, evaluators must recognize the complex, foggy chains of causality in politics, which

The political process is chaotic and often takes years to unfold, making it difficult to use traditional measures to evaluate the effectiveness of advocacy organizations. There are, however, unconventional methods one can use to evaluate advocacy organizations and make strategic investments in that arena.

make evaluating particular projects—as opposed to entire fields or organizations—almost impossible.

If foundations embraced the judgment-laden character of the effort—rather than giving up on advocacy or feeling they are falling short when their evaluations lack the scientific patina of service delivery program evaluations—the benefits would be profound. Funders could structure programs, often involving multiple unlikely bets, in ways that are more likely to succeed. Advocates could feel comfortable changing course as necessary. And foundations would be more likely to take chances on big efforts to change policy and public assumptions, rather than retreating to the safer space of incremental change.

Advocacy Is Different

The word *advocacy* is in many ways a misnomer. Funders do not, for the most part, give organizations money to simply fly the flag or make the case for a particular policy. Their goal is to change actual social, policy, and political outcomes. And ultimately, advocacy efforts must show progress toward those outcomes. But the relationship between the work to create those outcomes, and the actual results or signs of that progress, can be elusive, because advocacy by its nature is complicated and its impact often indirect.

Consider, for example, the campaign for US health care reform. The effort that culminated in 2010 was the result of decades of work, including a previous, high-profile failure in the early 1990s, waves of state-based reform, and numerous incremental efforts at the national level. Advocates invested hundreds of millions of dollars in initiatives ranging from media campaigns encouraging television producers to include stories of the uninsured, coalition-building projects, university- and think tank-based research, and grassroots initiatives. The basic outlines of reform policies were worked out well in advance, in advocacy groups and think tanks, which delivered a workable plan to presidential candidates. Important interest groups who could block reform, such as small business, had been part of foundation-supported roundtables seeking common ground for years. Technical problems had been worked out. And tens of millions of dollars had been set aside as long ago as 2007 for politically savvy grassroots advocacy initiatives targeted at key legislators. After a very long slog, the outcome was the Patient Protection and Affordable Care Act.

Now consider the effort in the United States to pass legislation to control global warming, which in many ways resembled the strategy to pass health care reform. Advocates of cap and trade engaged in what can only be called a mammoth effort, over more than a decade. Among other things, environmentalists drew on the services of a former vice president who made an Oscar-winning movie, spread their message for more than a decade across a remarkable span of media (up to and including children's cartoons), corralled a wide range of well-funded environmental groups to support a single strategy for reducing carbon (cap and trade), and attracted substantial support

from large businesses. The movement used every trick in the book (as well as creating some new ones), and the result was legislation that never made it to the floor of the US Senate, with the very real possibility that action will be delayed by years, if not decades.

Most advocacy efforts look more like the push for cap and trade than like health reform. That is, even the best designed and resourced initiatives usually fail to achieve even the more modest of their goals. The American political system is profoundly wired for stasis, and competition for limited agenda space is fierce. In an overwhelming percentage of cases, organizations fail to get substantial traction on their agendas for change. Conversely, items often wind up on the political agenda by random and chaotic routes that may have little to do with advocacy campaigns. If it is hard to know whether advocacy played any part in a policy outcome, it is harder still to know whether any particular organization or strategy made the difference. The fact that in 2010 Congress passed health care reform and not global warming legislation may have been mainly a function of dumb luck, rather than an indication of which advocacy campaign was better executed. Or it may tell us more about the enthusiasm or skill with which the campaign was implemented, rather than the general applicability of its tactics or strategies.

Despite the number of groups that will present themselves as the decisive force behind any legislative accomplishment, no successful advocacy effort is the result of any one organization or initiative. Health care legislation, for example, owes its passage to many efforts. Some were far from government, such as the academic work at Dartmouth College that showed how escalating health care costs could be contained while improving services. Some weren't directly focused on health care at all, such as political organizing around a broad progressive agenda and candidates.

When specific forms of advocacy—an aggressive grassroots campaign, or a behind-the-scenes, cross-partisan strategy involving paid lobbyists—receive credit for changes in policy, advocates adopting those strategies in the future may claim the strategies themselves are a marker on the road to success. But tactics that may have worked in one instance are not necessarily more likely to succeed in another. What matters is whether advocates can choose the tactic appropriate to a particular conflict and adapt to the shifting moves of the opposition.

Sometimes the most effective effort might be a disruptive innovation that does not follow known strategies. Consider MoveOn.org, for a time one of the country's most effective multi-issue advocacy organizations. But for years after its establishment in 1998, the organization attracted skepticism, because its primary strategy—repeated, small actions by members—was so different from the organizational membership model previously considered the standard of success.

Disruptive innovators may require a long period of trial and error, during which the policy landscape, and what strategies work within it, may change significantly. For example, when the Tax Reform Act of 1986—an iconic example of unlikely, bipartisan success—was passed, the political climate in Washington, D.C., was characterized by extremely weak parties, strong congressional committees and subcommittees, significant room for bureaucratic and interest group entrepreneurship, and pervasive cross-party coalition building. Best practices based on those conditions would make little sense in

STEVEN TELES is an associate professor of political science at Johns Hopkins University. He is the author and co-editor of several books, including *The Rise of the Conservative Legal Movement: The Battle for Control of Law*.

MARK SCHMITT is a senior fellow and director of the fellows program at the Roosevelt Institute. He was previously executive editor at *The American Prospect*, director of policy and research at the Open Society Institute, and policy director for former Sen. Bill Bradley.

today's national policy process, characterized by polarized, highly disciplined political parties.

Advocacy efforts almost always involve a fight against a strategic adversary capable of adapting over time. Practices that once worked beautifully get stale once the losers figure out how to adopt the winner's strategy or discover an effective counterstrategy. There was a time when bombarding the Congress with phone calls was an effective way of exercising influence by indicating mass support, but it became nearly useless once everyone did it. Strategic litigation was a genuinely disruptive innovation in the 1970s but declined in impact as its targets developed their own organizations and figured out ways to push back against public interest lawyers. The declining returns on political tactics that are a result of the repeated, competitive nature of advocacy makes it almost impossible to evaluate advocacy strategies against the metric of best practices.

Evaluating Advocacy Is Hard

Similar resources and advocacy strategies, therefore, can generate very different results. Sometimes political outputs are reasonably proximate and traceable to inputs, but sometimes results are quite indirectly related and take decades to come to fruition. Some advocacy efforts have a specific goal in mind, but in other cases the objective is broader and the benefits are reaped by groups other than those who paid the costs. Any effort to evaluate advocacy must be able to account for these and other complicating features of the terrain of policy and institutional change, but these facts in themselves can't help us evaluate advocacy. Indeed, they are more useful as guides to what not to do.

When evaluating service delivery programs, such as food banks and after-school enrichment programs, it is relatively easy to establish benchmarks to measure a program's effectiveness. By and large, most organizations are able to show some visible progress toward reaching their goal every day (even if it's just one hungry person getting breakfast). But the chaotic, nonlinear character of policy agendas means that funders cannot pretend to know where they are in the process. Most of the time, very little seems to be happening. As University of California, Santa Barbara, sociology professor Verta Taylor points out in her classic study of the women's movement, a cause can remain in "abeyance" for decades, but if the fires are kept burning, it's possible to get things moving when conditions become more permissive.¹

These long periods off the agenda can be broken quite abruptly and without warning. That is why it is important to continue to fund and pursue the quiet work—such as the long process of slow persuasion and litigation that led to the repeal of the "don't ask, don't tell" policy in 2010—even when attention is elsewhere. If one doesn't, then opportunities may be missed when the political weather changes.

Advocacy strategists, conditioned by funders, are accustomed to presenting a plan of action in which a large change is preceded by interim goals and achievements. A plan to achieve nationwide reform on a key issue might have as interim goals the passage of state ballot referenda, a specified number of co-sponsors for legislation, or passage of an incremental reform. An organization that can present a plan for advocacy with a well-marked path to success seems like a

business with a coherent plan pointing to profitability, and thus the safest bet for strategic grantmaking. Such a project can be evaluated as to whether it is achieving its projected interim goals.

But successful advocates know that such plans are at best loose guides, and the path to change may branch off in any number of directions. Interim achievements, such as the passage of a state ballot initiative, can be idiosyncratic victories. Incremental legislation often satisfies politicians that they have dealt with a problem while exhausting the capacity of grassroots advocates to keep pushing forward. Given the competitive nature of advocacy, such early under-the-radar successes may even have the unintended consequence of mobilizing the opposition, making later change more difficult.

Successful advocacy efforts are characterized not by their ability to proceed along a predefined track, but by their capacity to adapt to changing circumstances. The most effective advocacy and idea-generating organizations, such as the Center on Budget and Policy Priorities or the Institute for Justice, are not defined by a single measurable goal, but by a general organizing principle that can be adapted to hundreds of situations. Rather than focusing on an organization's logic model (which can only say what they will do if the most likely scenarios come to pass), funders need to determine whether the organization can nimbly and creatively react to unanticipated challenges or opportunities. The key is not strategy so much as strategic capacity: the ability to read the shifting environment of politics for subtle signals of change, to understand the opposition, and to adapt deftly.

The US system of government is characterized by parallel, loosely coupled agenda-setting processes at work simultaneously at different levels of government and across institutions. In sharp contrast to service delivery programs, then, advocacy projects cannot realistically experiment in one place in the hopes that successes can be scaled up. Successful advocacy projects must simultaneously pursue opportunities at the local, state, and federal level, as well as across governmental institutions. Sometimes these efforts need to be organized into a well-coordinated network, whereas in other cases they are best left uncoupled, pursued as a portfolio of distinct bets on the assumption that donors have little or no idea which strategy is likely to be successful. Under such conditions, it makes sense to evaluate the portfolio as a whole, not the individual projects.

Successful efforts to change public policy often require grassroots as well as elite strategies, because opposition in either quarter could derail the idea. For example, decades of work within the medical profession built elite support for comparative effectiveness standards to ensure appropriate treatment, but with no grassroots effort, it was effectively mischaracterized as "death panels."

Building advocacy projects that cover a range of political institutions and processes means that massive amounts of effort will seem wasted, because most will be unconnected to the final outcome. This waste, however, is unavoidable, because neither funders nor the organizations they support can know which strategy will be effective ahead of time.

Because funding is finite, there can be a tendency to view issues and advocacy efforts as if they are in competition for a limited amount of political capital or public attention. But success on one issue often builds a foundation for others by creating a sense of political momentum, restoring faith in government, establishing a precedent, or

creating habits of cooperation within legislative institutions. Even failure to achieve an identified goal can leave energy and momentum to achieve the goal in other ways. The massive push for the Equal Rights Amendment, for example, fell short in its constitutional goals but led to change through the courts that realized much of its larger ambitions.

Issue domains that may seem quite distinct in a donor's mind are rarely so in politics. Because issues spill over from one domain to another (issues of poverty affect health and education), particular issues are almost impossible to disentangle from general ideas and broader governing philosophies. Consequently, the fortunes of issue-specific mobilization may be due to actions conducted within that domain, but they may be reinforced by mobilization in another domain entirely, by generic, ideological activity, or by more neutral scholarly research. For example, over the past decade, efforts to reform K-12 education have focused on innovations such as charter schools and performance pay for teachers that are often opposed by teachers unions. The perception of teachers unions as powerful and intransigent was then transformed into a backlash against all public employee unions, manifested most recently in the proposals to end collective bargaining in Wisconsin and other states.

That is why it is difficult to accurately attribute the success of any advocacy project to a particular organization (or even issue-specific network). External effects of organizational activity (benefits created by one organization that are reaped by another) are pervasive in advocacy in a way that they are not in service delivery programs. Evaluators are faced, therefore, with the challenge of capturing all the benefits that an organization is generating, as well as preventing it from taking credit for benefits that are produced by others or that are due to good fortune rather than skill.

Evaluating Advocacy

Despite these many challenges, there are ways grantmakers can effectively invest in and evaluate the success of advocacy campaigns. One thing grantmakers can do is to use a spread-betting approach to making grants. A spread-betting approach invests in a wide range of organizations, strategies, scenarios, and even issues. Failing to fund the seemingly quirky, unproven strategy that turns out to be appropriate to the circumstances is just as big a loss as funding something that does not work out. Spread betting, therefore, requires that funders have an organizational culture that does not punish even a considerable number of failures, so long as they are balanced over the long term by a few notable successes.

Grantmakers should also focus on the aggregate return on investment of their entire portfolio of grants, not the success or payoff of any one grant. An investment in an issue in which no action has occurred, even for a long time, may not be a bad use of resources. But this will only be clear when a particular issue is judged in the context of a range of other bets put down by the donor. Only then can a donor have a sense of whether his resources are generating what investors call “alpha”—excess returns over the average. Portfolio evaluation, by averaging out a number of investments over a longer period of time, also prevents the risk of over-attribution of success or failure to factors that are entirely exogenous to the activities of those they are investing in.

Funders should evaluate their portfolio of investments using the longest feasible time horizon, recognizing that the political process does not end after a piece of legislation passes or a court decision is handed down. It allows for the assessment of what University of Virginia political science professor Eric Patashnik calls “policy durability”—whether a reform actually sticks or creates a platform for further change.² Some reforms, such as airline deregulation and tradable permits for sulfur dioxide emissions, generated powerful reinforcing dynamics that kept the policies from being clawed back, even in the face of initially strong opposition. But other changes that seemed momentous at the time, such as the Tax Reform Act of 1986, unraveled bit by bit over the years. Viewing policy enactment as only one step in a much longer process focuses donors' attention on what really matters—whether a policy sinks deeply into society and political routines. Funders may not be able to wait for years after reforms pass to judge whether their investment was worth it. But at the very least they should consider the possibility of reversal (or extension) in their evaluations, and evaluate the strategies of advocates by whether they have a plausible plan for protecting what they have won.

Some policy changes matter because they change the playing field on which subsequent action can occur. For example, the state welfare reforms of the 1980s and early 1990s did much more than change policy in the states where experiments were carried out: They altered the entire debate and emboldened policymakers to try more ambitious national changes.

In many cases, policy changes have political feedback as one of their primary objectives. Investment in green jobs, for example, when it emerged as a policy priority in 2005 or so, was billed by its supporters as a “strategic initiative” that, in addition to being good policy, might create a lasting labor-environmentalist alliance, mobilize voters around an optimistic economic vision, put a bright face on the tough choices of carbon pricing, and create a message of reduced oil dependence—not just create jobs and improve the environment. To some extent, the idea achieved those goals even as the policy itself fell short. Libertarians' litigation on issues like school choice and property rights was designed to detoxify their brand among racial minorities, along with achieving substantive policy and legal goals. The long-term effects of policy change on the character of politics can be at least as important as those that are produced by the policies themselves.

Evaluate the Advocates

We have argued that grantmakers should evaluate the success of advocacy efforts by thinking of them as long-term, portfolio-based, and inclusive of diffuse and indirect effects. We would now like to take this argument one step further—making perhaps our most radical suggestion—that funders may be better off eschewing evaluating particular acts of *advocacy*, and instead focus on evaluating *advocates*. We believe that the proper focus for evaluation is the long-term adaptability, strategic capacity, and ultimately influence of organizations themselves. This is the grantmaking model that the Sandler Family Supporting Foundation used to help create the Center for American Progress and ProPublica, and that the Walton Family Foundation uses to promote educational competition.

Evaluating advocacy organizations means paying close attention to the value they generate for others, rather than only focusing on their direct impacts. For example, the Center for American Progress's Campus Progress and *The American Prospect's* writing fellows program focus a great deal of effort on developing the talent of their younger advocates, writers, and activists. As a result, these organizations regularly lose their younger staff to more prominent organizations. Although this approach doesn't add much direct value to the two organizations, it does create enormous value for the larger ecosystem. In this instance, the advocacy evaluator needs to understand that in some cases staff turnover reflects organizational success, not failure.

The best way to evaluate an organization whose influence is extremely diffuse is for grant officers to be close to the political action and thus able to make informed judgment calls on how it conducts its core activities. This was the practice of many conservative foundations, whose staff devoted much of their time to simply reading the primary work of their grantees, rather than asking them to generate problematic metrics and lengthy reports designed solely for purposes of evaluation. Empowered by their boards or donors to trust their own judgment of good, appropriate work, this foundation strategy has been vindicated many times over in the real world of politics and the marketplace of ideas.

Equally important is an organization's strategic capacity, which can be defined not only as its formal strategic plan, or the wisdom of its senior leadership (two factors that funders tend to focus on), but also the organization's overall ability to think and act collectively, and adapt to opportunities and challenges. A good organization has a coherent and inspiring internal culture, the ability to consistently identify and motivate talented people, acquire and process intelligence, and effectively coordinate its actions. Effective advocacy organizations—such as Planned Parenthood, which recently maneuvered through a significant shift in their political alliances on reproductive rights—have a record of innovating and reorganizing when their tactics don't work as well as they once did.

Yet another way to measure an organization's quality and influence is through "network evaluation"—figuring out its reputation and influence in its policy space. Although this is probably the most important form of knowledge, it is also the most difficult to acquire. Where organizations are in competition with each other for resources, peer evaluations may be too harsh. When organizational leaders have close personal links, their assessments are likely to be too generous. And of course, all advocates have profound incentives to overstate their own importance.

Participants in a policy network may be hesitant to share accurate information with outsiders with whom they lack ongoing relationships, such as consultants hired by the foundation. Advocates may reveal their challenges only to those whom they trust profoundly. Nonetheless, members of policy networks generally do develop reasonably accurate assessments of which of their peers they listen to and trust, who does good work, and who policymakers take seriously. What donors are really looking for is network centrality—which actors play vital roles in issue networks. It is not too difficult to use network mapping to figure out these connections. The real art of advocacy evaluation, which is beyond the reach of quantitative methods, is *assessing influence*, which is what funders are really paying for.

What Makes a Good Evaluator

Advocacy evaluation is a craft—an exercise in trained judgment—one in which tacit knowledge, skill, and networks are more useful than the application of an all-purpose methodology. Evaluators must acquire and accurately weigh and synthesize imperfect information, from biased sources with incomplete knowledge, under rapidly changing circumstances where causal links are almost impossible to establish. There is a natural temptation to formalize this process in order to create at least the appearance of objective criteria, but it is far better to acknowledge that tacit knowledge and situational judgment are what really underlie good advocacy evaluation, and to find evaluators who can exercise that judgment well. It's the evaluator, rather than the formal qualities of the evaluation, that matters.

If scientific method is an inappropriate model, where can grantmakers look for an analogy that sheds light on the intensely judgmental quality of advocacy evaluation? One possibility is the skilled foreign intelligence analyst. She consumes official government reports and statistics, which she knows provide a picture of the world with significant gaps. She talks to insiders, some of whom she trusts, and others whose information she has learned to take with a grain of salt. In many cases, she learns as much from what she knows are lies as from the truth. It is the web of all of these imperfect sources of information, instead of a single measure, that helps the analyst figure out what is actually happening. And it is the quality and experience of the analyst—her tacit knowledge—that allows her to create an authoritative picture.

The best intelligence analysts are really applied anthropologists. They study a particular culture, in a particular place, that works differently in practice than it does on paper. Cultures are often characterized by a "hidden structure" that is largely invisible to outsiders and sometimes poorly understood even by insiders. Many cultures actually develop a lack of transparency precisely to prevent comprehension by outsiders. Discovering how a culture works requires one to create networks of informants and use research methods such as participant observation. This requires trust, which may take years to develop.

What marks a good intelligence analyst, and a good grantmaker in the field of advocacy, is the ability to penetrate those opaque surfaces to detect patterns of influence. Foundations engaged in advocacy need to build this capacity internally, strive for substantial continuity (and thus institutional memory) among those who possess these skills, and respect the value of trained, subjective judgment in making key decisions.

The characteristic features of the terrain of politics—chaotic agenda setting, pervasive misinformation, overlapping responsibility—mean that no one metric can capture the reality of influence. Donors do themselves a disservice by even looking for one. Only by trying to make sense of policymaking activity through the simultaneous application of multiple ways of knowing can donors get closer to finding out what they need to know. ■

Notes

- 1 Verta Taylor, "Social Movement Continuity: The Women's Movement in Abeyance," *American Sociological Review*, 1989: 761–75.
- 2 Eric Patashnik, *Reforms at Risk*, Princeton, N.J.: Princeton University Press, 2008.