



Cornell University
ILR School

Cornell University ILR School
DigitalCommons@ILR

Conference Proceedings, Presentations, and
Speeches

ILR Collection

1-2007

Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs

David B. Lipsky
Cornell University, DBL4@CORNELL.EDU

Ronald L. Seeber
Cornell University, rs60@cornell.edu

Ariel C. Avgar
Cornell University

Rocco M. Scanza
Cornell University

Follow this and additional works at: <http://digitalcommons.ilr.cornell.edu/conference>

Thank you for downloading an article from DigitalCommons@ILR.

Support this valuable resource today!

This Article is brought to you for free and open access by the ILR Collection at DigitalCommons@ILR. It has been accepted for inclusion in Conference Proceedings, Presentations, and Speeches by an authorized administrator of DigitalCommons@ILR. For more information, please contact hlmdigital@cornell.edu.

Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs

Abstract

[Excerpt] The growth of alternative dispute resolution (ADR) has been one of the most significant developments in the U.S. workplace in the past twenty-five years. There is a significant and growing body of research tracking the development of ADR in U.S. employment relations, its effects on organizations and workers, and its implications for the community of neutrals and the providers of neutral services (Lipsky, Seeber, and Fincher 2003; Seeber and Lipsky 2006; Bingham and Chachere 1999; Bingham 2004; Colvin et al. 2006; Lewin 2004). The intense debates that have arisen over the desirability of ADR have caused both practitioners and researchers to recognize the need for the evaluation of all types of ADR programs, including those mandated by the courts, statutes, and other public policies and those established by private sector organizations (Seeber, Schmidle, and Smith 2001; Lipsky and Seeber 2006; Bingham 2004).

The authors of this paper have conducted numerous evaluations of ADR programs in both the public and private sectors, and it has been our experience that the desire of the evaluators and the program sponsors to have an impartial and objective evaluation of a program has often been frustrated by political considerations. This paper will focus on the politics of the evaluation of ADR systems and programs. We maintain that there are three types of political factors affecting ADR programs: One type involves ideological and policy debates about the desirability of ADR; a second type involves political factors within an organization (whether private or public) that affect the adoption, implementation, and maintenance of an ADR program; and the third type is the political struggle that can sometimes emerge between the managers and practitioners who sponsor ADR programs and the academics and consultants who evaluate them. Here we concentrate particularly on the differences that arise between program sponsors and outside evaluators. Academic evaluators, for example, value the so-called purity of their research and strive to conduct evaluations consistent with accepted social science standards; program sponsors and administrators have evaluation objectives that are much more instrumental and pragmatic. All parties in an evaluation may very well have legitimate objectives, but political differences can arise out of the incompatibility of those objectives, an incompatibility that is often the consequence of the "clash of cultures" between academics and practitioners.

In our view, political factors will invariably influence program evaluation. It is clear that some of the political differences that affect the evaluation of ADR programs (for example, ideological debates) are beyond the control of either the program sponsors and administrators or the evaluators. But there are other political factors that the parties can potentially manage, or at least influence. On the one hand, political differences can threaten the integrity of an evaluation. On the other hand, not all political factors have negative consequences for an evaluation. The trick for sponsors and evaluators alike is, first, to recognize the political factors they can control and, second, to distinguish between those that have positive effects on the evaluation and those that have negative effects.

Keywords

alternative dispute resolution, ADR, conflict, industrial relations, policy

Comments

Suggested Citation

Lipsky, D. B., Seeber, R. L., Avgar, A. C., & Scanza, R. M. (2007). Managing the politics of evaluation: Lessons

from the evaluation of ADR programs [Electronic version]. *Proceedings of the Fifty-Ninth Annual Meeting of the Labor and Employment Relations Association*, 116-129.

Required Publisher's Statement

© Labor and Employment Relations Association. Reprinted with permission. All rights reserved.

Managing the Politics of Evaluation: Lessons from the Evaluation of ADR Programs

**David B. Lipsky, Ronald L. Seeber, Ariel C. Avgar,
and Rocco M. Scanza**
Cornell University

The growth of alternative dispute resolution (ADR) has been one of the most significant developments in the U.S. workplace in the past twenty-five years. There is a significant and growing body of research tracking the development of ADR in U.S. employment relations, its effects on organizations and workers, and its implications for the community of neutrals and the providers of neutral services (Lipsky, Seeber, and Fincher 2003; Seeber and Lipsky 2006; Bingham and Chachere 1999; Bingham 2004; Colvin et al. 2006; Lewin 2004). The intense debates that have arisen over the desirability of ADR have caused both practitioners and researchers to recognize the need for the evaluation of all types of ADR programs, including those mandated by the courts, statutes, and other public policies and those established by private sector organizations (Seeber, Schmidle, and Smith 2001; Lipsky and Seeber 2006; Bingham 2004).

The authors of this paper have conducted numerous evaluations of ADR programs in both the public and private sectors, and it has been our experience that the desire of the evaluators and the program sponsors to have an impartial and objective evaluation of a program has often been frustrated by political considerations. This paper will focus on the politics of the evaluation of ADR systems and programs. We maintain that there are three types of political factors affecting ADR programs: One type involves ideological and policy debates about the desirability of ADR; a second type involves political factors within an organization (whether private or public) that affect the adoption, implementation, and maintenance of an ADR program; and the third type is the political struggle that can sometimes emerge between the managers and practitioners who sponsor ADR programs and the academics and consultants who evaluate them. Here we concentrate particularly on the differences that arise between program sponsors and outside evaluators. Academic evaluators, for example, value the so-called purity of their research and strive to conduct evaluations consistent with accepted social science standards; program sponsors and administrators have evaluation objectives that are much more instrumental and pragmatic. All parties in an evaluation may very well have legitimate objectives, but political differences can arise

out of the incompatibility of those objectives, an incompatibility that is often the consequence of the "clash of cultures" between academics and practitioners.

In our view, political factors will invariably influence program evaluation. It is clear that some of the political differences that affect the evaluation of ADR programs (for example, ideological debates) are beyond the control of either the program sponsors and administrators or the evaluators. But there are other political factors that the parties can potentially manage, or at least influence. On the one hand, political differences can threaten the integrity of an evaluation. On the other hand, not all political factors have negative consequences for an evaluation. The trick for sponsors and evaluators alike is, first, to recognize the political factors they can control and, second, to distinguish between those that have positive effects on the evaluation and those that have negative effects.

Evaluation Research at the Institute on Conflict Resolution

The authors of this paper are affiliated with the Institute on Conflict Resolution (ICR) at Cornell University. Since ICR's founding in 1996, we have not only conducted research on ADR programs and systems in numerous private and public sector organizations; we have also been called upon to assist in the design, implementation, and evaluation of ADR programs in some of those organizations. For example, we have evaluated ADR programs for the New York State Workers' Compensation Board, the Unified Court System of the State of New York, the U.S. Department of Labor, the U.S. Equal Employment Opportunity Commission (EEOC), the U.S. Nuclear Regulatory Commission, and several private sector organizations (Lipsky, Seeber, and Fincher 2003; Seeber et al. 2001; Lipsky et al. 2003; Lipsky, Scanza, and Avgar 2006). We believe our experiences have provided us with insights about the politics of program evaluation, and in this paper we attempt to translate those insights into analytical and prescriptive terms. Admittedly, our paper is written from the standpoint of outside evaluators, but we think we have learned some lessons that are valuable not only for outside evaluators but also for program sponsors and administrators. In the more analytical portion of the paper we deal with both the substantive and process issues that invariably arise in every program evaluation and how political factors can affect those issues. In the prescriptive part of the paper we offer advice on how evaluators and sponsors can manage and derive benefits from the political factors that they are capable of influencing. We will use illustrations drawn from the evaluations we have conducted, in some cases naming the organization in question but in other cases purposely maintaining the anonymity of the organization.

Pressures for Program Evaluation

Research suggests that many organizations turned to ADR because they were certain it would save the time and money associated with litigation and maintain management's control over dispute resolution (Lipsky and Seeber 1998, Colvin 2003). We found that many major corporations adopted ADR policies after they had endured a crisis (a multimillion dollar lawsuit, for example) and a "champion" within the organization persuaded top management that ADR was preferable to alternative means of handling disputes (Lipsky, Seeber, and Fincher 2003). In the early days many of these organizations gave little thought to the systematic evaluation of the ADR processes they had adopted. Most of the organizations we have studied have been content to install tracking software that produces basic information about the operation of the ADR program but falls far short of providing the data needed for a comprehensive evaluation. Over time, however, managers in many corporations demanded more evidence that there was a good business case for using ADR; they increasingly wanted to know whether ADR programs produced an acceptable return on investment (ROI) and were cost effective. It was easy for a chief financial officer (CFO), for example, to observe that an ADR program added costs to the company's operations (usually in the form of additional staff), but it was difficult for the CFO to tell whether the ADR program produced tangible benefits. Top managers demanded accountability and consequently required that more careful evaluations be undertaken of their ADR programs.

A parallel development occurred in public sector agencies. In New York State, for example, the executive and legislative branches of the state government had searched for many years for means of reforming the state's multibillion dollar workers' compensation system. A bipartisan consensus was reached that the use of ADR, along with other reforms, could potentially save the state's employers and taxpayers millions of dollars. But the reform bill's sponsors realized that the controversial reforms they favored would only gain broad acceptance if it could be proven to all stakeholders that they were actually effective. Consequently, the legislators inserted into the reform bill a provision calling for Cornell University to conduct a systematic evaluation of the ADR experiment (Seeber et al. 2001).

Other lawmakers, at both the federal and state levels, have mandated that ADR programs be evaluated by impartial outsiders. The policy makers' growing insistence on evaluation is itself a response to the political pressures they face. By supporting the use of ADR, they risk alienating important constituencies that are often opposed to its use (for example, the union movement, the plaintiffs' bar, etc.). Inserting an evaluation requirement in a statute is a means of legitimizing the use of ADR and an attempt to remove it from the political fray. Legislators can claim, perhaps truthfully, that they seek transparency in the use of public funds and will support only those programs that are proven, by means of careful evaluation, to be truly effective.

Criteria for an Ideal Evaluation: A Clash of Cultures

The methodological standards and evidentiary criteria that are essential for a sound evaluation are well understood and accepted within the academy. In most of the organizations in which we have worked, however, we found that the managers and administrators of the programs we evaluated seldom had a firm grasp of these standards and criteria. Most social science researchers, beginning with their Ph.D. training, are immersed in the technical aspects of program evaluation, and program administrators cannot be faulted for pursuing a different career path. The sponsors and administrators of ADR programs have more pragmatic and instrumental objectives. They prefer evaluations that produce results that can be easily understood by all the stakeholder groups that need to be satisfied, and they seek recommendations that can be translated into practice.

There are four principal differences that distinguish academic culture from practitioner culture. First, academic researchers are taught to appreciate the value of experimental design and whenever possible want to conduct evaluations that have both "treatment" groups and "control" groups. Although pure experiments are actually quite rare in the social sciences, researchers strive to design an evaluation that closely mirrors an authentic experiment. In the evaluations we have conducted, we have almost always made efforts to set up an appropriate control group so that we could more precisely determine the effects of the "treatment"—the ADR program—we were evaluating. Occasionally we have succeeded, but usually practical considerations limited the design of our evaluation.

For example, in the evaluation we conducted of the New York State Workers' Compensation system, we were able to arrange for a reasonable control group (Seeber et al. 2001). But in our evaluation of a pilot ADR program conducted by the Solicitor's Office in the U.S. Department of Labor, we were never able to do so. In that project a mediation program was introduced in every one of the Solicitor's regional offices, and so there were no obvious regions or offices that could serve as a contemporaneous control group. We discovered in the course of doing our evaluation, however, that the use of mediation varied systematically across regions, and we were able to make use of that variation to explain the effectiveness of the pilot program. The variation was explained in part by the commitment of the regional solicitor to the program and in part by the nature of the cases handled in the region (Lipsky et al. 2003).

A second broad area of potential conflict between researcher and client is the extent to which there is a need for statistical rigor. Although the value of data and evidence is apparent to most policy makers, they rarely appreciate the need to collect the appropriate independent and intermediary variables necessary to create the

statistical rigor demanded by evaluators. Thus, the evaluators' desire to collect appropriate data through surveys, questionnaires, and data from already existing organizational records is seldom greeted with enthusiasm by organizations. In every evaluation study in which we have been involved, we have driven the data collection process much further than the sponsors thought necessary or useful.

Third, the integrity of an evaluation, in the view of academic researchers, depends on several factors relating to the evaluators' ability to maintain their independence and impartiality. For example, independence and impartiality depend in part on the evaluators' ability to produce a final report that is based entirely on the facts and evidence obtained. Academic researchers essentially strive to satisfy two audiences. One audience consists of the organizations that have sponsored and paid for the evaluation. The other audience consists of the researchers' colleagues and peers. To a degree, the objectives and interests of these two audiences do not correspond. Sponsors of an evaluation are seldom interested in the finer points of evaluation methodology, whereas academics'—their ability to get promoted and gain tenure—will depend on their ability to convince colleagues and peers that they have satisfied the highest standards of scholarly research.

Finally, academics are committed to reporting all the relevant facts and evidence, good and bad, and they believe withholding any important findings potentially compromises the integrity of the evaluation. The academic culture frowns on hyperbole, while organizational clients, in our experience, nearly always wish to emphasize the best features of their program and to deemphasize the worst. We do not mean to say that they are dishonest or intend to cover up bad results. Rather, the differences are often the consequence of the differences in the time perspective of evaluators and clients. When academics submit a final report to the organization, they usually sever their relationship with the organization, whereas responsible organizational officials must live with the consequences of the evaluation. They need to worry about the effect of the evaluation not only on the future of the program that has been evaluated but also on other programs, people, and policies.

The differences between evaluators and clients regarding an ideal evaluation represent largely a difference in cultures between the two groups. These cultural differences often result in the researcher questioning the sponsor's organizational commitment to a sound evaluation, while the sponsor views the evaluators as overly obsessed by data and analysis that they view as useless or irrelevant. In the end, these cultural differences nearly always require negotiation between the evaluators and the sponsors over precisely the level of organizational support for the evaluation.

The Unavoidable and Inherent Influence of Politics in Evaluations

There is little doubt about a fundamental source of political influence in program evaluations. Behind the introduction of an ADR program—whether it is a consequence of a change in public policy or is an innovation introduced by managers—is a group of individuals who believe that the program will have positive effects for the organization, for organizational outcomes the individuals value, or possibly for the individuals themselves. The individuals that trigger the innovation always have a stake in the success of the program they have sponsored. They are naturally concerned about the operation and effectiveness of the program and, therefore, the results of an evaluation. Here we need to distinguish between the *politics of policy change* and the *politics of organizational change*. The politics of policy change result from differing political and ideological views regarding competing public policies and their implementation. The politics of organizational change result from the differing views of organizational members and stakeholders over the merits of organizational innovations. The former pertains principally to politics outside the organization and the latter to politics inside the organization. For the evaluator both sources of political influence can threaten the integrity of an evaluation. But an evaluator is well advised to recognize the distinction between these two sources of political tension. There may be little or nothing evaluators can do about external political factors, but they may be able to limit and channel internal political factors.

ADR programs almost always pit some organizational stakeholders against others, and this can result in frequently intense turf wars. For example, in one organization we studied, the "champion" of the ADR program was an attorney in the counsel's office. After persuading his superiors to adopt the program, he became its director. In that role he clashed repeatedly with managers in the human resource function who thought he was usurping their prerogatives. By almost any measure, the ADR program was a success, but eventually the director of the program was forced to leave the organization. In our research we also discovered that middle managers occasionally resist the introduction of ADR programs favored by top managers because they fear that such programs will undermine their authority (Lipsky, Seeber, and Fincher 2003). Employees and their unions (if they are represented) are often skeptical about the value of ADR programs and do not share management's interest in the effectiveness of such programs. In a corporation stockholders and other external stakeholders can have a very different focus from internal stakeholders. In a public agency ADR programs may pit the interests of legislators against the interests of agency managers (Lipsky, Seeber, and Fincher 2003).

Moreover, ADR programs operate in environments rife with uncertainty and ambiguity. All of the programs we have evaluated were begun under conditions in which not all the critical procedural and policy questions about the implementation of

the program had been answered. For example, we have entered organizations in which the program sponsors and administrators could not clearly specify the objectives they hoped the program would achieve. In more than one organization, we suspect, the objectives remained ambiguous because the internal stakeholders themselves could not agree on them and purposely kept them vague. In some organizations the establishment of the program itself represents "success" for the principal stakeholders, and whether the program actually achieves anything of value is of secondary importance. Thus, evaluators inevitably operate in a politically charged environment requiring negotiation over the terms and conditions of their engagement.

This is not to say that the political nature of ADR program evaluation is all bad. There is a spectrum of political intensity that ranges, on the one end, from politics that is normal, expected, and even possibly constructive to, on the other end, politics that is unhealthy, destructive, and jeopardizes the integrity of an evaluation. The containment of unhealthy political factors is critical to the success of the evaluation, and it is a central task of both sponsors and evaluators to minimize these potentially destructive political influences. Ultimately, the evaluators at least must seek to gain the stakeholders' respect and their commitment to the evaluation process.

Negotiating the "Bread-and-Butter" Issues in Evaluation

In every evaluation the critical issues for negotiation can be reduced to four factors: the scope and clarity of the evaluators' mandate, the standards and criteria used for the evaluation, the minimization of conflicts of interest, and the establishment of buffers between the evaluators and the sponsors of the evaluation.

The mandate for the evaluations we have conducted varied from project to project. For example, the mandate for our evaluation of the New York State Workers' Compensation ADR program was contained in the statute creating the program (Seeber et al. 2001). The mandate for our evaluation of the ADR program administered by the Solicitor's Office of the U.S. Department of Labor was the consequence of an interagency agreement (Lipsky et al. 2003). The mandate for our evaluation of the EEOC's internal dispute resolution program resulted from the desire of the program's principal administrators to find out whether the program was (by their definition) an effective one (Lipsky et al. 2006). In the case of the workers' compensation program, the statutory mandate gave us a tool we could use to control the agency's desire to alter the scope of our evaluation. By contrast, in the case of the Solicitor's ADR program, the multiparty interagency agreement had not sufficiently clarified our mandate and was subject to continuing renegotiation and interpretation. In the case of the EEOC's program, our mandate was specified in a contract between the agency and Cornell. However, disagreements arose during the course of the project regarding the interpretation of some of the key provisions in the contract. It

may be an obvious truism to say that it is extraordinarily helpful to specify as precisely as possible the scope of the evaluators' mandate before the evaluation actually begins. But experience demonstrates (both in program evaluation and, we might add, in labor relations) that even the most careful attempts to negotiate the mandate in advance of the evaluation will not necessarily prevent disagreements from arising later.

A second topic often requiring negotiation between the sponsors and the evaluators centers on the appropriate standards and criteria the researchers will use to evaluate the ADR program. For example, in one of our evaluations the sponsors were quite content to limit the evaluation to rudimentary measures of participant satisfaction with the program, whereas we maintained that the measures they preferred would not be sufficient to allow them to understand critical dimensions of their program's operation. The sponsors also did not want us to collect basic demographic data about the users of the program. We pointed out that program satisfaction was likely to vary significantly by the gender, age, and race of the users, but nevertheless the sponsors insisted on a narrow definition of the criteria used for the evaluation. We negotiated these issues at some length, but in the end we had to yield to the sponsors' wishes. We did our best to produce the type of evaluation the sponsors wanted, but clearly the standards and criteria we used fell far short of meeting the textbook definition of a valid evaluation. The important point is not that the evaluation failed to meet our standards; rather, the important point is that in our judgment the evaluation failed to serve the sponsors' purposes.

A third topic necessitating negotiation between the sponsors and the evaluators centers on the need to minimize conflicts of interests that might undermine the integrity of the evaluation. Of course, even the appearance of a conflict of interest can undermine the evaluation. For many evaluations the sponsor or a funding agency issues a "request for proposals" (RFP), and the uninitiated may assume that each bidder has a fair chance of winning the competition based on the merit of the bidder's proposal. It is no secret, however, that many evaluation grants and contracts are "wired." That is, the RFP has been tailored to give an advantage in the competition to a preferred bidder, virtually guaranteeing the bidder will obtain the grant or contract. An alternative approach is for the sponsors to arrange for the evaluation to be done on the basis of so-called sole-source funding. We acknowledge that we have occasionally benefited from such arrangements. In some situations the pool of qualified evaluators may be very small, therefore justifying the targeted nature of the bidding process. Some sponsors may prefer to turn to a particular institution for an evaluation because the institution's reputation for quality work is well known and the sponsors want to take advantage of the institution's "brand name." Whether these considerations justify wired RFPs and sole-source funding is of course problematic. To say the least,

maintaining the independence and impartiality of the evaluation can be challenging if the evaluators have too cozy a relationship with the sponsors. Sponsors who channel grants and contracts to favored evaluators may have an expectation of reciprocity, usually in the form of the evaluators' "seal of approval."

We always do our utmost to protect our own integrity, and we assume the vast majority of evaluators do the same. The importance of protecting one's integrity is not only a matter of ethical principle; it is also a pragmatic consideration as well. In the long run we realize we will lose our reputations if we compromise the integrity of our work. We also know we will be judged by our colleagues and peers in the research community, and our careers will be damaged if they come to believe that we have crossed the boundary of accepted practice. The best response to these issues is, again, to negotiate them up front, preferably before the grant or contract is awarded. The preferred approach, we believe, is for evaluators to make it clear to the sponsors that they intend to conduct an independent and impartial evaluation, and to disabuse them of any notion they may have of receiving special favors because no other evaluators were considered for the project.

This discussion illustrates a common problem in program evaluation, namely, the lack of "buffers" between the evaluators and the sponsors. Some might argue that, in a perfect world, a firewall should separate the evaluators from the sponsors. But the evaluation of ADR programs requires communication between the evaluators and the sponsors on a continuing basis. In many ways, the best evaluation is a collaborative effort, characterized by ongoing interactions among the evaluators, sponsors, program administrators, and other stakeholders. All those interactions, however, offer opportunities for the stakeholders to put pressure on the evaluators in order to slant their work to suit the stakeholders' interests. The dilemma is to protect the evaluators from these political pressures while at the same time allowing them to have adequate communication with the sponsors and other stakeholders.

The solution is to create devices that serve as buffers. A contract is one device that can serve as a buffer, provided the contract adequately defines the parties' expectations and the limits on what is and is not permissible in the evaluation. We believe the parties should give consideration to specifying in the contract the means they will use to resolve disputes over the interpretation of the contract—a device that has the added benefit of mimicking a common feature of an ADR program. (We acknowledge that the contract office at Cornell would not approve a contract provision that delegated disputes over contract interpretation to third parties.)

Another buffer is to involve third parties in various phases of the evaluation process. For example, the sponsors and evaluators may agree to have an advisory board or council oversee the evaluation. Such entities were established for some of the

projects we conducted, including our evaluation of the workers' compensation ADR program and our evaluation of the Solicitor's program in the Department of Labor. In the former case the advisory board did an admirable job of serving both the sponsors' interests and our own; in the latter case the advisory body became the arena for intense debates and squabbling. Nevertheless, it was preferable to have these conflicts surface in the advisory body rather than in our direct relationship with the Department of Labor.

For some of our projects surveys and interviews were conducted "in house" by faculty and staff associated with the Institute on Conflict Resolution. For other projects the scale of the evaluation necessitated the use of Cornell's Survey Research Institute (SRI). A by-product of subcontracting survey work to SRI was that SRI served as an effective buffer between the sponsors, its members, and us. In effect, we could rely on SRI's expertise, and not just our own, to deflect pressures emanating from the sponsor during the course of the evaluation.

Throughout our discussion we have referred to the so-called sponsors of the evaluation. However, it is often not a straightforward matter to identify the sponsor. For example, the nominal sponsor of our evaluation of the EEOC program (and the signatory to the contract) was the agency itself. However, we needed to deal principally with two individuals who were the de facto sponsors: one was the director of the ADR program and the other was the special assistant to the chair of the commission. We admire greatly both individuals, but in time we discovered that they did not necessarily share the same goals and interests. The director of the ADR program understandably focused on what she perceived to be the best interests of her program, whereas the special assistant to the chair, for equally understandable reasons, focused on what she considered to be the best interests of her boss and of the agency itself. Arguably, our two principal contacts at the EEOC were not the only "sponsors" of our evaluation; the entire commission and the senior staff of the agency were also sponsors, although we had little or no contact with them. In the evaluations we have conducted the sponsors were almost always a group of individuals who had a vital stake in the success or failure of the program we were evaluating.

Ultimately, it is up to the evaluators to resist political pressures that threaten to compromise the integrity of their evaluation. The more experiences we have accumulated, the more we have realized the importance of negotiating the issues that are the potential source of disputes at the beginning of the evaluation process. Resolving key "bread-and-butter" issues before the evaluation begins can be the key to the success of the evaluation. Although up-front negotiation is essential, we believe, disagreements are bound to arise in the course of the evaluation. When these disagreements threaten to compromise the evaluation, it is critical for the evaluators to

hold their ground and rely on their own professional standards as well as the agreements they reached with the sponsors at the start of the project.

The Cassandra Effect: Delivering Bad News

The last step in an evaluation is the preparation and delivery of the final report. Some final reports are "dead on arrival"—filed away by the sponsor and never seen again. For some organizations the utility of the evaluation is not related to the evaluators' substantive findings and recommendations but rather to the fact that an evaluation was conducted in the first place. Most of the evaluation projects we have conducted, however, have had substantive and not merely symbolic significance. In the best of all possible worlds, the substantive findings allow the evaluators to write a final report that satisfies the needs and objectives of the program's sponsors and stakeholders. In our experience, however, the findings rarely result in a final report that contains only good news for the sponsors. Even the most favorable evaluations we have conducted identified some program weaknesses that needed to be addressed.

For example, our evaluation of the Department of Labor program demonstrated that the program was clearly achieving its objectives; the benefits of using mediation in enforcement cases, for example, clearly exceeded the costs, and virtually all users of the program were highly satisfied with both the mediation process and its outcomes. However, in our final report we offered some constructive recommendations for improving the administration of the program (Lipsky et al. 2003). Although our recommendations were received in the spirit in which they were intended, we nevertheless experienced some degree of discomfort in delivering them to the people responsible for administering the program. In other evaluations we had bad news to report to the sponsors. Evaluators need to recognize that bad news can often create serious political problems for the sponsor. The survival of the program may be at stake, and the careers of the individuals who administer those programs can be seriously affected.

In one evaluation we discovered that many complainants using the ADR program feared retaliation; we were surprised to learn that many of the managers and supervisors who usually were the respondents in these complaints also feared retaliation—that is, they were afraid that they would be penalized for having complaints filed against them. We knew these findings would be regarded as truly bad news by the evaluation's sponsors. The problem was not inherent in the ADR program itself, nor was it the fault of the program's administrators. Rather, the source of the problem seemed to be in the culture of the organization. Yet reporting our findings could jeopardize the future of the ADR program, and we truly believed that the best interests of the organization and its employees would be served if the ADR program survived. We pondered at length how to report our findings to the sponsors. Could we

possibly sugar-coat the truth and maintain our integrity? Were we prepared to take responsibility for damaging the careers of the program's sponsors, especially when we did not believe the sponsors were responsible for the problems we discovered? Before we submitted our final report, we had private sessions with the sponsors and gave them an honest account of our findings. The findings naturally disturbed the sponsors, and they did identify some of the limitations in our analysis. We agreed to modify our final report in ways that we believed did not compromise its fundamental integrity. Achieving acceptability of our findings in this and other projects was not a matter of covering up or varnishing the truth but rather was a matter of properly "framing" the results (Tversky and Kahneman 1981). Subsequently, the organization undertook training programs that helped to improve supervisory practices and did, to some degree, address the problems we identified in our evaluation.

Cassandra was executed for telling the truth. In the evaluations we have conducted, we have always been conscious of the possibility that we could pay a price for telling the truth. If delivering bad news results in the demise of a program, it is not only the sponsors and administrators who may suffer as a result: the evaluators' reputations and prospects for future work may also be affected. The likelihood of evaluators receiving grants and contracts for future work can be diminished if potential sponsors learn that the evaluators had conducted projects that damaged the interests of previous sponsors. But covering up the truth can be equally detrimental to the evaluators. Evaluators who discover that a program has serious weaknesses but nevertheless declare the program to be a success need to anticipate that eventually "the truth will come out." The short-term gain of satisfying the sponsors' needs and desires will be far outweighed by the longer-term harm to the evaluators' credibility in both the scholarly and the practitioner communities. On both ethical and practical grounds, honesty is the best policy.

Conclusion: Managing Expectations

In this paper we have identified some of the political issues that can affect the evaluation of ADR programs, and we have suggested some of the means sponsors and evaluators can use to deal with those issues. We believe political pressures can be minimized but never fully eliminated. Both sponsors and evaluators should recognize that evaluations involve multiple stakeholders, and those stakeholders are likely to have different objectives, values, and needs that make the evaluation process an inherently political one. Evaluators should certainly strive to conduct their work in a fashion that conforms to the standards of social science research, but maintaining the "purity" of their evaluations is likely to be more the ideal than the reality. Evaluators and sponsors almost invariably are required to negotiate issues that are critical to the design and execution of the evaluation. Negotiations almost always result in

compromises, but in evaluation work as in other walks of life, compromise should not mean the sacrifice of important principles.

It is better for evaluators and sponsors to launch an evaluation fully aware of the potential political pitfalls they are likely to encounter than to proceed naively and face unanticipated problems in the midst of the project. Our experience has taught us that dealing effectively with political issues is largely an exercise in managing expectations. Sponsors and evaluators should understand that it is in their mutual interests to deal with issues that may cause political problems up front. A joint problem-solving approach at the start of the project can help avoid an adversarial relationship later in the project.

We hold the conviction that most organizations that have sponsored evaluations are better off for having done so. In the organizations in which we have worked, attempts to minimize the importance of the political aspects of the evaluation allowed us to produce better studies than we might have otherwise. Resolving political issues up front helps, but both evaluators and sponsors need to understand that they are still likely to face political issues that arise in the course of the evaluation process that they could not possibly anticipate at the start. Approaching a project in a collaborative, problem-solving fashion, however, makes it easier to deal with political issues that may arise later.

References

- Bingham, Lisa B. 2004. "Employment Dispute Resolution: The Case for Mediation." *Conflict Resolution Quarterly*, Vol. 22, nos. 1&2, pp. 145&174.
- Bingham, Lisa B., and Denise R. Chachere. 1999. "Dispute Resolution in Employment: The Need for Research." In Adrienne E. Eaton and Jeffrey H. Keefe, eds., *Employment Dispute Resolution and Worker Rights in the Changing Workplace*. Champaign, IL: Industrial Relations Research Association, pp. 95&135.
- Colvin, Alex J. 2003. "Institutional Pressures, Human Resource Strategies, and the Rise of Nonunion Dispute Resolution Procedures." *Industrial and Labor Relations Review*, Vol. 56, no. 3 (April), pp. 375&392.
- Colvin, Alexander J. S., Brian Klaas, and Douglas Mahony. 2006. "Research on Alternative Dispute Resolution Procedures." In David Lewin, ed., *Contemporary Issues in Employment Relations*. Champaign, IL: Labor and Employment Relations Association, pp. 103&147.

Lewin, David. 2004. "Dispute Resolution in Nonunion Organizations: Key Empirical Findings." In Samuel Estreicher and David Sherwyn, eds. *Alternative Dispute Resolution in the Employment Arena*. New York: Kluwer, pp. 379-403.

Lipsky, David B., Rocco M. Scanza, and Ariel C. Avgar. 2006. *Final Report: The U.S. Equal Employment Opportunity Commission Internal Dispute Resolution Program: An Evaluation of RESOLVE*. Submitted to the Equal Employment Opportunity Commission, March.

Lipsky, David B., and Ronald L. Seeber. 1998. "In Search of Control: The Corporate Embrace of ADR." *Journal of Labor and Employment Law*, Vol. 1, no. 1 (Spring), pp. 133-157.

Lipsky, David B., and Ronald L. Seeber. 2006. "Managing Organizational Conflicts." In John Oetzel and Stella Ting-Toomey, eds., *The Sage Handbook of Conflict Communication*. Thousand Oaks, CA: Sage Publications, pp. 359-390.

Lipsky, David B., Ronald L. Seeber, and Richard Fincher. 2003. *Emerging Systems for Managing Workplace Conflict: Lessons from American Corporations for Managers and Dispute Resolution Professionals*. San Francisco, CA: Jossey-Bass.

Lipsky, David B., Ronald L. Seeber, Les Hough, and Rocco Scanza. 2003. *Final Report: Developing a Mediation Program for the U.S. Department of Labor: An Evaluation of a Pilot Program for Enforcement Cases*. Submitted to the U.S. Department of Labor, August.

Seeber, Ronald L., and David B. Lipsky. 2006. "The Ascendancy of Employment Arbitrators in U.S. Employment Relations: A New Actor in the American System?" *British Journal of Industrial Relations*, Vol. 44, no. 4, pp. 719-756.

Seeber, Ronald L., Timothy P. Schmidle, and Robert S. Smith. 2001. *An Evaluation of the New York State Workers' Compensation Pilot Program for Alternative Dispute Resolution*. Submitted to the New York State Workers' Compensation Board, December.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology Choice." *Science*, Vol. 40, pp. 453-463.