

Environmental Conflict Resolution: Evaluating Performance Outcomes and Contributing Factors

KIRK EMERSON

PATRICIA J. ORR

DALE L. KEYES

KATHERINE M. MCKNIGHT

This empirical study of fifty-two environmental conflict resolution (ECR) processes is based on an evaluation framework that specifies key conditions and factors that contribute to ECR outcomes. Data were collected on a range of ECR processes and applications. This article reports on findings from a multilevel modeling analysis that focuses on three primary outcomes: reaching agreement, the quality of agreement, and improved working relationships among parties. Effective engagement of parties is identified as a major contributor to all three outcomes. Other key factors that operate directly and indirectly through effective engagement are involvement of appropriate parties, the skills and practices of ECR mediators and facilitators, and incorporation of relevant and high-quality information. Findings generally support the ECR evaluation framework.

NOTE: This article could not have been written without the financial support of the William and Flora Hewlett Foundation and the U.S. Institute for Environmental Conflict Resolution, a program of the Morris K. Udall Foundation, an independent federal agency. The project support and partnership of the Collaborative Action and Dispute Resolution of the U.S. Department of Interior, Conflict Prevention and Resolution Center of the Environmental Protection Agency, the Federal Highway Administration, USDA Forest Service, Oregon Consensus Program, Policy Consensus Initiative, and practitioners Dick La Fever, Terry Stimson, Harry Seraydarian, Debra Drecksel, and Lisa Beutler were also instrumental.



Public managers make daily choices about how to manage conflict, not only within and between public agencies but with other levels of government and among stakeholders with competing interests in specific public policy decisions. Since the 1990s, federal and state policies and institutions have encouraged use of mediation and other forms of alternative dispute resolution to address challenging public policy conflicts (Mills, 1991; Singer, 2004). In the environmental and natural resources management arena, an array of tools and processes have emerged since the passage of the National Environmental Policy Act in 1969 to enable the public and affected stakeholders to influence public environmental decision making and address environmental challenges more constructively (National Environmental Conflict Resolution Advisory Committee, 2005).

Environmental conflict resolution (ECR) comprises a subset of those tools and processes to prevent, manage, and resolve conflicts involving environmental quality or natural resources management. In its broadest terms, ECR may be characterized as third-party assisted deliberations over environmental issues among affected parties, intended to arrive at the best possible mutually beneficial outcomes. ECR has been used in settings where broad or deep disagreement is anticipated, has begun to escalate, or has already led to impasse. ECR has also been drawn on to build agreement or reach consensus among multiple parties over policy development, planning, and complex regulatory negotiations. For example, ECR has been used to assist in enforcing cleanup standards on industrial sites, mitigating the impact of a new highway on wetlands, planning for the recovery of an endangered species or an ecosystem, developing pollution control rules for sources of air pollution, and managing off-road vehicle use on public lands. In these kinds of situations, where people are trying to reach agreement or resolve a dispute, ECR facilitators or mediators bring impartiality and independence to their role in helping people negotiate their differences (Bingham, 1986; Emerson, Nabatchi, O'Leary, and Stephens, 2003; O'Leary and Bingham, 2003).

Because ECR requires stepping beyond business as usual, public managers are looking for evidence to inform their management decisions before embarking on an ECR process. They want to know (1) what they can reasonably expect to achieve through ECR and (2) what they would need to do, provide, or prepare for, to ensure the most positive results in an uncertain, often conflictual situation (see, for example, Kessler, 2004; Irwin and Stansbury, 2004). This article attempts to address these two linked questions.

For the past twenty-two years, since Gail Bingham's landmark empirical study of environmental mediation cases (Bingham, 1986), researchers and practitioners have been studying, theorizing, and practicing ECR (Dukes, 2004b). Many dimensions of successful outcomes have been posited or attributed to ECR (D'Estree and Colby, 2004; Leach, 2007; Beierle and Cayford, 2002; Armbruster, 2008). In one recent study of 111 collaborative groups in the Intermountain West, twenty-two indicators of success were developed (McKinney and Field, 2006). D'Estree and Colby (2004) presented twenty-eight indicators of successful outcomes for resolving water disputes.

We have learned some things about ECR performance from case studies, comparative case analyses, and the growing number of multicase research projects (O'Leary and Bingham, 2003). For example, we have learned that mediation generally gets high marks from parties on satisfaction with the process and somewhat lower ratings on satisfaction with the outcomes (Coglianese, 2002; Dukes, 2004b). We know that the agreement rate for environmental and natural resource conflict resolution ranges considerably, from 61 to 93 percent depending on who is reporting and when the data are collected (Bingham, 1986; Buckle and Thomas-Buckle, 1986; Sipe and Stiffler, 1995; Sipe, 1998; Susskind, McKernan, and Thomas-Larmer, 1999; Andrew, 2001; O'Leary and Husar, 2002; Berry, Stiffler, and Dedekorkut, 2003; Frame, Gunton, and Day, 2004). There are increasingly sophisticated, albeit not widely used, tools for measuring the quality of ECR agreements (D'Estree and Colby, 2004; Armbruster, 2008), drawing on antecedent negotiation theory and experimental studies of cooperation and competition (Raiffa, 1982; Axelrod, 1984). We also have consistent evidence for improvements in relationships among parties in ECR processes (Talbot, 1983; Buckle and Thomas-Buckle, 1986; Susskind, McKernan, and Thomas-Larmer, 1999; Innes and Booher, 1999; O'Leary and Raines, 2001; O'Leary and Bingham, 2003; Dukes 2004b).

We know less from the research about what conditions and factors lead to high ECR performance. Most of what we know is drawn from largely qualitative analysis of individual and small comparative case studies combined with theory (Susskind, McKernan, and Thomas-Larmer, 1999; Kolb, 1997) and collective wisdom from practice (Society of Professionals in Dispute Resolution, 1997). However, empirical evidence from predictive, multicase analysis that directly links precursor conditions and factors with expected ECR outcomes is limited.

These linkages are the focus of this study. What Innes and Booher (2004) emphasized with respect to collaborative processes in general holds true for ECR as well: “Legitimacy of the collaborative methods . . . will remain contested until they have proved themselves more widely” (p. 430). This study and the evaluation system on which it is based are intended to contribute to our understanding of what specific conditions and factors improve ECR performance.

Study Background

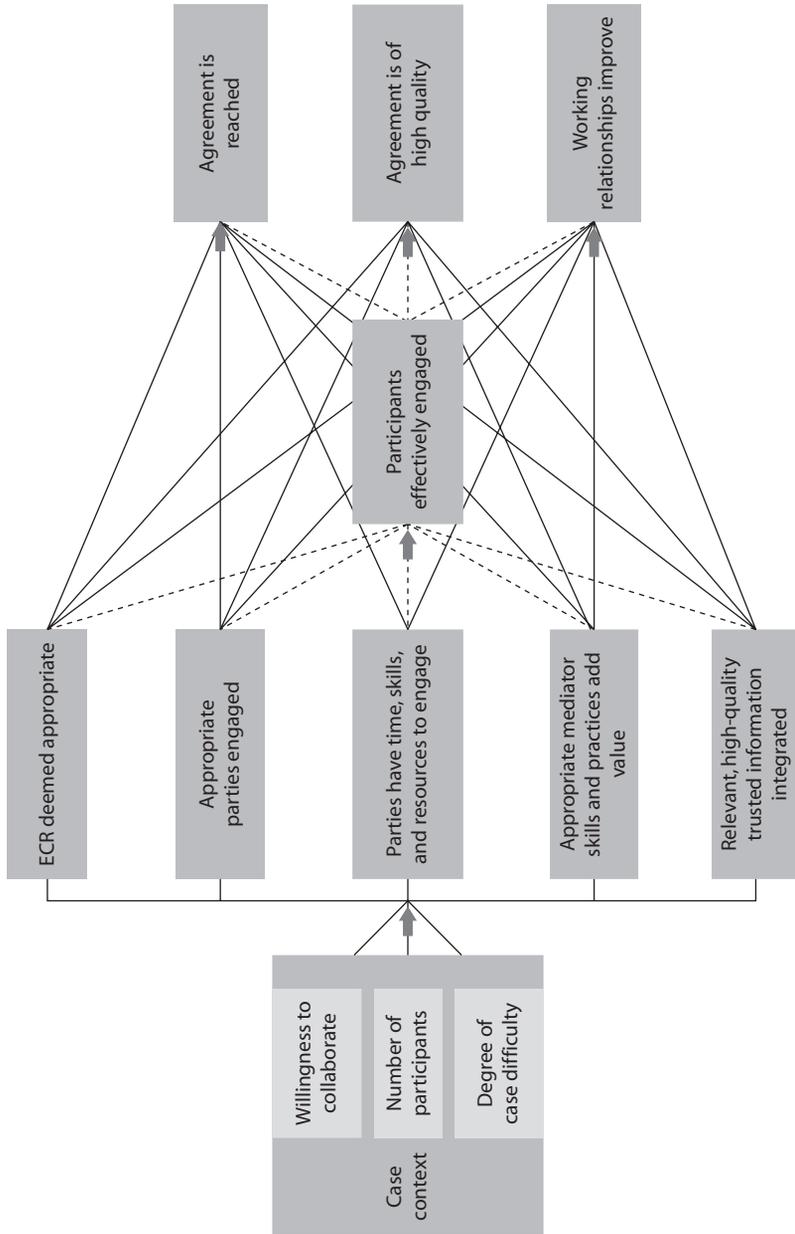
The opportunity to evaluate the workings and outcomes of a large number of ECR processes arose from a partnership of federal and state public policy dispute resolution programs that began in 1999.¹ The impetus for this partnership and its evolution has been described elsewhere (Emerson and Carlson, 2003; Orr, Emerson, and Keyes, 2008). The coordinated effort to collect data through shared procedures and evaluation survey instruments, led by the U.S. Institute for Environmental Conflict Resolution, resulted in assembling fifty-two ECR cases concluded between 2005 and 2007 involving multiple public agencies and stakeholders concerning a broad array of environmental issues around the country. This evaluation effort has become known as the Multiagency ECR Evaluation Study (MAES).

The MAES evaluation framework was itself a product of collaboration among program managers interested in evaluating how their programmatic efforts were contributing to effective use of ECR. This required articulation of a shared working theory about expected ECR outcomes and those conditions and factors over which the managers had some control (Orr, Emerson, and Keyes, 2008). The evaluation framework supplies the conceptual basis for the design of the postprocess survey instruments, specification of variables, and subsequent analysis presented in this article. It is essentially this framework that we are testing in this study.

Figure 1 shows how we conceptualized the relationships among the variables in the framework. Each of the key framework variables can act directly to influence outcomes. In addition, they can work indirectly through one of the variables (participants effectively engaged). Finally, contextual variables can affect the outcomes in nonspecified ways.

The key outcomes of interest are whether an agreement was reached, the quality of the agreement, and the improvement in working relationships among the participants. The key explanatory variables include:

Figure 1. Relationships Among the Variables in the Evaluation Framework



(1) whether ECR was deemed appropriate at the outset; (2) whether the appropriate participants were involved; (3) the participants' capacity to engage; (4) the mediators' skills and practices; (5) the use of relevant, high-quality information; and (6) the extent to which participants were effectively engaged. The diagram indicates that the first five of these factors work directly and indirectly through *effective engagement* to influence the outcome variables. There are also three contextual variables measuring the case challenge controlled for in this analysis: the number of participants in the case, the mediators' rating of case difficulty, and the willingness of the participants to collaborate at the beginning of the process.

Specifying Variables

The construction of each variable represented in Figure 1 and the results of the validity and reliability testing are described in Exhibit 1.

Exhibit 1. Construction of Variables for Exploring Factors That Contribute to Outcomes

Following are descriptions of the variables used in the statistical analysis. Unless otherwise noted, the variables were constructed from items scored on an 11-point scale (for example, where 0 = "not at all" and 10 = "completely"), and the average z-score for the items used to construct each variable was used as a unit-weighted factor score to represent the construct of interest. For these unit-weighted factor scores, Cronbach's alpha was used to assess reliability. This index assesses inter-item consistency, or the extent to which the items correlate with one another. When a factor is assumed to be unidimensional, its items or indicators are expected to be highly correlated. For factors that are assumed to be multidimensional, inter-item consistency is not expected to be high; items may be correlated, but that is not a requirement as would be the case for a unidimensional factor.

Preliminary assessments of convergent validity for the following variables were carried out by correlating the variable of interest with variables with which it is expected to correlate. Because most of the available data are from the study survey, assessments of convergence were the best available for measuring validity. For example, there was no external evidence about the extent to which agreement was reached, a key dependent variable in the study; therefore, we correlated the participants' rating of this variable with the mediator's because they were done independently. The magnitude of the correlation is reported for the validity analyses.

Outcomes (Dependent Variables)

Agreement achieved. A single item reflecting a 5-point scale ranging from 0 ("No agreement, we ended the process without making much progress") to 4 ("Agreement reached

on all key issues”). This outcome was measured using the participants’ responses only. To assess validity, participant responses were correlated with the mediators’ responses to the same item, with the hypothesis that there ought to be a strong positive correlation between the ratings. The resulting correlation was positive and significant ($r = 0.67, p < 0.0001$).

Quality of the agreement. Seven items, each assessing a feature of the agreement on the basis of participant responses only (you understand the terms of the agreement, it takes account of your interests, can be modified if needed, deals with the key issues, will address issues or resolve conflict if implemented, can be implemented, and participants have built relationships to ensure it will last). Cronbach’s alpha for this variable is high ($\alpha = 0.91$). Correlations between this variable and mediator’s ratings of three items related to quality (the agreement is implementable, is adaptable to unanticipated circumstances or changed conditions, and addresses all critical issues were addressed) were unexpectedly low, ranging from 0.19 to 0.26.

Improvement in capacity to work together (working relationships). A difference score created by subtracting each participant’s response to a question about his or her recollection of ability to work together cooperatively on issues “for this case or project” *before the process began* from the person’s response to the same question *as the result of the process*. To assess validity, this difference score was correlated with the change in participants’ trust of one another, before the process compared with after, with the hypothesis that improvement in capacity to work together should be positively correlated with improvement in trust. The resulting correlation was strong, positive, and significant ($r = 0.81, p < 0.0001$).

Explanatory Variables (Independent Variables): Contextual

Case challenge: number of participants. Obtained from the case sponsor or the mediator.

Case challenge: mediators’ rating of case difficulty. One item, rated by the mediator, assessing the difficulty of developing and implementing an effective collaborative process for the case or project compared with other similar ones.

Participant’s willingness to collaborate. One item rated by the participants about their attitude at the start of the process. This item was correlated with an item asking respondents to report the extent to which the case participants were able to work together cooperatively at the start of the process. It was expected that the items would be moderately correlated. Unexpectedly, the correlation was low ($r = 0.16$), perhaps reflecting a difference between the desire for future collaboration and the past behavior of many participants.

Explanatory Variables: ECR Framework

ECR determined to be appropriate. Measured by one surrogate: whether a situation or conflict assessment was conducted either separately from or as part of the process; a binary item (yes or no response) as reported by the mediator.

(Continued)

Exhibit 1. (Continued)

Appropriate participants involved. Four items, rated by the participants, regarding the extent to which all the needed participants were involved and appropriately engaged. One item was binary (yes or no; “Were all participants that were needed part of the process?”), and three items were rated on an 11-point scale measuring the extent to which the absence of some had a negative impact, participants had authority to represent their group, and the groups together reflected the necessary range of interests. The average z-score of the items was used to represent this variable. With respect to reliability, Cronbach’s alpha indicates moderate internal consistency across the four items ($\alpha = 0.70, p < 0.0001$). This is likely due to the items measuring a variety of attributes related to the presence of the appropriate participants.

Participants’ capacity to engage. Three items, rated by the participants, assessing the extent to which they had the skills and time, and they or their organization had the resources to participate effectively. We did not calculate Cronbach’s alpha for this variable because it is not necessary that available time, skills, and resources would correlate. Participants’ ratings of the availability of skills, time, and resources were positively correlated with mediators’ ratings of the same variables.

Mediators’ skills and practices. Six items, rated by the participants, assessing the effectiveness of specific mediator skills and practices (mediator had a work plan and timeline, dealt with me fairly, was able to help us move forward when things got tense, made sure the views of everyone were addressed, made sure no one dominated, and helped manage technical discussions efficiently). Regarding scale reliability, Cronbach’s alpha indicates high internal consistency ($\alpha = .95, p < .0001$). To assess validity, this variable was correlated with the extent to which the participant would recommend the mediator without hesitation to others in a similar situation. A strong, positive correlation was expected and confirmed ($r = 0.90, p < 0.0001$).

Use of relevant information. Four items, rated by participants, assessing the extent to which relevant, high-quality, and trusted information was presented, made accessible, and understood by the participants during the process (we worked to identify information needs, I understood the information, everyone had access to relevant information, the validity of the information was accepted by everyone). Regarding scale reliability, Cronbach’s alpha indicates high internal consistency ($\alpha = 0.80, p < 0.0001$). To assess validity, this variable was correlated with an item assessing the extent to which participants believed the process helped them gain a more complete understanding of the issues. A moderate to strong positive correlation was expected; results indicated a moderate correlation ($r = 0.46, p < 0.0001$). Additionally, the variable was correlated with the mediators’ ratings on a scale composed of four similar items, assessing the extent to which relevant information was used. A strong, positive correlation was expected, and a moderately strong one was found ($r = 0.60, p < 0.0001$).

Participants effectively engaged. Six items, rated by participants, assessing their ability to work together during the process and to clarify issues and perspectives (participants worked together cooperatively, participants sought solutions that met common needs, you gained a better understanding of others’ perspectives, others understood and could state your views, you could focus on the key issues, you gained an understanding of the

issues). Cronbach's alpha results indicate high internal consistency ($\alpha = 0.90$, $p < 0.0001$). To assess validity, the variable was correlated with the contextual variable "degree of case difficulty," mediators' ratings of the relative difficulty in developing and implementing an effective collaborative process (lower scores reflect less difficulty). The correlation was expected to be moderate to strong and negative. Results confirmed a moderate negative correlation ($r = -0.51$, $p < 0.0001$). Note that this variable was also modeled as an intermediate outcome to better understand how the explanatory variables influenced the primary set of outcomes.

Outcomes

Reaching Agreement. Defining ECR as "an agreement-seeking process" requires us to address the performance outcome of whether an agreement was indeed reached. The rate of settlement or agreement as a single measure of ECR success has been a contested standard in the research and practice communities for some time (Bingham, 1986; Buckle and Thomas-Buckle, 1986; Dukes, 2004b). As ECR is applied across a broadening spectrum of applications, the very meaning of agreement must be broadened as well, from final settlements within a judicial context to agreements on policy recommendations, proposed rules, resource management plans, and facility siting proposals. Likewise, ECR is increasingly seen as a component of a larger complex system or life cycle of a policy conflict, not just as a discrete alternative to reaching agreement in some conventional administrative, judicial, or legislative pathway (Dukes, 2004a). This makes comparison across cases served by ECR and not served by ECR especially problematic.

Nonetheless, reaching agreement remains one instrumental outcome of ECR that deserves attention. We asked participants and mediators to rate the extent of their progress toward reaching agreement using five categories (agreement on all key issues, agreement on most key issues, agreement on some key issues, no agreement but progress made, and no agreement and little progress).

Quality of Agreements. The literature on the quality of agreements reached through ECR goes back more than twenty-five years, to the early work of Susskind and Ozawa (1983). Numerous dimensions of high-quality agreements have been delineated since then, chief among them (1) the durability of such agreements according to their maximizing joint gains, taking into account all key issues, and addressing underlying interests; (2) the practicality and "implementability" of such agreements, taking into account available resources, legal precedents, and political acceptability; (3) the flexible nature

of such agreements, in that the agreements take into consideration changing and uncertain conditions in the future; and (4) the accounting for compliance through monitoring and evaluation provisions (see Susskind and Cruikshank, 1987; Hamilton, 1991; Sipe and Stiftel, 1995; Innes and Booher, 1999; Susskind, van der Wansem, and Ciccarelli, 2000; Todd, 2001; Kloppenburg, 2002; D'Estree and Colby, 2004; Dedekorkut, 2004).

In this study, we gathered information on most of these dimensions as rated after the process by the participants and the mediators. For the purpose of specifying a summary score for the modeling analysis, we selected seven substantive descriptors of quality agreements: the extent to which (1) the participants understand the terms of the agreement, (2) the agreement takes account of their interests, (3) it can be modified if needed, (4) it deals with the key issues, (5) it will address issues or resolve conflict if implemented, (6) it can be implemented in its current form, and (7) participants have built relationships to ensure it will last.

Improved Working Relationships. There have been consistent findings in the literature that ECR can lead to improved or transformed relationships among parties (Talbot, 1983; Buckle and Thomas-Buckle, 1986; Susskind, McKearnan, and Thomas-Larmer, 1999; Innes and Booher, 1999; O'Leary and Raines, 2001; O'Leary and Bingham, 2003; Dukes, 2004b). Relationship building is a particularly important outcome for site-specific or place-based issues where parties will be living in proximity to one another in the future, and in many instances they will be working together to implement or monitor the terms of the agreement. Developing or reestablishing mutual trust has been studied as a key component and outgrowth of relationship building in deliberative conflict resolution processes (Innes and Booher, 1999, 2004). Findings from two postprocess studies indicate moderate improvements in relationships (as reported by 42 percent of respondents in Berry and Stiftel's 2001 study) and success in generating trust (56 percent reporting to Frame, Gunton, and Day, 2004).

We derived the relationship outcome indicator from the post-hoc ratings by the participants of how well they worked together cooperatively on case-related issues before the ECR process began, as compared to after the process concluded.

Contributing Factors

Contextual Variables. In our preliminary work on evaluation, we did not adequately distinguish among cases on the basis of the degree of challenge

they posed to the participants and the mediators. Being able to control for certain context variables makes comparative analysis more useful. Case challenge has been measured in various ways, often indirectly by the number of parties engaged or the number and complexity of issues involved (Floyd, Germain, and Terhorst, 1996; Beierle and Cayford, 2002; Leach, 2006).

We developed three indicators of case challenge. One relies on the mediators' judgment of the degree of case difficulty compared to other similar cases. A second is based on the respondents' post-hoc assessment of how willing they thought all parties were to participate at the outset of the process. The third is the number of participants in the case.

Determining If ECR Is Appropriate. Our working theory of ECR includes the critical first step of assessing whether the problem or conflict at hand can benefit from ECR. This requires an informed determination through some preliminary assessment that the issue is ripe and the parties are ready to engage (McCarthy and Shorett, 1984; Moore, 1986; Army, 1987; Carpenter and Kennedy, 1988; Oregon Department of Land Conservation and Development, 1996; SPIDR, 1997; Susskind and Thomas-Larmer, 1999; Susskind, McKernan, and Thomas-Larmer, 1999; Policy Consensus Initiative, 1999; Dedekorkut, 2004). This initial determination, often followed by more extensive third-party investigations, is an important service that agency-based programs provide. Many practitioners view assessment work as not only a prerequisite to any process but as part of the process itself, yielding front-end benefits to the parties by eliciting views and framing issues objectively.

Our indicator for ECR being deemed appropriate was whether a situation or conflict assessment was conducted either separately from or as part of the process, as reported by the mediator.

Engaging the Appropriate Participants. Figuring out who should be at the table representing which interests has been a longstanding threshold question for ECR conveners. This has not been made easier by the inevitable tension between the principle of inclusive representation of a broad range of affected and interested parties and the principle of informed commitment that binds people to membership rules that may become exclusive (Dukes, 2004b; Leach, 2007). Striking the right balance here is a case-specific judgment call that affects not only subsequent process design decisions but ultimately the perceived internal and external legitimacy of the group.

We derived the index for this variable on the basis of the respondents' assessment of parties' representation (the necessary range of interests and authority to represent their group or constituency) and the extent to which it mattered (were all participants who were needed part of the process, and did the absence of any have a negative impact?).

Contribution of Facilitators and Mediators. The value of the third-party neutral in conflict resolution processes has been virtually axiomatic in the literature (Dukes, 2004b). In a more recent review of empirical studies on public participation in the USDA Forest Service, Leach noted that "the presence of an effective facilitator/coordinator is one of the most frequently cited keys to success" (2006, p. 46).

We characterized the contribution of the skills and practices of the mediators with an index that included rankings on their efficiency (mediator had a work plan and timeline; helped manage technical discussions efficiently), fairness (dealt with me fairly; made sure the views of everyone were addressed), and conflict management skills (was able to help us move forward when things got tense; made sure no one dominated).

Use of Information. One of the distinctions frequently drawn between deliberative processes such as ECR and conventional public agency decision making with standard notice and comment is that ECR enables fuller consideration of more sources and kinds of relevant information. Through the greater scrutiny afforded by more diverse interests and both expert and lay person review, the information itself is of higher quality, more useful, and, importantly, more trusted. In turn, this relevant, high-quality, and trusted information contributes to better negotiated agreements. For example, Leach (2006) finds numerous empirical studies reporting that having adequate information (both scientific and technical) is critical to success.

For this study, four dimensions of information were rated: the extent to which relevant, high-quality, and trusted information was presented; how accessible information was to all participants; how well it was understood by all participants; and how valid the information appeared to everyone.

Effective Engagement. This is not a term frequently used in the literature. Many researchers have, however, identified the importance of process dynamics such as opening lines of communication, developing mutual understanding, sustaining active engagement, and learning together (Innes

and Booher, 1999; Susskind, McKearnan, and Thomas-Larmer, 1999; Leach, Pelkey, and Sabatier, 2002; Dedekorkut, 2004; Daniels and Walker, 2001). Some argue that the essential ingredient of any successful process dynamic is collective recognition of mutual interdependence (Delli Priscoli, 1987; Susskind and Cruikshank, 1987; Hamilton, 1991); others recognize a transformational nature in collaborative engagement (Folger and Bush, 1996).

In our preliminary evaluation framework derived from iterative discussions among the MAES study partners, we identified a number of such process factors as dimensions of “effective engagement.” Among them were such indicators as how well participants communicate and collaborate, understand each other’s views and perspectives, and understand the issues at hand. We depicted these factors within a linear model of the independent variables in sequence, suggesting a serial progression of influences on outcomes. As we began to develop models for testing the framework, we wanted to better capture the interactive dimension of effective engagement as an intermediate variable (as depicted in Figure 1) toward which the other key independent variables contribute, but which in itself was an important precursor to successful outcomes.

Effective engagement for purposes of this study became a combination of process indicators rated by the participants, including their ability to work together during the process (participants worked together cooperatively, focused on key issues, and sought solutions that met common needs) and to clarify issues and understand one another’s views.

Data and Methods

The MAES process required extensive coordination with a number of federal and state agency partners and with members of the National Roster of ECR Practitioners.² They helped identify candidate ECR cases, which were then further investigated to determine whether they met these criteria for inclusion in the study:

- The case involved public lands, natural resource, or environmental issues, including related energy, land use, and transportation concerns
- The process was intended to seek agreement among parties (be it in the form of written or unwritten plans, proposals and recommendations, procedures, or settlements)

- The process was facilitated or mediated by a neutral third party, where the neutral could be internal or external as long as she or he was acting in a neutral capacity recognized by the parties
- The case concluded approximately six months prior to our contact and survey dissemination; that is, the stakeholders had reached agreement or decided to stop deliberating within six months of our inquiry

The first three criteria were consistent with the definition of ECR found in the 2005 ECR policy issued by the U.S. Office of Management and Budget and the President's Council on Environmental Quality.³ The study intention was to collect as many ECR cases as possible emphasizing the importance of gathering cases that represented a range of outcomes, not merely success stories. With respect to some of the contributing agencies, their cases represented 100 percent of the known ECR cases conducted during that period.

Survey Instruments

The survey instruments were developed over several years, with the last revision approved in 2004 by the U.S. Office of Management and Budget through the Information Collection Request process (Paperwork Reduction Act, 44 U.S.C. 3501 et seq.). Over this period, more than fifty researchers, practitioners, and evaluators from the ECR field were engaged in critiquing and refining the evaluation instruments. This extensive collaborative development process created the opportunity to design the instruments around a shared articulation of ECR practice and its intended outcomes and impacts. To this end, the instruments are used to gather feedback on the workings of ECR from the initiation of a process where critical process inputs (for example, appropriate participants engaged, suitable mediator selected to guide the process) combine to create desired process activities (such as participants communicating and collaborating) with the end goal of achieving desired outcomes (agreement to resolve an environmental controversy) and impacts (agreement goals are realized). The design and composition of these questionnaires are documented in Orr, Emerson, and Keyes (2008).

Data Collection

The U.S. Institute for Environmental Conflict Resolution managed the collection and merging of case evaluation data from multiple agencies, including a number of cases identified by independent case mediators and

facilitators (called “mediators” hereafter). The Dillman data collection methodology was used for all collections in an effort to achieve as high a questionnaire return rate as possible (Dillman, 2000). The evaluation questionnaires were mailed or electronically delivered to the mediators and to all participants in qualifying cases. Once the Dillman collection cycle for the fifty-two cases was complete, an aggregate dataset was created. The dataset was then cleaned (assessed for consistency between agency contributors and accuracy) and transferred to SPSS (for Windows, release 15.0, 2006) for the analysis phase.

Missing Data

In this study, there were two levels of missing data: person and item. At the person level, data were missing for those who participated in a conflict resolution process but failed to respond to the survey. Of the fifty-two cases in this study, the nonresponse rate ranged from 0 to 90 percent per case, with a mean of 57 percent. At the item level, survey respondents failed, either inadvertently or by choice, to respond to all items. In general, most respondents completed the entire survey.

For the inferential statistics reported in this article, missing data were handled differently for these two levels (items and persons). For scale scores, items were combined using the average z-score of all of the items within that scale. The average is calculated for all the observed items' z-scores, and this average becomes the scale score. Therefore all respondents have a scale score even if they are missing data for one or more items. This is advantageous over other methods of handling missing item-level data because the item weighting is generalizable (each item is weighted equally as opposed to weights based on sample-dependent variance), it is an easy and transparent data reduction procedure, and it is parsimonious (average z-scores correlate highly with differential weighting schemes such as factor analysis; McKnight, McKnight, Sidani, and Figueredo, 2007).

For data missing at the participant level, the EM algorithm for the maximum likelihood procedure used to estimate model parameters for the multilevel models used in this study is known to be efficient and unbiased when missing data are MAR (missing at random). The technical details are beyond the scope of this article.^{4, 5} Suffice it to say that experts on missing data acknowledge that maximum likelihood estimation is considered state of the art in conducting multilevel analyses, as in this study (Schafer and Graham, 2002).

Characteristics of the Dataset

The dataset contains responses from 523 participants and forty-eight mediators in fifty-two ECR cases. Cases reflect a wide geographic distribution, covering more than twenty states (some cases were regional in scope). Figure 2 displays the number of respondents and response rate based on the mediator’s identification of the affiliations of the participants, and the respondent’s self-identification of affiliation. Government representatives are the largest category for both participants and respondents.

Table 1 shows the distribution of the fifty-two cases by purpose and application. ECR purposes range from collaborative, consensus-building processes to focused, dispute resolution efforts. Eight categories of substantive application emerged from the case data. As shown, examples of each purpose and each application are present, though not every combination is represented; “Develop Plans or Site Facilities” is the most prevalent purpose and “Pollution or contamination” is the most prevalent application.

Cases varied in size from two to seventy-six participants, with a median of twenty-four. The mediators were asked to rate their case challenges compared to others in their experience (“difficulty developing and implementing an effective collaborative process”). Half the cases were rated moderately to extremely difficult (with the distribution between the two

Figure 2. Participants and Respondents by Affiliation

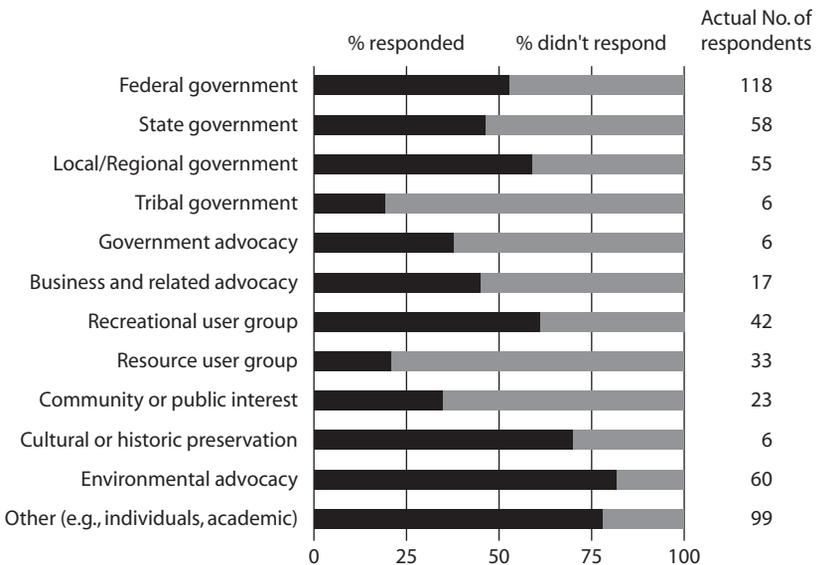


Table 1. Purpose and Application of the ECR Processes

<i>Substantive Area of Application</i>	<i>Purpose</i>						<i>Total</i>
	<i>Develop Guidance/ Policy</i>	<i>Develop Rules</i>	<i>Develop Plans or Site Facilities</i>	<i>Settle Enforcement Actions or Specific Disputes</i>	<i>Other</i>		
Transportation facilities	2		3				5
Off-road travel			3				3
Resource supply or use (energy, water, timber, fish)		3	1	3			7
Habitat or species protection or ecosystem restoration		1	5				6
Multiple use of resources		1	7				8
Pollution or contamination	2	1	1	8			12
Fire protection			3			1	4
Other			4	2	1		7
Total	4	6	27	13	2		52

slightly favoring the extreme end), and half were rated at a minor level of difficulty or not difficult at all.

Both the potential and the limitations of the dataset need to be acknowledged. The dataset itself represents one of the most comprehensive collections of evaluation data on ECR processes available to date. The vast majority of respondents appeared to have completed the questionnaires thoughtfully, using the entire range of the response scales. The numbers of cases and respondents are large enough to allow meaningful statistical analysis. Nevertheless, no claim is made that the dataset represents the universe of ECR cases or participants, despite our efforts to include a range of cases reflecting different applications, geographic settings, and degrees of challenge and success. Moreover, many participants chose not to respond, and many who did respond failed to answer all the questions. Nonrespondents and missing data may introduce unknown bias. Finally, the data are in large part perceptions of the participants and mediators, not independent measures of case characteristics and outcomes.

Descriptive Statistics Measuring Performance

For the majority of questions, respondents were asked to rate ECR performance on the basis of a 0 to 10 agreement scale, where 0 was labeled “not at all,” the midpoint 5 was labeled “moderately agree,” and 10 was labeled “completely agree.” Descriptive statistics (including the mean, standard deviation, median, and percentage frequencies) were used to summarize participant responses to evaluation statements. Respondent ratings of the evaluation questions were averaged across the entire dataset to generate participant-level outcomes. Respondent ratings were also averaged within each case to generate case-level outcomes. Case-level outcomes are reported as the percentage of cases where the average respondent score was above the *midpoint* in the scale (5.00). For example, if 80 percent of cases are reported to have achieved some attribute, this means that in 80 percent of the cases the average respondent score for that attribute was above 5.00. Case-level data are reported when process or case-level outcomes are of interest, and participant-level data are reported when respondent feedback is the measure of interest.

Multilevel Modeling to Analyze Contributing Conditions and Factors

In addition to learning about the performance of ECR in the fifty-two-case dataset, we wanted to test our working theory of ECR as captured in the ECR evaluation framework (see Orr, Emerson, and Keyes, 2008). By probing relationships among factors in the framework, we could better understand those that exert the most influence on ECR outcomes. Variables were constructed to represent these factors using answers to the questions in either the mediators’ or participants’ questionnaires. Validity tests were conducted by correlating the target variable with other relevant variables, an index of convergent validity. For unidimensional factors (variables created by combining items thought to measure the same underlying construct), Cronbach’s alpha was used to assess internal consistency, a measure of reliability. For many of the variables, the average z-score for the items on the 0 to 10 rating scales was used to generate a unit-weighted factor score. This method of constructing variables has been demonstrated to have desirable characteristics compared to factor analytic procedures: less variance and greater generalizability (values are less sample-specific). Moreover, using the average z-score is generally a better approach than summing the raw item values for a total score in the presence of missing data at the item level (see McKnight, McKnight, Sidani, and Figueredo, 2007).

Multilevel modeling (MLM) was chosen as the appropriate analytical method to investigate relationships among the variables on the basis of two key features of the dataset: (1) participants were not randomly distributed among the cases, as in an experimental design where participants are randomly assigned to conditions; and (2) characteristics of the cases as well as of the participants could contribute to the outcomes of interest. The first feature violates a key assumption of general linear modeling: that all the observations are independent. Thus, the familiar ANOVA and simple regression methods were ruled out as inappropriate for the analyses. Confirming the need for MLM, the correlation of respondents within cases (the intraclass correlation, or ICC) for the outcome variables was high, ranging from 0.29 to 0.61, confirming the nonrandom distribution of participants across cases. The second feature can be probed only by using methods that can mathematically tease out the differential effects at the case and participant levels. Given these data features, MLM is the method of choice (Arnold, 1992; Kreft and De Leeuw, 1998; Singer, 1998; Sullivan, Dukes, and Losina, 1999). MLM is increasingly used in the social sciences as researchers become aware of the advantages of these models in analyzing multilevel or “nested” data (O’Connell and McCoach, 2004; Ringdal, 1992; Singer, 1998). We believe this is the first application of MLM to ECR case analysis.

Using MLM, we then analyzed the extent to which the contextual conditions and key factors in our ECR framework contribute to these ECR outcomes. Note that our goal is not to explain as much variance as possible in outcomes across the cases and participants in our dataset. Rather, our focus is on understanding the influence of those factors that we believe could affect performance and over which program managers and ECR practitioners exercise some control.

As with any statistical method, care must be exercised in applying MLM and interpreting the results. The models are statistically complex and require expertise in specification and interpretation. The order in which explanatory variables are entered into a model is important and should be based on theory—in our case, the ECR evaluation framework. Collinearity among the variables is particularly problematic, decreasing the stability of model coefficients and making standard statistical tests of significance unreliable. To build and interpret the models, we used two conventional measures of how well the models explained the observations: model fit and percentage variance in the dependent variables explained by the independent variables (Kreft and De Leeuw, 1998; Singer, 1998).

Finally, model coefficients can be treated as fixed or random with respect to each variable. Whether a variable ought to be tested as a random or fixed effect is determined substantively as well as empirically. For variables that we assumed to be random effects, we tested the model with and without the random effect; if model fit was not improved with the random effect, the variable was tested only as a fixed-effect variable.

Multilevel modeling uses maximum likelihood for parameter estimation. In the presence of missing data, maximum likelihood is beneficial for technical reasons that go beyond the scope of this article (see Schafer and Graham, 2002; and McKnight, McKnight, Sidani, and Figueredo, 2007, for a discussion of the benefits of maximum likelihood for handling missing data).

In addition to using z-scores to construct the variables, as noted in Exhibit 1, all the variables in each of our models were specified using grand-mean-centered data. That is, each observation was expressed as the deviation from the grand mean. Centering is desirable because it removes high correlations between case-level and participant-level variables and increases model stability, among other technical effects (Kreft and De Leeuw, 1998; Singer, 1998).

Each model was designed to represent relationships between the three principal outcomes and the explanatory variables. The models were specified as linear relationships and the coefficients were estimated using the MLM algorithms in SPSS. Explanatory variables were entered sequentially with the order being based on relationships depicted in the ECR evaluation framework. Methods were used to hierarchically partition variance in order to account for multicollinearity among explanatory variables. Variables were retained in the models if model fit improved with their inclusion.

Findings

ECR Performance

In 82 percent of the cases, full agreement (agreement on all or most key issues) or partial agreement (agreement on some key issues) was reached, according to the participants. Progress was made without agreement in another 10 percent of cases, while in 8 percent of the cases no agreement was reached and little progress was made. Interestingly, the participants within a given case did not always concur on whether agreement was reached and the extent to which it was reached. As a consequence, participant median reported outcomes are used to summarize case outcomes

(Table 2). The mediator-reported outcomes indicate that in 85 percent of the cases full or partial agreement was reached.

With respect to agreement quality (Table 3), most of the respondents agreed that they understood the terms of the agreements, their interests were taken into account, and ultimately the agreement dealt with the key

Table 2. Levels of Agreement Based on Median Participant Responses

	Percentage of Cases	
	Participant Median Respondent Reported Outcomes	Mediator/Facilitator Reported Outcomes
Full agreement (agreement on <i>all</i> or <i>most</i> key issues)	61	77
Partial agreement (agreement on <i>some</i> key issues)	21	8
Progress but no agreement	10	9
No agreement and little progress	8	6

Table 3. Participant Perspectives on the Quality of Agreements Reached

	<i>n</i>	Percentage of Ratings ≥ 5	Median Ratings	Mean (Std. Dev.)
Participants understood the agreement terms	382	96	9.00	8.31 (1.90)
The interests of the participants were taken into account	381	86	7.00	6.83 (2.40)
Agreement effectively dealt with key issues	380	81	7.00	6.54 (2.52)
If implemented, the agreement will effectively address the issues or resolve the controversy	380	78	7.00	6.21 (2.58)
Participants are confident the agreement can be carried out in its current form	379	75	6.00	6.07 (2.75)
Agreement specifies how it can be changed or modified if things don't go as planned	372	69	6.00	5.67 (3.02)
Participants are confident they have built strong enough relationships to ensure the agreement lasts	376	75	6.00	5.96 (2.72)

Table 4. Mediator Perspectives on the Quality of Agreements Reached

	<i>n</i>	<i>Percentage of Ratings ≥ 5</i>	<i>Median Ratings</i>	<i>Mean (Std. Dev.)</i>
Agreement addresses all critical issues	37	95	9.00	8.15 (2.24)
Addressed more of the parties' interests than would otherwise without the process	36	100	10.00	9.62 (0.85)
Took full advantage of relevant information	36	100	10.00	9.14 (1.36)
Legal requirements were addressed	32	100	9.50	8.82 (1.48)
Contained clear and measurable standards and objectives	31	84	8.00	7.34 (2.36)
Identified roles and responsibilities for implementation	35	97	9.00	8.01 (2.33)
Includes a plan for monitoring implementation	27	78	8.00	6.96 (3.07)
Provided means for adapting to unanticipated events	33	91	8.00	7.80 (1.92)
Included conditions under which parties would reconvene if needed	24	58	7.50	5.81 (3.84)
Addressed the resources needed for implementation	31	74	7.00	6.39 (2.92)
Agreement is implementable	37	100	9.50	8.91 (1.44)
Specified how the participants would know when the agreement was fully implemented	29	79	7.00	6.84 (2.81)

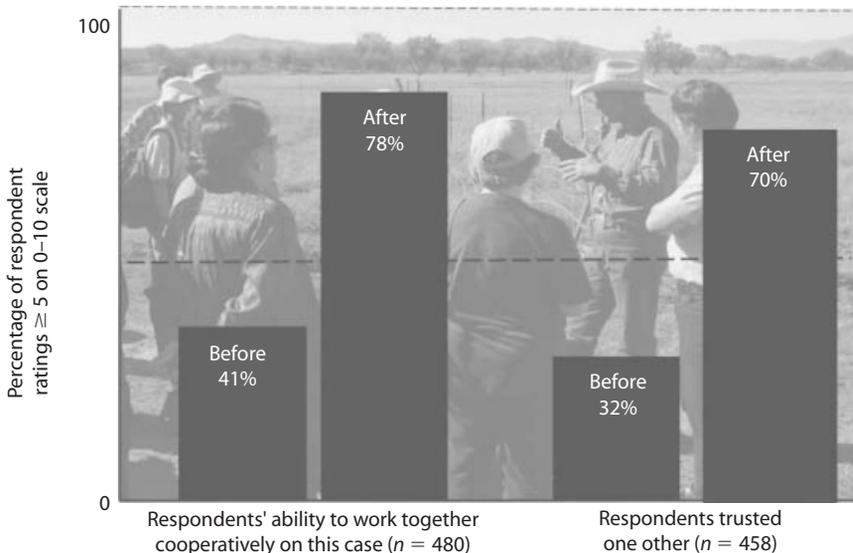
issues. The median ratings ranged from 7 to 9 (on a scale of 0 to 10, with higher values reflecting greater agreement with statements). Regarding agreement durability and “implementability,” the majority of respondents indicated they were confident their agreements could be carried out in the current form, the agreements specified how they can be changed or modified if things don’t go as planned, and if implemented the agreements will effectively address the issues or resolve the controversy. Three-quarters of the respondents also reported that they had built strong enough relationships to ensure the agreements would last. That said, when one looks at the median

ratings it appears that the parties have less confidence in the durability of the agreements than in other agreement characteristics.

The mediators provided a second and more detailed perspective on the quality of the agreements (Table 4). The mediator feedback addressed the many dimensions of agreement quality, notably (1) the durability of agreements based on their maximizing joint gains, taking into account all key issues, and addressing underlying interests; (2) the practicality and “implementability” of such agreements, taking into account available resources and legal precedents; (3) the flexible nature of such agreements, acknowledging changing and uncertain conditions in the future; and (4) accounting for compliance through monitoring and evaluation provisions. From the mediators’ perspectives, the agreements reached were of high quality, although agreement attributes that deal with future contingencies are not as well addressed.

In terms of working relationships, the third outcome variable measured, the majority of respondents reported an increase in their ability to work together on issues related to their case, and an increase in the level of trust among stakeholders as a result of ECR (Figure 3). Improvement in the working relationships is based on a difference score created by subtracting each participant’s response to a question about recollection of his or her ability to work cooperatively on issues “for this case or project” *before the process began*

Figure 3. ECR Improved Stakeholder Working Relationships



from the person's response to the same question *as the result of the process*. To assess validity, this difference score was correlated with the change in participants' trust of one another from before the process to after the process, with the hypothesis that improvement in capacity to work together should be positively correlated with improvement in trust. The resulting correlation was strong, positive, and significant ($r = 0.81, p < 0.0001$).

In addition, when asked to summarize what was accomplished as a result of ECR (from a list of nine outcomes ranging from "a crisis was averted" to "the process made the issue or dispute worse"), the respondents most frequently cited as an accomplishment that "relationships among parties in this process were improved."

Factors That Influence Outcomes

The results lend strong support to the structure of the ECR framework. Table 4 shows the MLM results for one intermediate and the three primary outcomes. The order of the explanatory variables reflects how they were entered into the models. A dashed line means the variable did not improve model fit and was not retained in the final model. The signs and the strengths of the variables are based on the final models after all variables were entered. Note that all of the explanatory variables are based on respondent-level data except for the number of participants in the case and the mediators' assessment of case difficulty.

We posited that the ECR variables work directly and indirectly (through contributing to effective engagement of parties) to influence outcomes. In this sense, *participants effectively engaged* can be considered an intermediate outcome. A model for predicting effective engagement is presented first in the shaded column in Table 5 and shows that all but two of the ECR framework variables contribute to whether the participants were effectively engaged. One context variable (case challenge: degree of difficulty), as expected, has a significant inverse or negative relationship to the parties being effectively engaged. In other words, the more difficult the case, as assessed by the mediators, the less likely effective engagement is achieved among the parties. Holding case difficulty constant, we find three ECR framework variables contributing significantly to effective engagement: parties have the capacity to engage; mediator skills and practices add value; and relevant, high-quality, and trusted information used. The more highly parties rated their capacity, mediator skill and practices, and relevant information, the more highly parties rated the effectiveness of their engagement.

Table 5. Modeling Results

Explanatory Variables	Intermediate Outcome		Primary Outcomes	
	Participants Effectively Engaged	Agreement Is Reached	Agreement Is of High Quality	Working Relationships Improve
Context				
Participants' willingness to collaborate at the start	+0.02 (0.02)	-0.04 (0.02)	+0.01 (0.01)	-0.16* (0.05)
Case challenge: number of participants in the case	—	—	—	—
Case challenge: degree of difficulty	-0.05* (0.02)	-0.14* (0.05)	-0.09* (0.03)	-0.02 (0.09)
ECR Framework				
ECR deemed appropriate	—	—	—	—
Appropriate parties engaged	+0.03 (0.05)	+0.24* (0.06)	+0.08* (0.04)	+0.24 (0.16)
Parties have capacity to engage	+0.03* (0.01)	-0.01 (0.02)	+0.01 (0.01)	-0.08 (0.05)
Mediator skills and practices add value	+0.23* (0.04)	+0.01 (0.05)	+0.06† (0.03)	+0.56* (0.15)
Relevant, high-quality, and trusted information used	+0.78* (0.09)	+0.17 (0.14)	+0.19* (0.08)	+0.50 (0.45)
Participants effectively engaged		+0.31* (0.08)	+0.14* (0.05)	+0.74* (0.24)
Model percentage between-case variance explained	0.3%	5.3%	1.3%	7.0%
Model percentage within-case variance explained ^a	45.6%	27.0%	33.2%	25.4%

Note: Coefficients (standard errors) are shown; statistical significance at the $p < 0.05$ is noted by an asterisk; trend toward significance ($p = 0.07$) is denoted by “+.” Solid lines indicate that the explanatory variable did not improve model fit and therefore was not included in the final model. More within-case than between-case variance was explained because most of the variables are measured at the participant level rather than the case level.

^aThe variance explained by the four models from which these findings are derived ranges between roughly 25 and 45 percent (within-case variance). Although as we mentioned the models were not designed to explain as much of the variance in the outcomes as possible, these results are comparable to noteworthy findings from studies in related fields such as education and social psychology.

With respect to the other two variables noted in the “effectively engaged” model (participants’ willingness to collaborate at the start; appropriate parties engaged), they were not found to be significant (although they contributed to model performance, their coefficients are not significantly different from 0, and thus their signs do not have any reliable meaning). That “appropriate parties engaged” did not have a significant coefficient may indicate that many of the respondents focused on how well the actual participants were engaged irrespective of whether or not all the appropriate parties were represented. The finding that neither “case challenge: number of participants in the case” nor “ECR deemed appropriate” contributes to this intermediate outcome model or to the three primary outcome models is addressed below.

In the next three models presented in Table 5, we explore primary ECR outcomes and introduce the intermediate outcome (effectively engaged) as an explanatory variable entered last. Turning first to the “agreement reached” outcome, two of the three context variables and five of the six ECR framework variables are included in the final model. Once again, “case challenge: degree of difficulty” is statistically significant and in the expected negative direction; the more difficult the case, the less likely an agreement is reached. Two of the ECR framework variables are also statistically significant in the expected direction: the higher the rating for “appropriate parties are engaged” and for “participants are effectively engaged,” the more likely an agreement was reached.

Note that the other three ECR framework variables in the “agreement reached” model (parties have the capacity to engage; mediator skills and practices add value; and relevant, high-quality, and trusted information used), although not statistically significant on their own, improved model fit and therefore contributed to predicting the “agreement reached” outcome. Finally, the fact that “participants effectively engaged” is included in the model indicates that this variable exerts an influence above and beyond the influences of the other ECR framework variables.

The model results for “agreement is of high quality” are very similar to those for “agreement is reached.” The notable difference is the heightened importance of “relevant, high-quality, and trusted information used,” which exerts a statistically significant positive influence over the “agreement is of high quality” outcome. In addition, “mediator skills and practices add value” shows a trend toward statistical significance ($p = 0.07$) in a positive direction.

A somewhat different pattern arises from the results for the third primary outcome, “working relationships improve.” First, the influence of the context variables is different: “participants’ willingness to collaborate at the start” is statistically significant and negative (that is, the lower the starting point, the greater the likelihood of improvement), while “case challenge: degree of difficulty” still contributes to the outcome (improves model fit), but not to a statistically significant degree. Second, with respect to the ECR framework variables “mediator skills and practices add value” and “participants effectively engaged,” both are significant, positive contributors to improved relationships.

Looking across all the primary outcomes, two variables were not included in any of the final models: “case challenge: number of participants in the case” and “ECR deemed appropriate.” The number of participants is often assumed to be a surrogate for case complexity, but it did not exert an influence on any outcomes in our study. The fact that “ECR deemed appropriate” did not appear in any of the final models is also surprising. It was measured using a surrogate indicator: whether or not a situation assessment was conducted either separately from or as part of the process. There are several potential explanations: judging the appropriateness of ECR is not important to any of the outcomes, *or* the assessment variable is misspecified. The fact that all these cases did move forward after somebody’s decision that it was appropriate to do so points to misspecification; that is, this surrogate is not measuring the underlying variable. The binary specification of the variable (no assessment and some type of assessment) is another possible explanation for the nonsignificant finding; the variance for this variable was less than optimal (19 percent indicated no assessment, 81 percent indicated some type of assessment). We do not think it is appropriate to conclude that the conduct of an assessment does not matter.

Taking the results for all three primary outcomes together, we conclude that the evidence strongly supports key elements of the ECR framework. With the possible exception of determining that ECR is appropriate before embarking on ECR, all of the ECR framework factors contribute to explaining the outcomes specified through the MLM models.

In addition, our expectations for the importance of effective engagement are also borne out across all three primary outcomes. Effective engagement is a significant contributor to the parties’ perceptions of each positive outcome reported.

Discussion and Conclusion

As noted at the outset of this article, public managers (as well as those who sponsor, fund, and convene ECR processes) are interested in knowing what general results they can reasonably expect from ECR, and what conditions and factors they need to address in order to optimize those results. What have we learned from this study of fifty-two recent ECR cases that can inform those two questions?

Expected Outcomes

To the first question, we can affirmatively respond that participants in ECR processes appear to be able to reach agreements of high quality, even in complex and challenging situations, and they improve their working relationships as a consequence of ECR processes. Despite the complexity of diverse contexts and the breadth of substance and procedures for which agreements are being sought in this fifty-two-case dataset, the extent to which parties perceive agreement was reached is relatively high (82 percent) and generally in line with other empirical studies that address ECR agreement rates (Table 6). Although we cannot fully generalize to the universe of cases from this sample, given that the cases were not randomly selected, the database does represent a sizable number and range of cases, including the level of difficulty as reported by the mediator, the number of parties, case duration, and both purpose and application of ECR. For some agencies, these represent 100 percent of their cases completed within the study timeframe.

Mediators confirmed the participants' judgments, further strengthening the overall finding here. However, mediators appear to be more positive in their view that agreement is reached on all or most issues than the parties themselves. This might suggest that mediators probe the parties about the agreement more thoroughly to confirm that their interests and concerns are being addressed adequately in the agreement being forged. Differing perspectives between mediators and participants, and among participants in the same case, are issues that merit future research.

An important MAES finding is the degree to which parties had differing perspectives regarding the extent to which agreement was reached. This finding echoes some earlier research (Buckle and Thomas-Buckle, 1986) and deserves further attention in the future. It stands to reason that parties with diverse interests will appraise the negotiated agreement through their particular frame of reference, especially with respect to whether they have

Table 6. Comparison of Agreement Findings in the Literature

ECR Studies of Agreement Rate		Agreement Rate
Emerson, Orr, Keyes, and McKnight (2009)	Environmental and natural resource conflicts	82% of cases reached full or partial agreement
Frame, Gunton, and Day (2004)	Land and resource management planning	85% of responding mediators reported full or partial agreement reached
Berry, Stiffler, and Dedekorkut (2003)	Multiparty state agency administrative mediation	93% of completed cases reached agreement
O'Leary and Husar (2002)	Environmental and natural resource conflicts (primarily waste cleanup)	89% of respondents (mediators and participants) reported their cases as having settled at mediation or after mediation
Andrew (2001)	Waste management	61% of responding attorneys reported that use of ADR successfully resolved the dispute
Susskind, McKearnan, and Thomas-Larmer (1999)	Land use	81% of cases resulted in a final settlement
Sipe (1998)	Environmental enforcement cases	61% of responding participants stated that their cases settled
Sipe and Stiffler (1995)	Environmental enforcement cases	85% of cases resulted in a settlement
Buckle and Thomas-Buckle (1986)	Environmental cases where mediation was extensively pursued	74% of cases reached agreement
Bingham (1986)	Agreement-seeking policy and site-specific environmental cases	75% of cases had signed agreement
		78% of cases resulted in an agreement

Source: Table adapted and updated from Andrew (2001).

met their hoped-for outcomes in the case. Consensus endorsements among contending interests are rarely unanimously enthusiastic. Many include those who have accepted the agreement or provision on the table because it is the best they can get at that time or because they can at least live with the agreement for now.

Looking more closely at three case-specific examples (Table 7), we see how this variation across participant perspectives plays out. In all three case examples, the majority of respondents felt that full or partial agreement was reached. However, in all cases there was at least one participant who felt “progress was made but no agreement” or saw “no agreement and little progress.” The mediator’s feedback provides perspective on the differences. For case example one, a few stakeholders stepped aside to let the agreement go ahead. These stakeholders didn’t agree with the resolution, but they felt progress was made because the agreement signed by the other stakeholders was in their opinion better than what would likely have happened in the absence of such an agreement. In the second case example, the mediator noted that one stakeholder withdrew from the process and did not endorse the agreement. In the third example, the case product was a stakeholder report that noted areas of agreement and disagreement among the parties.

Table 7. Case Examples Illustrating Differing Perspectives Regarding Extent of Agreement

<i>Participant Feedback</i>	<i>Differing Perspectives on Agreement: Case Examples</i>		
	<i>Case 1</i>	<i>Case 2</i>	<i>Case 3</i>
Full agreement, on <i>all</i> or <i>most</i> key issues	20% (<i>n</i> = 2)	84% (<i>n</i> = 16)	15% (<i>n</i> = 2)
Partial agreement, on <i>some</i> key issues	60% (<i>n</i> = 6)	11% (<i>n</i> = 2)	54% (<i>n</i> = 7)
Progress but no agreement	20% (<i>n</i> = 2)	0% (<i>n</i> = 0)	23% (<i>n</i> = 3)
No agreement and little progress	0% (<i>n</i> = 0)	5% (<i>n</i> = 1)	8% (<i>n</i> = 1)
Mediator feedback	Agreement on <i>most</i> key issues	Agreement on <i>all</i> key issues	Agreement on <i>all</i> key issues

Note: Additional insights from mediator: in case 1, some stakeholders stepped aside to let the agreement go ahead; in case 2, one stakeholder did not endorse the agreement; in case 3, there was agreement with report, which noted areas of agreement and disagreement.

With respect to agreement quality, public managers can be reassured by this study's findings that, according to the opinions of most of the case participants, the agreements reached through ECR processes took their interests into account and effectively dealt with key issues. These findings are consistent with the respondent ratings in two other studies on environmental enforcement and land use cases (Sipe and Stiffler, 1995; Susskind, van der Wansem, and Ciccarelli, 2000).

Mediators also rated the quality of these agreements very high with respect to how they addressed relevant information and legal requirements, included measurable standards and objectives, clarified roles and responsibility for implementation, and provided means for monitoring. The mediators unanimously reported that the agreements were implementable, although the parties were somewhat less sanguine and more divided over whether the agreements would be carried out and less certain that they would last or how they could be modified if need be. Generally, descriptors for conditions around agreement implementation (for example, a plan for monitoring implementation is included in the agreement, addresses implementation resource needs) and dealing with future contingencies (agreement includes conditions under which the participants will reconvene) received positive but relatively lower ratings than the other agreement descriptors. This might suggest more explicit attention be paid to these issues within ECR processes in the future.

It is important to acknowledge that these ratings are subjective; independent assessments of the agreements themselves would be a more accurate measure of quality. We have collected the written agreements for many of the fifty-two cases, anticipating that future content analyses of the agreements will furnish comparison to the ratings of the participants and the mediators.

On the third key ECR outcome, our study findings indicate substantial reported improvement in the level of trust and participants' ability to work together on issues in the future. This finding will be of particular interest to public managers doing ongoing work with other agencies or stakeholders with an interest in building or restoring constructive partnerships for co-management or shared public investments. Where relationship and trust building have long-term value, ECR can help create the foundation.

These three demonstrably positive outcomes should be informative to public managers as they make decisions about whether to pursue an ECR process. The findings raise two important qualifying questions, however, that go beyond the scope of this article (but that are intended for future study).

First, despite the generally strong performance across cases and participants, are all stakeholders as positively impressed with the ECR process, or are there patterns among the parties that suggest differential perceived benefits? A cursory analysis suggests that positive ratings generally occur across all categories of stakeholders, but this certainly merits further investigation.

A second question pertains to implementation, and the extent to which these agreement and relationship outcomes translate over time into the expected positive on-the-ground impacts, environmental as well as social and economic. This too awaits further long-term follow-up study. What we have learned from the MLM analysis is that not all ECR processes are equal; depending on the context and key contributing factors, the process outcomes and most likely the long-term impacts as well will vary.

Key Factors Influencing Performance

The second question for public managers concerns what needs to be in place to ensure positive outcomes. This study basically confirms most of the elements of the ECR evaluation framework, suggesting that attention needs to be paid to all of them in order to optimize ECR performance. Effectively engaging the parties certainly appears to be a gateway factor and a key predictor of agreements reached, their quality, and improved working relationships.

Effective engagement is a complex dynamic, and our findings suggest that having the appropriate participants; the right mediator skills and practices; and relevant, high-quality information helps ensure that active engagement takes place and is productive.

These study findings should not come as a surprise. However, this empirical analysis confirms best-practice theory through a rigorous methodology that has not been applied before in this field. These results demonstrate why it is so important for public managers and those who convene and sponsor ECR processes to attend to getting the right parties to the table, making sure there is access to the pertinent information, and working with a third-party mediator or facilitator with the requisite skills.

Without addressing these elements, at the very least managers are diminishing the prospects of positive gains from ECR. Admittedly, as the findings reveal, the more challenging the case and the less willing parties are to engage at the outset, the tougher it is to succeed. But what the modeling reveals is that, with case difficulty held constant, the other factors continue to make significant contributions to positive case outcomes. Progress can be made in even the most challenging cases.

Disclaimer

The perspectives expressed in this article are those of the authors alone. The perspectives expressed do not necessarily reflect the views of the data contributors or the institutional perspectives of the U.S. Institute for Environmental Conflict Resolution of the Morris K. Udall Foundation, or Public Interest Research Service, Inc. Any errors or omissions are solely the responsibility of the authors.

Notes

1. The evaluation framework and instruments described in this document have been created collaboratively in several stages over the past few years by specialists from a number of agencies. In addition to U.S. Institute staff, particular thanks is due to key design contributors: Mike Niemeyer from the Oregon Department of Justice, Will Hall from the Environmental Protection Agency (Conflict Prevention and Resolution Center), Elena Gonzalez and Kathy Lynne from the Department of the Interior (Office of Collaborative Action and Dispute Resolution), Chris Pederson from the Florida Conflict Resolution Consortium, Kasha Helget from the Federal Energy and Regulatory Commission, Susan Brody from the National Policy Consensus Center, and Chris Carlson from the Policy Consensus Initiative. Evaluation consultants Kathy McKnight and Lee Sechrest from Public Interest Research Service have guided this effort since 2004, Tom Miller of the National Research Center helped with preliminary data analysis in 2003, and Andy Rowe of ARCEconomics guided development of the original model and provided input on the later revisions. The U.S. Institute would also like to acknowledge the many researchers and practitioners, particularly Bernie Mayer of CDR Associates and Julie Macfarlane of the University of Windsor, for their contributions along the way. Finally, our sincere thanks to the many case participants and facilitators who completed evaluation questionnaires and made this study possible.
2. The National Roster of Environmental Dispute Resolution and Consensus Building Professionals is managed by the U.S. Institute for Environmental Conflict Resolution and includes more than 270 qualified members. See www.ecr.gov for more information on the roster.
3. On November 28, 2005, Joshua Bolten, then director of the Office of Management and Budget (OMB), and James Connaughton, chairman of the President's Council on Environmental Quality (CEQ), issued a policy memorandum on environmental conflict resolution (ECR). This joint policy statement (found at <http://ecr.gov/Resources/FederalECRPolicy/MemorandumECR.aspx>) directs agencies to increase effective use and their institutional capacity for ECR and collaborative problem solving (U.S. Office of Management and Budget and President's Council on Environmental Quality, Nov. 5, 2005).

4. The interested reader should consult McKnight, McKnight, Sidani, and Figueredo (2007) for a nontechnical discussion of the maximum likelihood procedure and the EM algorithm for handling missing data.
5. Data on cost effectiveness and relative merits have been gathered from the respondents and will be reported elsewhere. Long-term tracking of cases by third-party researchers will be needed to more accurately get at these questions.

References

- Andrew, J. S. "Examining the Claims of Environmental ADR, Evidence from Waste Management Conflicts in Ontario and Massachusetts." *Journal of Planning Education and Research*, 2001, 21, 166–183.
- Armbruster, A. *Collaborative Versus Technocratic Policymaking: California's Statewide Water Plan*. Sacramento: Center for Collaborative Policy, California State University, 2008.
- Army, D. *The Politics of Environmental Mediation*. New York: Columbia University Press, 1987.
- Arnold, C. "Methods, Plainly Speaking: An Introduction to Hierarchical Linear Modeling." *Measurement and Education in Counseling and Development*, 1992, 25, 58–90.
- Axelrod, R. *The Evolution of Cooperation*. New York: Basic Books, 1984.
- Beierle, T. C., and Cayford, J. *Democracy in Practice: Public Participation in Environmental Decisions*. Washington D.C.: Resources for the Future Press, 2002.
- Berry, F., and Stiffler, B. "Mediation in State Agency Administrative Procedures: An Assessment of Its Use in Florida State Agencies." Paper presented at Syracuse University's Conference on Evaluating Environmental and Public Policy Dispute Resolution Program and Policies, Washington, D.C., March 2001.
- Berry, F., Stiffler, B., and Dedekorkut, A. "State Agency Administrative Mediation: A Florida Trial." In R. O'Leary and L. B. Bingham (eds.), *The Promise of Environmental Conflict Resolution*. Washington, D.C.: Resources for the Future Press, 2003.
- Bingham, G. *Resolving Environmental Disputes: A Decade of Experience*. Washington, D.C.: Conservation Foundation, 1986.
- Buckle, L. G., and Thomas-Buckle, S. R. "Placing Environmental Mediation in Context: Lessons from 'Failed' Mediations." *Environmental Impact Assessment Review*, 1986, 6(1), 55–60.
- Carpenter, S. L., and Kennedy, W.J.D. *Managing Public Disputes*. San Francisco: Jossey-Bass, 1988.
- Coglianesi, C. "Is Satisfaction Success? Evaluating Public Participation in Regulatory Policymaking." John F. Kennedy School of Government Harvard University Faculty Research Working Papers Series. 2002, RWPO2–238.
- Daniels, S. E., and Walker, G. B. *Walking Through Environmental Conflict: The Learning Approach*. Westport, Conn.: Greenwood, 2001.

- Dedekorkut, A. "Determinants of Success in Interorganizational Collaboration for Natural Resource Management—Florida Planning and Development Laboratory Interagency Collaboration Survey, 2004." Retrieved July 27, 2009, from <http://etd.lib.fsu.edu/theses/available/etd-01062004-231112/>.
- Delli Priscoli, J. "Conflict Resolution for Water Resource Projects: Using Facilitation and Mediation to Write Section 404 General Permits." *Environmental Impact Assessment Review*, 1987, 12(7), 313–326.
- D'Estree, T. P., and Colby, B. G. *Braving the Currents: Evaluating Environmental Conflict Resolution in the River Basins of the American West*. Boston: Kluwer Academic, 2004.
- Dillman, D. A. *Mail and Internet Surveys: The Tailored Design Method*, New York: Wiley, 2000.
- Dukes, E. F. "From Enemies, to Higher Ground, to Allies: The Unlikely Partnership Between the Tobacco Farm and Public Health Communities." In W. R. Lovan, M. Murray, and R. Shaffer (eds.), *Participatory Governance: Planning, Conflict Mediation and Public Decision-Making in Civil Society*. London: Ashgate Press, 2004a.
- Dukes, E. F. "What We Know About Environmental Conflict Resolution: An Analysis Based on Research." *Conflict Resolution Quarterly*, 2004b, 22(1–2), 191–221.
- Emerson, K., and Carlson, C. "An Evaluation System for State and Federal Conflict Resolution Programs." In R. O'Leary and L. B. Bingham (eds.), *The Promise and Performance of Environmental Conflict Resolution*. Washington, D.C.: Resources for the Future, 2003.
- Emerson, K., Nabatchi, T., O'Leary, R., and Stephens, J. "The Challenges of Environmental Conflict Resolution." In R. O'Leary and L. B. Bingham (eds.), *The Promise and Performance of Environmental Conflict Resolution*. Washington, D.C.: Resources for the Future, 2003.
- Floyd, D. W., Germain, R. H., and Terhorst, K. "A Model for Assessing Negotiations and Mediation in Forest Resource Conflicts." *Journal of Forestry*, 1996, 94(5), 29–33.
- Folger, J. P., and Bush, R.A.B. "Transformative Mediation and Third-Party Intervention: Ten Hallmarks of a Transformative Approach to Practice." *Mediation Quarterly*, 1996, 13(4), 263–278.
- Frame, T. M., Gunton, T., and Day, J. C. "The Role of Collaborative Planning in Environmental Management: An Evaluation of Land and Resource Management Planning in British Columbia." *Journal of Environmental Planning and Management*, 2004, 47(1), 59–82.
- Hamilton, M. S. "Environmental Mediation: Requirements for Successful Institutionalization." In M. K. Mills (ed.), *Alternative Dispute Resolution in the Public Sector*. Chicago: Nelson-Hall, 1991.
- Innes, J. E., and Booher, D. E. "Consensus Building and Complex Adaptive Systems: A Framework for Evaluating Collaborative Planning." *Journal of the American Planning Association*, 1999, 65(4), 412–423.

- Innes, J. E., and Booher, D. E. "Reframing Public Participation: Strategies for the 21st Century." *Planning Theory & Practice*, 2004, 5(4), 419–436.
- Irwin, R. A., and Stansbury, J. "Citizen Participation in Decision Making: Is It Worth the Effort?" *Public Administration Review*, 2004, 64(1), 55–65.
- Kessler, B. L. *Stakeholder Participation: A Synthesis of Current Literature* prepared by National Marine Protected Areas Center in cooperation with the National Oceanic and Atmospheric Administration Coastal Service Center, 2004. Retrieved from www.mpa.gov.
- Kloppenborg, L. A. "Implementation of Court-Annexed Environmental Mediation: The District of Oregon Pilot Project." *Ohio State Journal on Dispute Resolution*, 2002, 17(3), 559–596.
- Kolb, D. M. *When Talk Works: Profiles of Mediators*. San Francisco: Jossey-Bass, 1997.
- Kreft, I., and De Leeuw, J. *Introducing Multilevel Modeling*. Thousand Oaks, Calif.: Sage, 1998.
- Leach, W. D. "Public Involvement in USDA Forest Service Policymaking: A Literature Review." *Journal of Forestry*, Jan./Feb. 2006, 43–49.
- Leach, W. D. "Collaboration and Its Alternatives: What We Know from Comparative Research on Environmental Governance." Presented at Symposium on Collaborative Governance in the West, Oregon State University, Oct. 1, 2007.
- Leach, W. D., Pelkey, N. W., and Sabatier, P. A. "Stakeholder Partnerships as Collaborative Policymaking: Evaluation Criteria Applied to Watershed Management in California and Washington." *Journal of Policy Analysis and Management*. 2002, 21(4), 645–670.
- McCarthy, J., and Shorett, A. *Negotiating Environmental Settlements: A Guide to Environmental Mediation*. New York: American Arbitration Association, 1984.
- McKinney, M., and Field, P. "Evaluating Community-Based Collaboration on Federal Land and Resources." *Society and Natural Resources*, 2006, 21(5), 419–429.
- McKnight, P. E., McKnight, K. M., Sidani, S., and Figueredo, A. J. *Missing Data: A Gentle Introduction*. New York: Guilford Press, 2007.
- Mills, M. K. *Alternative Dispute Resolution in the Public Sector*. Chicago: Nelson-Hall, 1991.
- Moore, C. W. *The Mediation Process: Practical Strategies for Resolving Conflict*. San Francisco: Jossey-Bass, 1986.
- National Environmental Conflict Resolution Advisory Committee. *Report to the U.S. Institute for Environmental Conflict Resolution*, 2005. Retrieved from www.ecr.gov.
- O'Connell, A. A., and McCoach, D. B. "Applications of Hierarchical Linear Modeling for Evaluations of Health Interventions: Demystifying the Methods and Interpretations of Multilevel Models." *Evaluation in the Health Professions*, 2004, 27, 119–151.

- O'Leary, R., and Bingham, L. B. (eds.). *The Promise and Performance of Environmental Conflict Resolution*. Washington, D.C.: Resources for the Future, 2003.
- O'Leary, R., and Husar, M. "What Environmental and Natural Resource Attorneys Really Think About ADR: A National Survey." *Natural Resources and Environment*, 2002, 16(4), 262–264.
- O'Leary, R., and Raines, S. S. "Lessons Learned from Two Decades of Alternative Dispute Resolution Programs and Processes at the United States Environmental Protection Agency." *Public Administration Review*, 2001, 61(6), 682–711.
- Oregon Department of Land Conservation and Development. *Collaborative Approaches to Decision Making and Conflict Resolution for Natural Resource and Land Use Issues*. Salem: Oregon Department of Land Conservation and Development, 1996.
- Orr, P., Emerson, K., and Keyes, D. "Environmental Conflict Resolution Practice and Performance: An Evaluation Framework." *Conflict Resolution Quarterly*, 2008, 25(3), 283–301.
- Policy Consensus Initiative. *A Practical Guide to Consensus*. Bismarck, N.D.: Policy Consensus Initiative, 1999.
- Raiffa, H. *The Art and Science of Negotiation*. Cambridge, MA: Harvard University Press, 1982.
- Ringdal, K. "Methods for Multilevel Analysis." *Acta Sociologica*, 1992, 35, 234–243.
- Schafer, J. L., and Graham, J. W. "Missing Data: Our View of the State of the Art." *Psychological Methods*, 2002, 7(2), 147–177.
- Singer, J. D. "Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models." *Journal of Educational and Behavioral Statistics*, 1998, 24, 323–355.
- Singer, J. D. *Federal Dispute Resolution Using ADR with the United States Government*. San Francisco: Jossey-Bass, 2004.
- Sipe, N., and Stifel, B. "Mediating Environmental Enforcement Disputes: How Well Does It Work?" *Environmental Impact Assessment Review*, 1995, 25, 139–156.
- Sipe, N. G. "An Empirical Analysis of Environmental Mediation." *Journal of American Planning Association*, 1998, 64(3), 275–285.
- Society of Professionals in Dispute Resolution Environment, Public Disputes Sector, Critical Issues Committee (SPIDR). *Best Practices for Government Agencies: Guidelines for Using Collaborative Agreement-Seeking Processes*. Washington, D.C.: Society of Professionals in Dispute Resolution, 1997.
- Sullivan, L. M., Dukes, A. K., and Losina, E. "Tutorial in Biostatistics: An Introduction to Hierarchical Linear Modeling." *Statistical Medicine*, 1999, 18, 855–888.
- Susskind, L., and Cruikshank, J. *Breaking the Impasse: Consensual Approaches to Resolving Public Disputes*, New York: Basic Books, 1987.

- Susskind, L., McKernan, S., and Thomas-Larmer, J. (eds.). *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*. Thousand Oaks, Calif.: Sage, 1999.
- Susskind, L., and Ozawa, C. "Mediated Negotiation in the Public Sector." *American Behavioral Scientist*, Nov./Dec. 1983, 27, 255–279.
- Susskind, L., and Thomas-Larmer, J. "Conducting a Conflict Assessment." In L. Susskind, S. McKernan, and J. Thomas-Larmer (eds.), *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*. Thousand Oaks, Calif.: Sage, 1999.
- Susskind, L. M., van der Wansem, and Ciccarelli, A. *Mediating Land Use Disputes*. Cambridge, Mass.: Lincoln Institute of Land Policy, 2000.
- Talbot, A. R. *Settling Things: Six Case Studies in Environmental Mediation*. Washington, D.C.: Conservation Foundation, 1983.
- Todd, S. "Measuring the Effectiveness of Environmental Dispute Settlement Efforts." *Environmental Impact Assessment Review*, 2001, 21(11), 97–110.
- U.S. Office of Management and Budget and President's Council on Environmental Quality. "Memorandum on Environmental Conflict Resolution." Nov. 5, 2005. Retrieved from <http://www.whitehouse.gov/ceq/joint-statement.html>.

Kirk Emerson is the former director of the U.S. Institute for Environmental Conflict Resolution of the Morris K. Udall Foundation. She is currently a visiting professor at the University of Arizona's School of Public Administration and Policy and the Udall Center for Studies in Public Policy and maintains a small ECR practice.

Patricia J. Orr is a program manager for evaluation at the U.S. Institute for Environmental Conflict Resolution.

Dale L. Keyes is a former senior program manager at the U.S. Institute for Environmental Conflict Resolution and is currently practicing as an independent mediator.

Katherine M. McKnight is a program evaluator and statistical consultant with Public Interest Research Service.