(How Well) Do Collaborative Processes Work? Scale Effects in Evaluation Approaches

Sanda Kaufman (Cleveland State University)
Connie Ozawa (Portland State University)
Deborah Shmueli (University of Haifa)

Paper Presented at the
23rd Annual International Association of Conflict Management Conference
Boston, Massachusetts
June 24 – 27, 2010

Abstract:
Evaluation of collaborative processes is an important source of insights for improving public decisions and conflict management practice, as well as a basis for validation of individual interventions for stakeholders, funders, and the field of conflict management as a whole. In this article, we conducted an exploration of how best to use information about past collaborative processes to inform future collaborative practices in environmental and other public decisions. Our ultimate purpose is to characterize the extent and kinds of information about collaborative processes that might help to fully understand and characterize what worked and what did not work in specific collaborative planning processes. We aim to pinpoint the kind of evaluation that contributes to all stages of collaborative processes from assessment to the implementation of agreements, as well as to the reflective practice of interveners. The U.S Institute of Environmental Conflict Resolution (USIECR) engaged in the careful design of a data collection instrument and of a data base for accumulating information about environmental conflicts handled through intervention and consensus building. We discuss the USIECR survey instrument and what it can yield. We describe briefly one Oregon mediation case (Aggregate Mining) for which participants had responded to the USIECR survey. After conducting interviews with some of the participants and interveners in this case, we compare the insights gained through direct interviewing to those that can be derived from the USIECR database, to assess what each approach can contribute to improving collaborative decision making practice.

Extended Abstract

(How Well) Do Collaborative Processes Work? Scale Effects in Evaluation Approaches

In conflicts over environmental and other public decisions, evaluation is an important source of validation for interventions, their funders, and the field of conflict management as a whole.

Evaluation is inevitably a political act in that it serves someone's purpose - be it to promote, enhance or fund collaborative processes in the public domain, or to encourage resolution of public disputes by other means deemed more effective than through collaboration. It is often argued, however, that evaluation contributes to systematic reflection, which in turn plays a key role in efforts to manage conflicts and to engage in collaborative public decision making. Evaluation results are essential for reflective practice, as they accumulate and distil the learning and knowledge necessary to design, implement and recalibrate collaborative decision processes.

Although project evaluation is an established practice in other areas, collaborative environmental and other public decision processes are still coming to terms with evaluation as a routine component of any project. These conflict contexts have posed challenges to evaluation design, which can be characterized as plagued by methodological anarchy (Lewis, 2004). As a result, evaluation is mostly summative, occurring sometimes well after a project has been concluded. That in turn creates difficulties for the data collection and capacity to capture in-the-moment process features that may account for the outcomes. When studied retrospectively with 20/20 hindsight, these features may distort our understanding of relationships, stakes, and moves that led to success or failure of collaboratives. Participants are difficult to locate, reluctant to participate, and their memories are “contaminated” with after-the-fact knowledge and more recent related events. On the other hand, retrospective explorations have some benefits when problems tackled by the collaborative have been resolved, or are no longer an issue, and respondents are not involved in an active dispute. Those who participate in evaluation may be in a better position to reflect on their own and others’ actions. With some distance, they may discern turning points and critical events that affected the course of the collaborative, and that are more difficult to recognize when one is immersed in the day-to-day details of a process.

Evaluation efforts are often funder-driven, and therefore necessarily motivated by the funders’ interest in ensuring accountability (and, at times, success) of interventions. Not surprisingly, such evaluations “deliver.” However, they do not necessarily provide the kinds of insights that can contribute to the improvement of conflict management or to public decision making practice. Those who try to promote and inform these practices often rely on anecdotal evidence that collaborative processes and various kinds of intervention – facilitation, mediation – work, in that they yield better decisions, are less costly and faster, and foster positive relationships, compared to adversarial alternatives such as litigation. However, those who have designed, conducted, facilitated or participated in

collaborative decision processes know how likely it is for at least one of these touted advantages not to materialize in specific cases. They also know that collaboratives yield some valuable intangibles that resist quantification or clear tradeoffs with time or resources.

Theory and experience tell us that to be useful an evaluation requires a clear sense of: the outcomes it seeks, how these outcomes are to be used, and the intended audiences. It is also important to identify who might best conduct the evaluations, and what tools and techniques will be used to conduct them. In practice, we see effort, time and resources spent developing “generic” instruments that are flexible enough to work in a broad range of contexts and serve multiple audiences and purposes, possibly serving none very well. The quest for broad applicability tends to minimize important situational factors that comprise the core of environmental and other public and conflicts. Although various aspects or components of one situation are frequently found in others, together they make for unique cases whose understanding may be flawed in the absence of information about the unique traits.

Why are context specifics so necessary for understanding why collaborative processes succeed or fail? We argue that it is not possible to understand these processes or to draw any practice lessons for the future without being able to link them to the contexts that shaped them. Otherwise, we might end up with “generic,” context-free extrapolated recommendations that do not work particularly well in any context. We also propose that because of this linkage our recommendations to intervention practitioners should be offered in contingent form: “if… then…” that enables both understanding and tailoring process design to enhance the likelihood of success.

Since evaluation is a crucial contributor to the improvement of collaborative decision making in the environmental and other public areas, and since it can be resource- and labor-intensive, it is necessary to understand and perfect our approaches, to make the best of such efforts in informing conflict management theory and practice. To assist efforts to evaluate environmental and public decision collaboratives we explore in this article the usefulness of one evaluation tool that strives to serve multiple audiences. It was designed by the U.S Institute of Environmental Conflict Resolution (USIECR) for the description and evaluation of a large number of cases. We find that the resulting database is useful for extracting general trends of some specific elements across varying contexts. However, to build up its database, USIECR depends on the willingness of participants to various projects to use this instrument. Herein lies an inherent clash between the tool, developed on a national scale for national purposes, and its reliance on data from projects implemented at local scales that do not always generalize meaningfully to the intended national scale.

The opportunity to test our observation regarding the limited usefulness of "generic" evaluation instruments arose at the confluence of a request for help from an organization providing assistance for collaborative processes in Oregon, and the USIECR evaluation initiative. USIECR engaged in the careful design of a data collection instrument and of a data base for accumulating information about environmental conflicts handled through intervention and consensus building. The purpose of this data base was to enable evaluation and comparison of a large

number of cases by overcoming the difficulties associated with their context dependence and variability along numerous dimensions such as location, issues, stakes, scale, duration, number and nature of parties, type and quality of the intervention. The USIECR's primary concern was with characterizing the usefulness of intervention and consensus building approaches, and confirming the largely anecdotal evidence that on average consensus building and mediation are more successful, satisfying and effective than alternatives such as litigation. For example, if data indicate that in many of the cases where consensus building was used respondents are satisfied with process and outcome, it might be easier to make a case for using and funding such processes, incorporating them into process design from the outset, and even institutionalizing them at the federal, state and local levels.

In this article, we explore how best to use information about past collaborative efforts to inform future collaborative decision practices. Our ultimate purpose is to characterize the extent and kinds of information that might help to more fully understand what worked and what did not work in specific collaborative processes. We aim to pinpoint the kind of evaluation that contributes to all stages of collaborative processes from assessment to the implementation of agreements, as well as to the interveners' reflective practice.

This is admittedly a tall order. Therefore, we began by assessing which existing evaluation practices already serve this purpose either partially or entirely. Since the USIECR instrument was the result of contributions by numerous conflict management researchers and practitioners, it is arguably at the cutting edge of comparative evaluation. Therefore, it is the natural term of comparison when trying to assess the extent to which such instruments can be used to enhance understanding and improve conflict management and collaborative practices. The USIECR instrument, which has been in use since 2005 and addresses itself to conflicting parties as well as interveners, is quite extensive (77 open- and closed-ended questions for participants, and 40 for interveners), leading to the hope that it might help illuminate intervention and collaborative practices. We describe here the USIECR instrument and what it can yield. Next we describe briefly one case in which participants had responded to the USIECR survey. After conducting interviews with some of the participants and interveners in this case, we compare the insights gained through direct interviewing to those that can be derived from the USIECR database, to assess what each approach can contribute to improving collaborative decision making.

In examining the USIECR survey instrument, we were particularly interested in the extent to which the responses might help inform the programs offering facilitation assistance (such as the OCP), or the facilitator/mediators involved. We find that having aimed for process and outcome data to the near-total exclusion of context and content, USIECR formulated the questions in ways that ensured a good fit with current claims about the benefits of collaboratives and of mediation. In the process inadvertently (or by design?) the survey designers made it almost impossible for any kind of information to seep in that might surprise, contradict the current collective mental model of what a good collaborative process is and how it works, or offer new insights that are not currently part of this mental model but maybe should. In other words, it is a confirmatory rather than exploratory (learning) tool,

Thus it is very unlikely that we might discover anything specific to a case that the survey designers have not thought about in advance (and asked). As well, unless a large proportion of the cases ended in failure to achieve agreement (which is unlikely), the designers' mental model of the collaborative processes cannot but be confirmed. It is as if the survey asked "Isn't it true that collaboratives work exactly as we think they do?" with a broad range of answers being consistent with this model (as well as with others, however, whence the inability to test this model meaningfully).

For example, the survey asks about adequate representation, improved relationships, inclusion of issues – all argued in the conflict management and mediation literature to affect the likelihood and quality of agreements - and satisfaction with the mediation, However, in the absence of any other information about the case, we cannot judge, for example, whether the adequate representation, improved relationships or coverage of issues were due to a skillful mediator, or whether mediation was successful because a convener had taken care of ensuring good representation and a complete issue set (or even because the case was rather simple, or had few or disengaged participants). Neither can we pick up trade-offs or test if they are possible, given that no process is ever perfect: is it possible that if some of the ideal conditions are not fulfilled they can be compensated by others? For example, if there is no adequate representation is the process doomed or is there anything that can be done (and was done in some cases) to make up for this shortcoming? Is representation as important as comprehensive inclusion of issues, or might it be the case at times that with strong representation issues that were not all considered from the outset can be added later?

The USIECR open-ended questions can yield some interesting insights (we discuss some examples in connection with the case on which we focused), but while they are useful in diffusing the respondents' frustration by allowing them to qualify their responses on the scales, they are very unlikely to be useful in any other way than as vignettes in a report. They are not in a format given to statistical analysis, refer to context-, case- or participant-specifics (so they do not generalize) and without intense processing they cannot readily be used for the USIECR goals. This is an unfortunate waste of information, as well as a problematic request for respondents to take (quite a lot of) time to offer responses (in which they invest themselves emotionally) that will not be very important or useful to the USIECR endeavor.

Since the USIECR instrument uses mostly Likert-type scales the results are seemingly amenable to quantitative analyses across cases. To what extent is this instrument useful despite the limited statistical analysis possible with the small numbers of cases or the small groups? To what extent is the format of the survey – aimed at capturing a variety of contexts - reducing the information to comparables, and in the process losing specifics key to our understanding? Is it possible to enhance the usefulness of the elicited information through slight modifications that might yield substantially more beneficial information for a specific program or facilitator/mediator? Or is such a survey, designed as it is for statistical comparability, inherently inadequate for informing the reflective practice in specific contexts?

At the encouragement of USIECR, the Oregon Consensus Program (OCP), a program of the National Policy Consensus Center at Portland State University, assessed ten of their collaborative problem solving cases using the USIECR instrument. At this time, the relatively small number of Oregon cases does not allow the OCP to draw statistically meaningful conclusions from the data collected for the USIECR. As the OCP gathers more data, this limitation may be overcome. We examine here to what extent using the USIECR instrument (which requires time and resources) can serve the local participants or mediators, and we find that its current contribution is rather marginal.

Can the current USIECR (participant and mediator) survey instruments be adapted to better inform the Oregon Consensus Program, while still providing USIECR the data it seeks? Are there data that can be useful at both the national and the local scale (USIECR and Oregon)? If not, what must be added to enable formative, summative and knowledge-based evaluations (Elliot et al. 2003) at the more immediate level of local program effectiveness?

To address these questions we describe here one of the OCP cases – the Aggregate Mining (failed) collaborative - which we selected for detailed analysis to explore the value added by eliciting context and more elaborate open-ended responses from some of those who participated in the USIECR survey. Our approach involves analysis of the participant and mediator USIECR survey and follow-up interviews with participants and mediators. We develop recommendations for the kinds of data apt to provide insight that can make generic multi-scale evaluation instruments more effective for their own (national-scale) purposes as well as for the local scales from which the data are collected.

References

Ackterman, G. and team, 2005. Preliminary Summary of Aggregate Mining in Oregon with emphasis in the Willamette River Basin, August, Institute of Natural Resources, Oregon State University, unpublished???

Church, C. and J. Shouldice, 2002. The Evaluation of Conflict Resolution Interventions: Framing the State of Play. Letterkenny: INCORE.

d'Estree, T.P. , Fast, L.A., Weiss, J.N., & Jakobsen, M. S., 2001. Changing the Debate about "Success" in Conflict Resolution Efforts. *Negotiation Journal,* 17(2), 101-113.

Elliott, M., d'Estree T.P. and S. Kaufman, 2003. The Role of evaluation in Resolving Intractable Conflicts. *Beyond Intractability.* Ed. Guy Burgess and Heidi Burgess. DATE. Conflict Research Consortium, University of Colorado, Boulder, Colorado, USA.

Innes, J. E., 1999. Evaluating consensus building, In *The Consensus Building Handbook: a Comprehensive Guide to Reaching Agreement*. L. Susskind, S. McKearnan and J. Thomas-Larmer. Thousand Oaks, CA: Sage.

Lewis, H., 2004. Evaluation and Assessment of Interventions. *Beyond Intractability.* Ed. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder, Colorado, USA.

Orr, P., 2009. Environmental Conflict Resolution: Performance Evidence from the Field, memo?? January. U.S. Institute for Environmental Conflict Resolution, Tucson, AR.

Orr, P., Emerson, K. and D. Keyes, 2008. Environmental Conflict Resolution Practice and Performance: An Evaluation Framework, *Conflict Resolution Quarterly*, 25(3), 283-301.