

PROGRAM EVALUATION

Program evaluation traditionally asks two types of questions: Summative evaluation asks questions such as “does the program work.” Formative evaluation asks questions such as “Is the program constructed appropriately to achieve its desired goals.” From the start, case by case project evaluation faces several challenges. It is near impossible to conduct a randomized experiment, because no two disputes are alike and no two ADR projects are alike. Case-by-case project evaluation of ADR processes cannot achieve quasi-experimental status, either, because there is no control group to be had. Because of these limitations it is difficult for ADR evaluations to offer substantive inference and attribution. Yet another is that project evaluation often seeks to answer both formative and summative questions simultaneously, which causes significant design and usability hurdles.

However, this is not an impasse. Large-scale evaluations of ADR as a form of intervention have relied on broadly gathered empirical research. There have been successful theoretical comparisons made with control groups and control processes (alternatives to the ADR process such as litigation). This is beyond our current purpose. The purpose of creating a project evaluation instrument for the Ruckelshaus Center is to provide The Center with case by case feedback so that they may capture lessons learned to continually improve their project processes.

Nevertheless, in order to make sure that the project evaluation instrument is sound, I consider several standard program evaluation principles and strive to avoid certain pitfalls. I first briefly describe the principals and pitfalls. In the Tradeoffs section I analyze instrument style, instrument implementation, and instrument question-types accordingly. In addition, I will analyze the aforementioned three areas of instrumentation in terms of transactional cost, the amplitude with which they can achieve the Ruckelshaus Center’s project evaluation goals and in some cases the usability of the instrument.

Program Evaluation Principles

The project evaluation instrument that this inquiry will produce must take many issues into consideration in its design in order to maximize its usefulness. Typically in program evaluation there are more issues to consider than the following: these are the most primary to a project evaluation instrument for the Ruckelshaus Center.

Reliability

In the case of a project evaluation, the reliability of an instrument is the extent that an instrument produces the same results when assessing one measure multiple times. For example, a thermometer that returned different readings with multiple contiguous uses would be an unreliable indicator of a fever. In that light, an evaluation instrument must contain a degree of specificity that minimizes rapid changes in interpretation in the individual responding to the evaluation inquiry.

Validity

Validity in evaluation is the degree that a measure captures the information it is intending to capture. In the ambiguous category of an alternative dispute resolution process's success, what success means, or what various outcomes could be reasonably thought of as successful, would have to be clearly spelled out ahead of time. If not, trying to capture the degree to which a process succeeded would be left to the mercy of interpretations that could be contrary to the processes intended purposes.

Sensitivity

Sensitivity is essentially the level of nuance that a measure is able to incorporate into its findings. For example, a project evaluation seeking to understand the level of participant's satisfaction asking only if participants were 'satisfied or not' would result in potentially superficial results. On the other hand, if that same measure asked participants to report their level of satisfaction with specific activities, it would generate more detailed and usable information. Another example is if an ADR process was interested in transforming relationships between participants, then the evaluation of the process would have to be able to capture the nature of participant's relationships before and after the ADR process, and it must be sensitive enough to discern these changes.

*Biases to project evaluation**Unintended Consequences*

In general, project evaluation must be mindful of creating effects that the evaluation does not seek to create. For example, an evaluation of a dispute resolution process implemented in an untimely manner could create political pressures among the stakeholders. An evaluation of the skill of the facilitator could fall into the wrong hands and create negative consequences for that facilitator's career.

The Hawthorne Effect

A famous unintended consequence of evaluation is the Hawthorne Effect, in which participants of the evaluation alter their behavior as a result of their knowing they are being evaluated.

History Biases

History biases occur when an event outside of the project impacts project participants. For example, if a stakeholder group in an ongoing land use dispute secures a federal grant during an ongoing mediation, they may suddenly be amenable to an entirely different range of financial outcomes. Their increase in ability to come to agreement on financial fronts was not caused by the mediation process, so an evaluation of the mediation should not attribute all of this group's change in attitude to the mediation process. Because of this a project evaluation has to separate outside from inside events, and weed out irrelevant information. It needs to be constructed in such a way that captures only process relevant information. Otherwise the evaluation will create reports based not on the process it seeks to evaluate, but on causes far beyond the process scope.

Maturation Biases

People continue to grow and learn and age through a project. A Maturation issue is when a participant gains skills outside of the project being evaluated. A basic example: if a program were to attempt to increase the math skills of a young person, the pre and post evaluation may capture not only the benefits from the math program, but also the normal course of childhood intellectual development. A participant in a facilitated consensus-seeking dispute may become curious and decide to read books on consensus building. That participant will gain skills that aid them in the facilitation from a source outside of the facilitation process. Because of this an evaluation instrument that sought to capture a participant's gain in negotiation skills as a result of the facilitation process would potentially misattribute its findings.

Instrumentation Biases

If a neutral third party observer is used to evaluate a process, that observer may become tired or bored over time. In this case the instrument alters during the process and the results produced are altered as well. In the case of a pre and post evaluation, if changes are made to the evaluation instrument in the interim, the results produced could be attributed to the altered instrument rather than the project being evaluated.

Dropout Biases

Dropout threats are very important to project evaluation for The Center. If some number of participants in a Center project quit halfway through, and only those participants who complete the process participate in the evaluation, then a very important segment of stakeholders are not represented and the evaluation results become skewed. This is important as The Center's purpose in evaluation is to create lessons learned and foster process improvements: project participants who drop out are a highly valuable source of information and feedback on what went wrong.

Reasonability Biases

If the project evaluation relies on responses from participants, then what the participants are asked must be within the realm of their ability to respond. Overly complex or theoretical inquires, though interesting to the analyst, could be unanswerable by project participants.

Selection Biases

Since project participants submit data voluntarily, participants self-select both who will respond and what data researchers will and will not receive.

Transactional Cost

Transactional cost is a measure of the time, effort and resources required by any method of project evaluation weighted against the quality of information produced. If the project evaluation instrument requires a great deal of time and effort from either The Center or respondents, the quality and/or completeness of information gathered may be compromised. For example, electronic resources like Survey Monkey are inexpensive, but may result in low response rates. A resulting data

set may be too small to be useful. If the instrument is administered by hand via one-on-one interview techniques it may generate a high quantity and quality of information, but at high cost in terms of time, budget and effort for both The Center and the interviewee.

Usefulness of Information

Because of reasonability and transactional cost, The Center faces some constraints. Project evaluation cannot be so large or complex that evaluation participants are unable to provide credible feedback. Further, The Center has specific purposes for evaluating its projects: as stated, lessons learned and process improvements. The Center's project evaluation instrument must be mindful of this and use the *types* of questions and *method* of evaluation that help respondents to provide good feedback that directly helps The Center meet its goals.