

## Introduction to Rapid Impact Evaluation (RIE)

**Andy Rowe**

ARCeconomics and School of Public Administration, University of Victoria, Victoria, British Columbia, Canada

**First Published** September 19, 2019 Research Article

<https://doi.org/10.1177/1356389019870213>

Rapid Impact Evaluation (RIE) offers the potential to evaluate impacts in both ex ante and ex post settings providing utility for developmental and formative evaluation as well as the usual summative settings. RIE triangulates judgments of three separate groups of experts to assess the incremental change in effects attributable to the program. Three methodological innovations are central to the method; the scenario-based counterfactual, a simplified approach to measuring change in effects and an interest-based approach to stakeholder engagement. In evaluations to date RIE has proved a very cost effective and nimble approach to assessing impacts and does not intrude on design or implementation of the program. By applying recent thinking on use-seeking research emphasizing joint knowledge processes over knowledge products RIE promotes salience, legitimacy and credibility with decision makers and key stakeholders. Applications of RIE show it to be fit for purpose.

L'évaluation rapide des impacts (ERI) offre la possibilité d'évaluer les impacts dans les contextes ex ante et ex post, ce qui est utile pour les évaluations évolutive et formative ainsi que pour les travaux sommatifs habituels. L'ERI triangule les jugements de trois groupes d'experts distincts afin d'évaluer les effets spécifiquement attribuables au programme. Trois innovations méthodologiques sont au coeur de la méthode : le scénario contrefactuel, une approche simplifiée pour mesurer le changement dans les effets et un engagement des parties prenantes selon une approche basée sur les intérêts. À ce jour, dans les évaluations, l'ERI s'est avérée une approche rentable et souple pour évaluer les impacts; de plus, elle n'empiète pas sur la conception ou la mise en oeuvre du programme. Appliquant une réflexion récente sur la recherche axée sur l'utilisation qui met l'accent sur les processus conjoints de connaissance plutôt que sur les produits de connaissance, l'ERI favorise la visibilité, la légitimité et la crédibilité auprès des décideurs et des principaux acteurs. Les applications de l'ERI montrent qu'elle est adaptée à son objectif.

## Introduction

Evaluating impacts is often associated with high stakes summative evaluation settings addressing questions about expanding, replicating or curtailing an intervention or identifying which approaches are most effective at addressing a problem. High stakes questions invite methods with requirements that limit the applicability of impact evaluation such as higher costs, intruding into the intervention, limiting opportunities for participation in the evaluation by staff and program participants, timeliness and legal and ethical limitations. Commissioning and using some forms of impact evaluation can be beyond the capacity of many programs and for many the impacts of concern might not be observable for decades. Impact evaluation methods can also be too costly or insufficiently adaptable for many formative and developmental (ex ante) evaluation settings. These are some of the limitations on the feasibility and use of impact evaluation and mean that impact evaluation is very much a work in progress (Stern, et al., 2012) (Blatterman, 2008) (White, 2006) (GAO, 2009).

Rapid Impact Evaluation (RIE) is a theory-based approach (Rogers, Petrosino, Huebner, & Hacsí, 2000) developed to provide a structured assessment of impacts in settings where the main existing impact evaluation methods are not possible or feasible but evaluation of impacts is still needed. The approach was designed to be low cost and nimble so that it could be applied in a wide range of evaluation settings such as sustainability with complex coupled human and natural system settings often having quite different spatial and temporal scales (Rowe, 2012). A happy consequence is that RIE can be usefully applied in ex ante as well as ex post settings and so can be used as part of formative or developmental evaluations.

The RIE approach is based on three new evaluation methods:

- New metrics for assessing impacts
- The scenario-based counterfactual, a new approach to counterfactuals
- Interest-based approaches to stakeholders

These are also proving useful for many mixed methods evaluations, for example the metrics have been incorporated into surveys, group processes and interviews to assess impacts and combined with scenario-based counterfactuals to assess the incremental contribution of the program to these impacts. Interest-based approaches can be applied to all evaluations.

Based on the premise that without use there is no utility in being rapid RIE is built around a use-seeking framework making this an evaluation approach with use and influence a central design feature. The full RIE approach triangulates the judgments of three separate and distinct groups of experts to assess the incremental change in effects attributable to the program. In hindsight and not recognized until the approach had been developed and applied, RIE has many similarities to other triangulation approaches such as the Delphi method and other triangulation approaches (Carugi, 2016); to risk assessment that employs likelihood and magnitude as primary metric; Bayesian Networks to predict likelihood of contributions to an outcome from a connected network of observable outcomes (think theories of

change loaded with probabilities)<sup>1</sup>; and Structured Analogies where experts match determinative to target outcomes in forecasting (Green & Armstrong, 2004). RIE is distinct in (unwittingly) incorporating the main features of each into a single evaluation approach. The use-seeking approach is based on work done by the Conservation and Science Program at the Packard Foundation (Rowe & Lee, 2012) to develop and apply a use-inspired approach to science philanthropy itself based on the valuable contributions of Bill Clark and colleagues to understanding use of science knowledge (Clark, Mitchell, & Cash, 2006), Jacobs et al. (2007) and others.

Applications show RIE to be fit for purpose and a useful addition to the evaluator toolkit. RIE has been piloted and accepted as an approved evaluation approach for Canada's National Evaluation Policy and has been used in evaluations conducted for US EPA, US Interior, six federal departments in Canada and the Global Environment Facility (GEF). Of especial note is that RIE can be employed to assess impacts in two-system settings such as coupled human and natural systems and that RIE is one of the very few evaluation approaches addressing natural systems.. A separate paper will describe recent applications of RIE.

The purpose of this article is to introduce RIE to the evaluation field and to stimulate additional applications to test and refine the approach as part of RIE entering fully into the evaluation methodological establishment. An overview and short description of the new methods is followed by an extended discussion of how RIE works.

### *Background and Overview of Rapid Impact Evaluation*

#### **Background**

The RIE method was developed in 2003-4 to evaluate the impacts of natural resource management decisions reached using mediated decision processes, sometimes termed environmental conflict resolution or ECR (Susskind & Weinstein, 1980) (Bingham, Emerson, Nabatchi, & O'Leary, 2003) typically using a third-party facilitator or mediator<sup>2</sup>. Then existing approaches for evaluating the impacts of mediated decision making were recognized as inadequate (Rowe, 2004) (Brogden, 2003) (Koontz & Thomas, 2006) (Todd, 2001). Development of RIE was originally funded by the Conflict Resolution Program of the William and Flora Hewlett Foundation (Hovick, 2005) as part of the program's final round of grants to develop an approach to evaluating the contribution of different decision processes to substantive outcomes in human and natural systems. Ironically the Conflict Resolution program was terminated in part because its efficacy could not be demonstrated to the satisfaction of the Executive of the Foundation. While RIE was initially developed for ECR settings, it has been successfully applied to a broader range of evaluations as well as complex settings where many impact evaluation methods are

---

<sup>1</sup> The most useful overviews of Bayes Networks are web sites such as Bayesialab <https://www.bayesia.com/bayesian-networks-examples>

<sup>2</sup> Interestingly, the conflict resolution field employs a counterfactual-like concept, the BATNA (Best Alternative to a Negotiated Agreement) (Fisher, Ury, & Patton, 2011) as a minimum threshold that a negotiator should accept in an agreement.

seriously challenged such as natural resource management, climate and sustainable development interventions, research impacts and for evaluating multiple Sustainable Development Goals (SDGs).

RIE is a practitioner-developed approach. Practitioners do not have the resources for development, testing and dissemination of methods. The contributions of other evaluators and interest of evaluation commissioners in piloting this new method has enabled RIE to advance to the current stage, and while this has caused testing and dissemination to be episodic it has also contributed to very valuable adaptation of the approach and ensured that feasibility was always an important concern<sup>3</sup>.

RIE has methodological elements that have sometimes challenged those using the approach, training participants and readers of earlier drafts of this paper. Text boxes are used to highlight these challenging concepts.

### **Overview of RIE**

RIE incorporates three new methods in a use-seeking frame that can be applied together or individually in most mixed methods settings including ex ante as well as ex post. The new methods are:

1. A simplified metric to assess impacts based on the premise that the main sources of variation in estimating a given effect is the likelihood of the effect occurring and its magnitude. The metric generates an index of change for each effect referred to as the RIE index.
2. The scenario-based counterfactual is a new type of counterfactual for impact evaluation. It is based on alternative approaches to an intervention that are plausible, efficacious, feasible, legal and ethical sometimes drawing on alternatives such as those defined in regulations or legislation, or those applied in similar settings elsewhere, that were considered but not implemented for the intervention being evaluated, or those being considered as options for modifying the approach.
3. The underlying premise of interest-based approaches to stakeholders is that each interest has worldviews that shape and influence their assessments and that combining the worldviews of all interests who can influence the intervention and those affected by the intervention offsets potential bias from a narrower approach to stakeholders. RIE accords each interest one “vote” in calculating impacts regardless of the number of parties representing an interest.

RIE impact metrics are used by three distinct expert groups to assess the key effects under the current approach and the scenario-based counterfactual, the difference between the current approach and the counterfactual is the estimate of the contribution of the program to the impact. The resulting assessments of the three types of experts usually follow a similar pattern and collectively provide a range of achievement of effects.

### **RIE is built around use and influence**

RIE was developed with use strongly in mind, originally with guidance from the extensive literature on evaluation use and subsequently applying what we know about use of science knowledge. We use the

---

<sup>3</sup> FN for those who commented on drafts, supported applications

theory of change on use-inspired research developed by the author and Kai Lee of the Packard Foundation that has been applied in the Foundation's Science program for over five years drawn from the work of Bill Clark and colleagues, Jacobs et al., and others who observed that prospects for use are enhanced when decision makers and key stakeholders regard the research as salient (relevant, timely), legitimate (fair, unbiased, respectful, feasible), and credible (true, technically appropriate handling of evidence). The mechanism for achieving these values is a joint knowledge process with the scientists, decision makers and stakeholders and emphasizing the knowledge process over knowledge products (Clark, Mitchell, & Cash, 2006, pp. 14-15).

Joint knowledge production is central to all phases of RIE: for example the RIE approach seeks consensus among interests on the key characteristics of the intervention and engages all interests in assessing the effects. We refer to *convening* parties (the party or parties who initiate the decision process and usually have the actual authority to make the decision), *core* (parties who actively participate in the decision process and are able to affect the decision or its implementation) and also *core* those directly or strongly affected by the intervention and *non-core* (other parties who engage/observe but are not central to or strongly affected by the discussions).

### **RIE and expert judgement**

RIE facilitates judgements of experts in the science or subject matters and experts in the intervention to provide assessments of impacts of the intervention. RIE processes align closely with the guidance on expert assessment released several years after the development and many applications of RIE. This alignment is good news for both RIE and the guidance (California Ocean Science Trust, 2013). The guidance emphasizes the importance of the key features of RIE, for example what they term guiding values include credibility, legitimacy and salience; the importance of experts who are to work constructively with others as a criteria for joining in group processes, consideration of the limits to generalizing expert judgements, appropriate granularity of expert judgements and addressing disagreements and a range in views of experts. Other useful guidance on expert judgement is also consistent with the RIE approach (OECD, 2015)

### Interests and parties

An **interest** is defined by a common worldview. For example, many environmental organisations share a conservation worldview. This is a very different perspective from resource extraction organisations such as fishing, logging, mining or utilities, and from regulators, program managers, delivery and beneficiaries. There are multiple government interests, for example federal and Indigenous governments.

A **party** is an organization or individual(s) within an interest, for example individual environmental groups are parties that share a common interest. Within parties we refer to convening, core and non-core parties reflecting the strength and character of their relationship to the intervention and their standing in relation to decisions.

## How RIE Works

The joint knowledge process including the evaluator and representatives of interests is a thematic characteristic of RIE, beginning at the outset with a consensus-seeking approach to determine the necessary elements for the evaluation and continuing through the assessment processes. In this section we use the thematic joint knowledge processes as the storyline to describe how RIE assesses impacts, illustrated using an example that we now briefly describe.

### *Example*

The evaluation of the decommissioning of the Marmot Dam (Keller, 2009) was one of the Oregon pilots conducted in 2004 when RIE was still very much in the formative stage. The illustration has been updated slightly to reflect the methods as refined through subsequent application and reflection.

*In the early 1990's, Portland General Electric (PGE) faced an important relicensing decision relating to its Bull Run hydropower generating facilities requiring a proposal to the Federal Energy Regulatory Commission (FERC). PGE realized that as a condition of a new license, it would likely have to install costly fish passage upgrades, which would have raised the cost of operating a project that supplied less than one percent of Portland's electricity. Instead of relicensing, PGE decided to decommission the dams, but little was known about the impacts on water quality and fish habitat of the release of 90 years of sediment accumulation from behind the dams. To address these problems, PGE established a Decommissioning Work Group, which negotiated a consensus agreement for the removal of two dams and provided for donation of land and water rights by PGE to the Western Rivers Conservancy and the Oregon Water Resources Department (Rowe, Colby, Niemeyer, & Hall, Evaluating Environmental and Economic Effects of Collaborative Decisions, 2004).*

Some years earlier PGE had adopted collaborative processes for its licensing proposals. This differs from the approach favored by most utilities where the utility submits a proposal to FERC and external parties may choose to intervene with litigation, often brought under the Endangered Species Act (ESA). A mediator facilitated the Marmot dam process taking approximately nine months with participation of twenty-five participants from twenty-three parties. Despite wave tank simulations and other applied research undertaken as background to the decision the science was very inconclusive about what would happen to the Sandy River when the dam was removed (O'Connor, Major, & Grant, 2008). The decision was therefore undertaken with moderately high levels of uncertainty and ambiguity, knowing that do-overs were not possible once the decision was implemented by breaching the dam. Our evaluation occurred approximately eighteen months following the decision but prior to implementation.

Hydro generation cases often involve four different government interests (federal, tribal, state and local/regional), environmental conservation interests, commercial resource user interests (utilities, settler and tribal commercial fishing, irrigators), traditional and Indigenous users (e.g. ceremonial, archeological) and recreational resource users (anglers, white water users). A party can be represented by one or more participants and one or more parties might be associated with an interest.

## Staffing for a RIE

An experienced evaluator accustomed to working adaptively with mixed methods and having very strong facilitation skills and one or more Technical Advisors comprise the necessary staffing for a RIE. The technical advisors are identified and contracted early in the design once the evaluator understands the key knowledge domains and likely mechanisms of change for the intervention sufficiently to inform selection of appropriate Technical Advisor(s). The role of Technical Advisor(s) is to bring relevant subject matter knowledge to the evaluation contributing to the theory of change, important mechanisms and assumptions, effects and the counterfactual. For Marmot there were two Technical Advisors, a hydrologist and a fisheries ecologist to help us understand the complexity of effects in water (flow, seasonal flow, temperature, sedimentation) and of the targeted fish species (life cycle, habitat, mortality, contingencies with water and other factors). Technical Advisors can also represent the knowledge of important mechanisms, for example advisors to a recent evaluation of marine protected areas (MPAs) represented collaborative decision processes in marine settings and with First Nations, while on another evaluation that sought to take the intervention to scale diffusion of innovation was a relevant knowledge domain.

### Technical Advisor(s)

The evaluation requires a source of expert technical knowledge from the outset. The Technical Advisor(s) are identified and join the team prior to significant engagement with interests and parties. The Technical Advisor provides advice, identifies and helps interpret key technical documents to compliment the specialized evaluation expertise of the RIE evaluator.

They are most often academics or practitioners and cannot have a prior relationship to the intervention. On occasion government officials have been used so long as their independence is easily and commonly accepted, for example an ecologist from the Smithsonian or a Geological Survey scientist.

The Technical Advisors interpret science and technical documents and data where needed and provide inputs and advice in drafting the program summary document. They may also contribute to interpreting the results, for example calculating the change in greenhouse gas emission attributable to the program. For Marmot Dam the advisors helped the evaluators interpret planning science studies. If this evaluation were conducted now the advisors would also have helped extract and interpret coefficients from simulation models for the relevant fish populations enabling estimates of the change in stocks attributable to the removal of the dam.

A subject matter expert group<sup>4</sup> is involved in the impact assessments and are usually paid an honorarium and expenses. A RIE that addresses only a modest range of evaluation issues additional to impacts will require about 30 days plus travel from the evaluator and five days from the Technical Advisor; the subject matter panel is usually budgeted at \$5,000 USD. Assistance from a mid-level evaluator is helpful to review literature, coordinate interviews and the panel, and administer surveys. Application of RIE in international development settings encounter logistical issues that contribute to higher costs and longer durations for a RIE. Likewise multi-site interventions with important differences

---

<sup>4</sup> Described in Expert Groups text box below.

between sites can require additional design processes and a subject matter panel for each site, such as Vietnam and Indonesia for the GEF UNIDO energy efficiency program or the Maritimes and Pacific jurisdictions in the Canadian MPA evaluation.

### *Design of a RIE evaluation*

Following reviews of secondary and program documents and inputs from the Technical Advisor(s) convening interests and representatives of selected core interests are interviewed. A goal of this preliminary review is to sketch the main effects and logic pursued by the intervention and the key mechanisms of change that will give life to these. The evaluation design which is captured in a short summary of the intervention and its' key elements including: description of the program, effects and the logic of the program, the counterfactual, temporal and spatial scales, and a listing of the interests who can influence or are affected by the program.

### **Effects, outcomes and impacts**

The reach of a RIE to longer term outcomes and impacts is contingent on what the expert groups can reasonably assess. We recognize variability and ambiguity through use of the term *effects* that includes outcomes (long and short term) and impacts that are articulated in the theory of change. Like many evaluations RIE can be confronted by highly contingent and complex outcomes and impacts. However RIE is an expert judgement approach and needs to stay within the reach of what can be reasonably addressed by the approach.

Moreover as a use-seeking approach the reach of the evaluation is largely determined through the joint knowledge process during the design phase. Typically some interests are very ambitious about the reach and strength of the intervention while other interests can be somewhat churlish. Anecdotally it seems that distinguishing short and long term outcomes and impacts is often regarded as a hierarchy associated with the status or value of the intervention and making it more difficult to reach agreement on what should be attributed to the intervention. Focusing on direct and more influential contributions to the program goals has proven useful as has using the generic term *effects* rather than outcomes and impacts. Use of the general term effects is a strategic decision. It is useful to note that the effects addressed by the evaluation are those that apply to the intervention and these same effects are used in assessing the counterfactual.

Identification of effects starts with the literature and program documents, then discussions with representatives of convening and key core interests, and inputs from the Technical Advisors. Effects are part of the initial program summary that is then reviewed and potentially adapted through the joint knowledge process with representatives of all convening and core interests. On occasion the subject matter expert panel raises concerns about how an effect is being operationalized and the evaluator may adapt the expression of individual effects.

### **Scenario-Based Counterfactuals**

Many of the constraints on impact evaluation approaches arise from challenges identifying and operationalizing the counterfactual. Some of these are; uneven temporal occurrence of impacts; technically feasible options such as with/without the intervention are often not regarded as efficacious,



legal or ethical (Boyd & Mason, 2011); complex, ambiguous and highly contingent causality often have too many moving parts and lack sufficient data; impacts can occur over widely differing spatial scales which certainly for natural systems do not align with the spatial scales of the program; and the program definition and accountability frames can disconnect them from many of the important resulting impacts (Rowe, 2018).

The Scenario-based counterfactual is efficacious (to decision makers and stakeholders); a plausible option considering politics, culture, capacity; it is feasible in the sense that there are no budgetary, timing, or technical reasons that would have prevented use; it is legal in the sense that the scenario represents an option within current law or regulations or with plausible changes to these and it is ethical. Often the scenario-based counterfactual is an approach that was strongly considered but not applied or an approach that is being applied elsewhere. Sometimes a with/without the intervention comparison is used as a scenario-based counterfactual so long as it is efficacious, plausible, feasible, ethical and legal option on which the program experts agree.

Scenario-based counterfactuals are developed jointly by the evaluator, technical advisors and program experts as part of the program summary in the first phase of a RIE. The scenario-based counterfactual specifies the alternative processes and any differences that would have occurred such as in timing or scale. The Scenario-based counterfactual to evaluate decommissioning the Marmot dam was:

*If PGE had developed the dam removal plan on their own and submitted their plan to the Federal Energy Regulatory Commission in 2001, environmental groups and possibly the Tribes would have litigated using the Endangered Species Act and Tribal legislation. Litigation would have delayed the initial decision by 3 years to 2004 and with additional technical studies required by the court, technical planning would have concluded in 2007. The judicial decision would have supported PGE's application to decommission the dam and PGE would have breached the dam using explosives much as they did. The judicial decision would not have included transfer of senior instream water rights to the state, deeding of 1,500 acres of shore lands to the Western Rivers Conservancy (providing the formative nucleus for the 9,000 acre Bureau of Land Management managed natural refuge and recreation area), nor would the decision have provided funding for the monitoring program and research on effects of dam removal. The effects of this decision occur in the watershed immediately below the dam and extending to its intersection with the Bull Run River. All*

#### Scenario-based counterfactuals

The RIE counterfactual is initially developed by the evaluator from secondary documents and discussions with selected convening and core parties. The counterfactual is part of the program summary document that is reviewed by the interests and revised as necessary. The counterfactual is an alternative scenario for the intervention and can range from a reduced or “no program” through to scenarios that enhance the program being evaluated.

Scenario-based counterfactuals are often based on a very viable option such as an option that was seriously considered but not selected during program design, or an option that was applied elsewhere. They must be efficacious, plausible, feasible, ethical and legal.

*existing human activities above and below the dam, other than those required to operate the dam and hydro generation facility, would remain the same.*

The counterfactual is developed jointly through discussion with representatives of convening and core organisations as part of the program summary and reviewed by all stakeholders and revised as needed. It represents a consensus of stakeholders involved in or affected by the decision about what a realistic alternative to the agreement-seeking process that was used to reach a decision on decommissioning the Marmot dam. The scenario-based counterfactual is used like other counterfactuals. The difference between assessments of the effects of the program and the alternative becomes the net incremental change (merit) for each outcome. A strong advantage of the scenario-based counterfactual is that they can apply in ex ante settings in developmental or formative evaluations.

Clearly the evaluator plays a strong role in developing the counterfactual and so it is possible that s/he could influence the assessment of incremental direct effects through the process of selecting the counterfactual. We have built in two controls on this: first the scenario-based counterfactual is reviewed by the program experts who must concur on the counterfactual as well as all other elements before they are used in the evaluation; second, where there are sufficient respondents in the program group we can use two scenario-based counterfactuals bracketing the actual decision and assigned randomly in the survey.

### **Program summary**

A draft program summary document of two to five pages is developed and reviewed with convening interests (e.g. program management, funders, delivery agencies) whose comments and suggestions are incorporated into the draft summary. The summary is revised and sent to representatives of all the convening and core interests involved in the program including interests affected by the program. Telephone interviews with each gain suggestions for improving the draft summary and identifies concerns that they feel need to be addressed in the evaluation. Where revisions potentially important to one of more interests are suggested the revised document is circulated to all convening and core interests for further comment and suggestions/objections. Normally only modest revisions are required to the initial draft summary. The process concludes when there is consensus on the summary which may then be sent to selected non-core interests followed by shorter interviews. This can be an extensive stakeholder engagement and design process seeking to develop a solid understanding of the program and establishing positive prospects for use of the evaluation. Efficient RIE information gathering and analysis processes offset the higher inception cost which typically requires about two-thirds of the evaluators total time.

The program summary should include:

- A short descriptive statement of the intervention, logic and mechanisms of change for the intervention including the primary and connected systems and with temporal and spatial scales
- The intended effects of the intervention including those that are within its' reach and contingent effects to which the intervention contributes. Effects should be stated as conceptually

observable outcomes including their temporal and spatial scales (e.g. spawning habit will be appropriate and sufficient for the salmon population)

- The interests that can affect success of the intervention and those affected by the intervention, with potential representatives of those interests and identification of convening, core and non-core status
- One or more scenario-based counterfactuals

The main ambitions for first phase are to engage parties in a joint knowledge process in developing a program summary identifying the key elements required for the evaluation that is agreeable to the evaluator and to all convening and core interests. The joint knowledge process is expected to promote participation in the next phase and use of the evaluation, the program summary provides the necessary technical inputs for the evaluation.

### *Implementing a RIE evaluation*

RIE triangulates across three distinct groups of experts to generate a range of the likely level of achievement of effects of the intervention. A facilitated workshop with the subject matter expert panel is the first leg of the triangulation. The program summary is sent to subject matter experts prior to the workshop which begins with an overview and discussion of the workshop approach and the program summary, subject matter experts then provide their assessments for each effect under the program and counterfactual using the RIE metrics on flip charts. Workshop participants are encouraged to voice questions and concerns or to simply muse aloud at the flipcharts all providing opportunities to facilitate discussion promoting a common understanding of the key concepts and assumptions. It is made clear that we are not seeking consensus on any of the questions and that the reason for encouraging dialogue is to promote a common understanding of the concepts that they are assessing thereby promoting reliability.

#### **Subject matter experts – Expert group 1**

This is a group of experts in the various knowledge domains important to the intervention and setting. Typically a subject-matter expert group has about five members (see Marmot Dam description above). Subject-matter experts should not have any relationship to the intervention, or the specific issue being addressed. They are convened in a one-day workshop, first reviewing the theory of change and the RIE metrics. The subject-matter experts then participate in a facilitated process to generate their assessments.

The results are compiled during the final coffee break and discussed when the group reconvenes with the intent of gaining their insights into the causality and limitations of the assessments they have provided and extending the joint knowledge process to interpretation.

The main function of the workshop with subject matter experts is to gain their assessments of the effects of the intervention using the RIE metrics. Their consideration and discussion early in the workshop of the intervention including the mechanisms of change and logic, effects and the operationalization of the impact metrics provides valuable insights and information about the underlying sciences as they apply to the intervention. On occasion the subject matter expert panel has

suggested modest adaptation to how an effect is operationalized in a question. The evaluator decides on whether to revise the question but must maintain fidelity to the concepts developed during the design phase. This means that the elements of the evaluation first developed by the evaluator and Technical Advisor, then considered and adapted by program experts are now considered from a more technical perspective by experts in the relevant subject matters.

The second triangulation leg is the program expert group using a web-based survey with representatives of all convening, all core and all or some non-core interests. This survey also provides a vehicle to gain inputs and information about other evaluation issues and to obtain valuable qualitative inputs.

The population size of the program expert group has ranged from under a dozen on high stakes environmental enforcement decisions such as with USEPA to many hundreds in a more programmatic setting such as the GEF evaluation of the UNIDO SE Asia energy efficiency program. By ensuring that all interests provide inputs the bias coming from the worldview and priorities of any individual interest are balanced by those of the other interests. It is not unusual for evaluations to prioritise program interests lending the evaluation a program-centric bias.

**Program experts – Expert group 2**

Representatives of interests and parties that can affect success of the intervention or who are affected by the intervention have strong knowledge of the intervention, its' setting and effects. They are thus experts in the program. It is important that all interests are represented to ensure that the evaluation adequately expresses the intervention, that the evaluation questions and approach are salient to all interests not just program officials. Program experts are involved throughout the evaluation, design, information gathering and interpretation. This approach promotes support for the evaluation and findings from relevant interests.

DRAFT - 10

The third and final triangulation leg is a web-based survey of the technical advisors addressing the same RIE metric questions as the program expert group. The technical advisors straddle the space between program experts (experts in the program and knowing a little of the sciences) and subject matter experts who know little about the program and a lot about the science. This completes the triangulation with the subject matter experts providing expert knowledge of the key science domains, the program experts providing expert knowledge of the intervention; the technical advisors drawn from one or more of the key science domains and gaining knowledge of the intervention through their engagement with the evaluation occupy a knowledge position bridging both the sciences and the intervention.

### Technical advisors – Expert group 3

The technical advisors, by virtue of their subject matter expertise and knowledge of the intervention gained through their advisory roles, represent a third source of expertise, bridging the other two groups. It is not unusual for readers to confuse the subject-matter expert group and technical advisors; the former are engaged in a facilitated process and have no prior engagement with the intervention or issue, the technical advisors will likely share subject matter expertise with some of these and have gained good knowledge of the intervention and its theory of change through their advisory function.

### Impact Evaluation Metrics

Four questions apply the RIE impact metrics to each effect addressing the likelihood and magnitude of the effects under the intervention and under the counterfactual, that is:

1. Likelihood under the intervention
2. Magnitude under the intervention
3. Likelihood under the counterfactual
4. Magnitude under the counterfactual.

This is illustrated using questions from a recent evaluation that applied the current RIE methods, evaluation of Marine Protected Areas in Canada. At the time of the evaluation Canada was lagging significantly behind international commitments (Internal Audit Fisheries and Oceans Canada, 2016) and the then newly elected government prioritised meeting international commitments in the overarching context of Truth and Reconciliation with Indigenous nations and co-management of oceans and fisheries (Trudeau, 2016). This placed the focus on accelerating the processes to designate new MPAs, a formative evaluation application of RIE. The mandate, research literature and program documents indicated a more inclusive, transparent and collaborative decision processes would improve the designation processes, operationalized for the evaluation as a shift from the then current setting where interests outside the federal government provided input to the designation decisions to an agreement-seeking process (International Association for Public Participation, 2018) as described in the scenario-based counterfactual:

*...for the purposes of this evaluation, please consider an alternative scenario where DFO (Fisheries and Oceans Canada) and interested/affected parties engage in processes to **agree** on the substance of the declaration of regulatory intent including recommendations for the*

*conservation objectives, the regulatory and management measures and the boundaries of the proposed MPA. The process includes all interests that will eventually be reflected in the regulatory impact analysis by Treasury Board and other Government of Canada agencies. Parties engaging in this process would ensure that their organisations are briefed on the discussions and raise issues that could affect their ability or willingness to support the agreement. Current processes including interim protection where needed and continue subsequent to development of the regulatory intent; DFO prepares key documents that are reviewed by DFO, Treasury Board and the Department of Justice. Department of Justice drafts the MPA regulations based on the statement of regulatory intent and the reviews of key documents developed by DFO. If the responses received are substantively different than the agreed to declaration of regulatory intent and conservation objective recommendations, DFO and interested/affected parties will engage again to agree on any modifications to the regulatory impact analysis statement and the draft regulations before the final regulations are published in Canada Gazette Part II.*

From the literature and consultations one of thirteen key results from an agreement seeking process was that all parties would effectively engage in the processes; for the RIE engagement translated to the likelihood concept that all convening and core interests would attend and a magnitude concept that they would actively engage in the process. The questions for the current program were phrased as follows:

Preamble

The questions ask you to assess the likely level of achievement of six factors known to be important in decision making. Questions are worded to direct you to consider each factor first from the perspective of likelihood of this factor occurring and then the magnitude or strength of the factor.

Stem

Please consider a situation where you might be involved with a process to designate a new MPA that is similar to the processes used for the {{Q2}} MPA in which you participated earlier. These processes address identification of the boundaries, conservation values, permitted uses and the monitoring and management plans for the new MPA. To what extent would you agree with the following statements? If you wish to elaborate, please use the comments box.

Questions

Q1 All parties who can affect the designation of the new MPA *will attend* the processes.

Q2 All parties who can affect the designation of the new MPA *will actively participate* in the processes.

The following scales are used for assessing likelihood and magnitude:

1. Likelihood uses a scale from 0 (no chance) to 3 (certain / has already occurred).
2. Magnitude uses a scale from -3 to +3 with 0 as a mid-point labeled “no effect”. The end points are labeled to indicate as “bad” and as “good” as it can get within the context of the program. Where negative effects are not realistic a 0 to 3 scale is used.

Likelihood and magnitude are combined to produce an index of change in each direct effect. We apply these measures and calculate the difference between the program and the scenario-based

counterfactual much as one would with a comparison group. This is a measure of evaluation merit, the net incremental change attributable to the program.

### Computations

The main steps in calculating impacts using these metrics are first to calculate the RIE indices, then normalize the program expert group responses to equalise the strength of voice of each interest, and finally to compute the effects:

For each expert group

1. Calculate the RIE index for each respondent and for each effect under the intervention and the counterfactual. The RIE index is the product of the scores for likelihood and magnitude and then divided by the maximum possible score thereby generating an index with values between 0 and 1. If likelihood was 3 (certain) and magnitude was 1 (some effect) the calculation would be  $(3*1)/9$  where 9 is the maximum score possible on the 3-point scales. If either likelihood or magnitude is zero, the value of the index is zero.

For the program expert group, because it is likely that some interests will be represented by more than one organization and that organisations will likely have different numbers of respondents:

2. Where an organization has more than one respondent calculate the mean RIE index values for the organization so that each organization has a single index value (e.g. Oregon Fish and Wildlife).
3. Where an interest is represented by more than one organization calculate the mean RIE index value for the interest so that each interest has a single index value (e.g. state government).

For each expert group

4. Calculate the contribution of the program to the effects (RIE contribution index) as the difference between the RIE index for the program and counterfactual.

#### Calculating the RIE index for each of the expert groups

**Subject-matter experts:** The RIE index is calculated for each expert and the results are presented as a range (e.g. subject-matter experts assessed impact A at between 12-18% improvement. A mode is used for graphical representations of their assessments.

**Program experts:** The intent is to provide an assessment that balances the perspectives and assessments of all affecting and effected interests. The process is to calculate the RIE index for each participant, average these by parties where there are multiple representatives, then by interest and then a single average is generated for all interests. This process ensures that interests and parties with more representatives do not receive more weight, and that the views of all interests are considered equally important.

## **Presentation of Results**

The assessment of effects is presented by the RIE index which readily translates to a percentage change (e.g. a RIE index of 0.25 is a 25% change in that effect) in an effect attributable to the intervention. The results are presented as the average for each of the three expert groups and discussed as a range. Results are also disaggregated by interest and other sources of variation such as region for the program experts to provide additional insights and as a base for further analysis. To illustrate, a recent evaluation of the current approaches to designating Marine Protected Areas in Canada posed a counterfactual that would put in place the then recently elected Prime Ministers' direction to the Minister of Fisheries and Oceans to address reconciliation with First Nations and seek co-management approaches. The counterfactual on which all interests agreed was an agreement seeking process compared to the current consultative process see (International Association for Public Participation, 2018) which the three groups of experts expect to improve results by 25-35%. Results included the time required to designate a new MPA, and the implementability and durability of agreements on key elements including the definition and management plan. The counterfactual represented a feasible, plausible, ethical and legal approach to designation that would fulfill the direction from the Prime Minister. The more detailed results around engagement (of interests) and the character of the designation process and provided guidance on the areas that were most determinative of the lower assessments and poor performance of the current process.

## **Addressing Challenging Effects**

As Stern et al (2012, section 5) comment evaluating program effects can be a complex undertaking because the program itself is complex or because the program is set in a complex system(s) that influences results. The Marmot example illustrates this, the effect of interest is the population of salmon in the river systems above the dam. Salmon are born in the rivers and return to the river three to five years later to reproduce. While at sea the salmon from our river are well beyond the reach<sup>5</sup> of the intervention and are there affected by many diverse forces such as ocean acidification and warming, legal and illegal harvests, ocean garbage (plastics, drift nets), and declining food stocks, none of which are affected by removal of the Marmot dam. To get to the sea as youngsters, and on their return from the sea as adults, salmon face a number of hurdles many of which are affected by the removal of the dam as well as by other factors: water flow and temperature especially during the sensitive migration periods, barriers to passage such as the Marmot dam and other barriers including dams downstream of the Marmot, quality of spawning habitat, availability and suitability of protected and shaded resting places, riverside human activities, and so on. Natural science turns to methods such as modelling to estimate the contribution of removing the dam to departing or returning salmon populations. These models can be very useful, but like traditional impact evaluation they are costly, time-consuming and data and capacity challenged.

While the expert judgements might sometimes or often not fully reach impacts, RIE uses an approach not dissimilar to other established methodologies such as Bayesian Networks and Structured Analogies and takes estimation further than related theory-based evaluation approaches such as Contribution

---

<sup>5</sup> Reach has both spatial and temporal dimensions. An important temporal factor is that the intervention removes the dam at a single temporal point but each affects each successive generation of salmon potentially in perpetuity.



Analysis. (Mayne, 2001). With the knowledge and guidance of the Technical Advisors and other consulted experts and the technical literature RIE can reach to or close to impacts. For example, simulation models and the research literature enable estimation of the effects of removal of a dam on salmon stocks based on combining the known determining outcomes that were assessed by RIE expert groups (habitat etc.); similarly Technical Advisors and technical studies and consultation with a panel of global experts in energy efficiency enable extension from RIE-estimated changes in enterprise use of energy efficient processes and equipment to changes in greenhouse gas emissions and from there to human health<sup>6</sup>.

### **Validity and Reliability**

RIE goes to considerable length to specify important elements such as the effects, counterfactuals and temporal and spatial scales. The intent is that each expert will understand these concepts in the same way and as a result RIE assessments will be reliable. Results have been encouraging, the Cronbach's Alpha statistic assessing internal consistency has always been above the threshold of 0.7 indicating relatively high levels of internal consistency. The Cronbach's Alpha statistics for the Oregon pilots all exceeded 0.9; in subsequent applications of RIE the lowest value for the statistic has been 0.84. Similar results have been achieved when RIE methods are used to assess impacts as part of mixed methods evaluations.

The internal reliability of RIE assessments is not surprising with an extensive consensus-seeking joint knowledge process. This process also contributes to validity. All of the concepts important to the intervention construct (program logic, mechanisms of change, effects, temporal and spatial scales) and to the evaluation (counterfactual, adjectives) are determined through a systematic process beginning with the evaluator and Technical Advisor drafting these elements from program documents and the literature and then a review and enhancement process with experts in the intervention (all convening and core interests) that only concludes when there is consensus on all elements, followed by review by the subject-matter (technical) experts. The ambition of RIE is to provide good quality assessments of the impacts from an intervention. A founding premise of RIE is that interventions often have unique qualities and settings and the unit of account for RIE is usually the specific intervention. It is thus not wise to generalize from a single RIE. Generalisation beyond the intervention would require more programmatic use of RIE to assess, for example, classes of interventions using a number of connected RIE evaluations of a representative sample of a program; RIE evaluations could also be incorporated into systematic reviews. However an individual RIE evaluation is specific to the intervention and RIE processes provide a solid and agreed foundation of the primary theoretical constructs and their operationalization and support valid inferences from the results to the intervention, that is, RIE has good construct validity.

Of course a main interest is how well do the judgments of the three groups of experts comport with actual observations? This is challenging, if there were data on the incremental impacts it is unlikely that an impact evaluation would be requested. However in two of the Oregon pilots we could compare the estimates for changes in fish populations from technical studies conducted for the intervention with the

---

<sup>6</sup> It is noteworthy that the estimates needed to be generated for each country, industry sector and enterprise size.

RIE forecasts, both showed the RIE method results to be compatible with the intervention forecasts (Rowe, Colby, Niemeyer, & Hall, 2004). The estimates from the GEF evaluation of the UNIDO energy efficiency program in SE Asia were reviewed positively by a global expert panel in energy efficiency. More often RIE evaluations have been conducted in highly contingent settings for which useful external estimates of effects are not available, for example rule-making for use of off-road vehicles on two National Seashores or decision processes used in USEPA enforcement actions at sites that are part of much larger, very polluted and highly connected sites.

From the testing undertaken RIE seems to generate assessments that can be considered reliable and valid. More importantly the use-seeking joint knowledge generation processes promote the credibility, legitimacy and salience of the evaluation.

## Summary

RIE is a systematic approach to triangulating expert judgments using three distinct groups of experts. The three expert groups use the same evaluation metrics for the effects of the program but approach the evaluation with different kinds and levels of knowledge and use different mechanisms to render their assessments.

Three methodological developments enable rapid evaluation of impacts, the scenario-based counterfactual, a simplified measurement approach for assessing impacts and an interest-based approach to engaging stakeholders in the evaluation. RIE is structured around a use-seeking framework and can provide valid and reliable assessments of impacts. The simplified metrics and the scenario-based counterfactual enable formative and developmental evaluations to assess likely future outcomes without using the full RIE approach. These can be incorporated into most mixed methods evaluations without significant cost and effort. These characteristics offer the possibility of extending impact evaluation to settings where current impacts approaches are not feasible or technically possible.

In applications to date RIE is a relatively low cost and efficient approach compared to other impact evaluation methods; it does not require much new empirical information and does not place any requirements on the design or implementation of the program. RIE offers an approach to evaluate impacts in both ex ante and ex post settings and so has utility for developmental, formative and summative evaluation.

## Works Cited

- Bingham, L., Emerson, K., Nabatchi, T., & O'Leary, R. (2003). The Challenges of Environmental Conflict Resolution. In E. Rosemary O'Leary and Lisa B. Bingham, *Promise and Performance of Environmental Conflict Resolution*. Resources for the Future.
- Blatterman, C. (2008, February 14). Impact Evaluation 2.0. *Presentation to the Department for International Development (DFID)*. London: unpublished presentation.
- Boyd, J., & Mason, C. (2011). *Attributing Benefits to Voluntary Programs in EPA's Office of Resource Conservation and Recovery*. Washington DC: Resources for the Future.
- Brogden, M. (2003). The Assessment of Environmental Outcomes. In E. Rosemary O'Leary and Lisa B. Bingham, *Promise and Performance of Environmental Conflict Resolution*. Resources for the Future.
- California Ocean Science Trust. (2013, December). *Putting the Pieces Together: Designing Expert Judgment Processes for Natural Resource Decision-Making*. Retrieved January 2014, from <http://calost.org/expert-judgment/>
- Carugi, C. (2016). Experiences with systematic triangulation at the Global Environment Facility. *Evaluation and Program Planning*, 55-66.
- Carugi, C., Rowe, A., & Anand, A. (in preparation). Lessons from applying Rapid Impact Evaluation in mixed methods evaluations.
- Clark, W. C., Mitchell, R. B., & Cash, D. W. (2006). Evaluating the Influence of Global Environmental Assessments. In R. B. Mitchell, W. C. Clark, D. W. Cash, & N. M. Dickson, *Global Environmental Assessments* (pp. 1-28). Cambridge: MIT Press.
- Fisher, R., Ury, W. L., & Patton, B. (2011). *Getting to Yes: Negotiating Agreement Without Giving In*. Penguin.
- GAO. (2009). *Program Evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions*. Washington, DC: United States Government Accountability Office.
- Green, K. C., & Armstrong, J. S. (2004). *Structured Analogies for Forecasting*. Victoria, Australia: Department of Econometrics and Business Monash University.
- Hearn, S. (2017). *Rapid Outcome Assessment*. Overseas Development Institute.
- Hovick, D. (2005). *The Hewlett Foundation's Conflict Resolution Program: Twenty Years of Field-Building 1984-2004*. Pao Alto: William and Flora Hewlett Foundation.

- International Association for Public Participation. (2018). *IAP2 Spectrum of Public Participation*. Retrieved from International Association for Public Participation: <https://www.iap2.org/page/pillars>
- Jacobs, J., Brasseur, G. P., Barron, E. J., Benedick, R., Chameides, W. L., Dietz, T., . . . Richels, R. (2007). *Analysis of Global Change Assessments: Lessons Learned*. National Academy of Sciences.
- Keller, T. M. (2009, September 1). DAM DECOMMISSIONING: Removing Marmot Dam. *Hydro Review*. Retrieved from <https://www.hydroworld.com/articles/hr/print/volume-28/issue-6/featured-articles/articles/dam-decommissioning.html>
- Koontz, T. M., & Thomas, C. M. (2006). What Do We Know and Need to Know about the Environmental Outcomes of Collaborative Management? *Public Administration Review*, 111-121.
- Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *16*(1), 1-24.
- O'Connor, J., Major, J., & Grant, G. (2008, March). *Down with the Dams: Unchaining U.S. Rivers*. Retrieved July 7, 2012, from Geotimes: Earth, Energy and Environment News: [http://www.geotimes.org/mar08/article.html?id=feature\\_dams.html](http://www.geotimes.org/mar08/article.html?id=feature_dams.html)
- OECD. (2015). *Scientific Advice for Policy Making: The Role and Responsibility of Expert Bodies and Individual Scientists*. Retrieved from OECD Science, Technology and Industry Policy Papers: <http://dx.doi.org/10.1787/5js3311jcpwb-en>
- Rogers, P., Petrosino, A., Huebner, T. A., & Hacsí, T. A. (2000). Program Theory Evaluation: Practice, Promise, and Problems. *New Directions in Evaluation*, 5-13.
- Rowe, A. (2004). Evaluation of Environmental Dispute Resolution Programs. In R. O'Leary, & L. Bingham, *Promise and Performance Of Environmental Conflict Resolution*. RFF Press.
- Rowe, A. (2012). Evaluation of Natural Resource Interventions. *American Journal of Evaluation*, 384-394.
- Rowe, A. (2018). Ecological Thinking as a Route to Sustainability-Ready Evaluation. In R. Hopson, & F. Cramm, *Solving Wicked Problems in Complex Evaluation Ecologies*. Stanford University Press.
- Rowe, A. (2019). Sustainability-Ready Evaluation: A call to action. *New Directions in Evaluation*.
- Rowe, A., & Lee, K. (2012). *Linking Knowledge with Action: Promoting Use of Science Knowledge*. Retrieved February 18, 2013, from Packard Foundation - Conservation and Science: [http://www.packard.org/wp-content/uploads/2013/05/LinkingKnowledgewithAction\\_ScienceCS2013.pdf](http://www.packard.org/wp-content/uploads/2013/05/LinkingKnowledgewithAction_ScienceCS2013.pdf)
- Rowe, A., Colby, B., Niemeyer, M., & Hall, W. (2004). *Evaluating Environmental and Economic Effects of Collaborative Decisions*. Hewlett Foundation.

- Server, B. (n.d.). *Bayesian networks - an introduction*. Retrieved June 05, 2019, from BayesServer.com: <https://www.bayesserver.com/docs/introduction/bayesian-networks>
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the Range of Designs and Methods for Impact Evaluation*. London: DFID.
- Susskind, L., & Weinstein, A. (1980, January 1). *Towards a Theory of Environmental Dispute Resolution*. Retrieved February 22, 2013, from Digital Commons at Boston College Law School: <http://lawdigitalcommons.bc.edu/cgi/viewcontent.cgi?article=1746&context=ealr>
- Todd, S. (2001). Measuring the effectiveness of environmental dispute settlement efforts. *Environmental Impact Assessment Review*, 97-110.
- White, H. (2006, March). Impact Evaluation: An Overview and Some Issues for Discussion. *DAC Network on Development Evaluation*.
- Wikipedia. (2012, November 19). *Bull Run Hydroelectric Project*. Retrieved February 15, 2013, from Wikipedia: [http://en.wikipedia.org/wiki/Bull\\_Run\\_Hydroelectric\\_Project](http://en.wikipedia.org/wiki/Bull_Run_Hydroelectric_Project)