Multi-level Evaluation Systems and Conflict Resolution in Environmental Settings

February 2010

Introduction

- First goal is to share what I know from working in North America, Europe, India, Western Pacific and Caribbean
- Second goal is to learn from your experiences in South Africa and elsewhere
- Leading characteristics of my approach to evaluation:
  - Mostly do formative or developmental evaluation,
  - I mostly use emergent designs,
  - I favor structured information gathering,
  - Programs see my role as an improvement ally,
  - Key decision makers learn of important observations and their implications, and my advice, during the evaluation. Reports synthesize what they have already heard.
- Speaking as individual, not representing any federal or state agencies or foundations, nor are the views I report necessarily theirs

Topics

- Will talk about the use of outcome mapping to articulate the first program theory for the field of environmental conflict resolution
- Demonstrate how that program theory was the foundation of the evaluation system that has now been applied by agencies for about eight years
- Describe a second approach, the SEEER method, used to evaluate the results of environmental decisions was first applied to collaborative decisions
- Describe use of the system and propose a framework, drawn from outside evaluation, which provides a useful framing for understanding use

Key Concepts

- Three types of evaluation:
  - **Summative** evaluation judges the merit or worth of a program. Decisions typically are about the continuation or replication of the program. Programs should already be known to be performing well, usually with assistance of formative evaluations.
  - **Formative** evaluation is about obtaining information, insights and providing advice to help programs improve. This is used in decisions to modify the program to improve its effectiveness.
  - **Developmental** evaluation helps programs navigate their way in very complex settings to identify and test approaches that will likely work.
- Program logic, logic models, and a theory of change are ways of capturing the problem the program is addressing and how it thinks it will succeed – i.e. what it needs to achieve to be successful.
  - Evaluation must be ethical and useful, feasible, and the quality of the information must be good enough for the decisions likely to be made. Evaluators aim to be able to judge what the program has contributed relative to a reasonable alternative, although we might not always address this specifically in any given evaluation undertaking.

Conflict Resolution

- Conflict Resolution variously referred to as Dispute Resolution, Environmental Conflict Resolution (ECR), Collaborative Decision Making, and Mediation
- Types of processes – agreement seeking, etc.
- State of the field in 2000
- My grand vision – with no evaluation or agreed program theory, formative evaluation approaches were needed to:
  - Facilitate articulation of a common approach
  - Provide information and insights that could be used to improve ECR processes
  - So that it could mature and be ready for evaluation that would look at the efficacy of ECR vs other approaches.

Evaluation of Conflict Resolution in the US

- Legal and conflict resolution academics have been describing individual cases for about three decades
- First systematic contribution was in the early 1990s by Paul Sabatier and Bill Leach looking at state water boards in California in the early 1990s
- Late 1990's a state organization (Policy Consensus Institute – PCI) and federal agency (US Institute for Environmental Conflict Resolution – USIECR) collaborated to develop evaluation for ECR
- First and still the only systematic evaluation of conflict resolution processes was the product of the PCI/USIECR partnership:
  - System developed by Andy Rowe collaboratively with agencies
  - Referred to as "…an excellent job of operationalizing the characteristics and properties of mediators".

Observations about Role and Use of Evaluation for State and Federal Agencies

- Process evaluation system:
  - Applied at most federal environmental agencies and several state conflict resolution agencies.
  - Subject of very favourable reflection and review resulting in peer reviewed publications.
  - The only evidence of use to improve the practice of ECR is by the Canadian provincial agency and Oregon state agency.
  - Extensive unintended use of the system for performance reporting.
  - Estimated cost to date is 1.5 million USD.
  - System has significant “neutral” bias.
- Results evaluation methods:
  - Provides statistically valid and reliable estimates of environmental effects in three applications to date
  - Captures economic effects

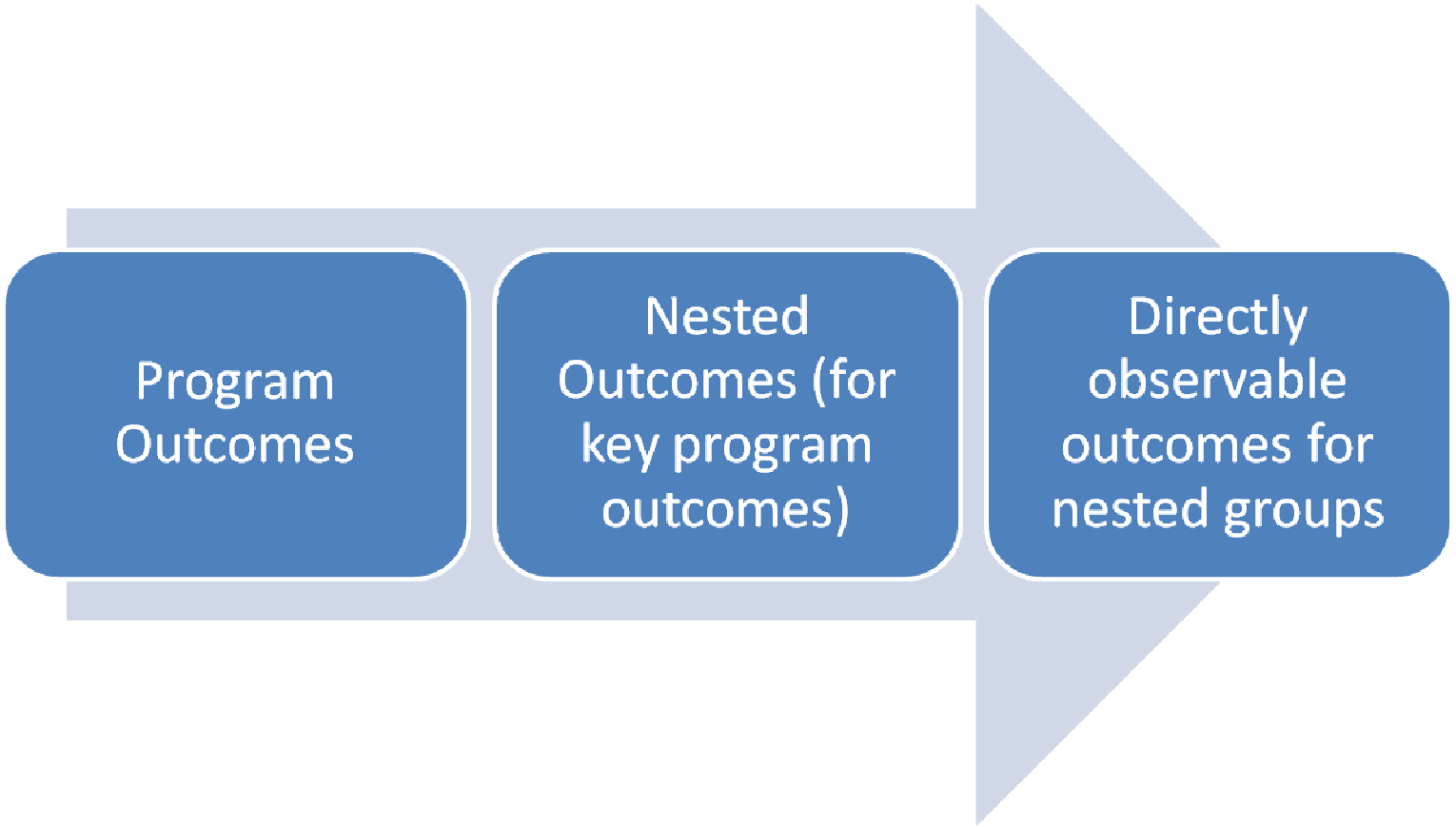
Examples of Cases in Process Evaluation

- This is a CERCLA cost recovery case filed in District of New Jersey under CERCLA Section 107(a). The US seeks reimbursement of site costs totaling $1.3 Million for removal and oversight activities related to the former * facility site. This matter involves contamination of soil at an industrial facility with long history of use by different operators.
- The Shoshone-Bannock Tribes (Tribes) of the Fort Hall Reservation in southeast Idaho submitted an application to the Environmental Protection Agency (EPA) on December 14, 2004 seeking eligibility for "Treatment in the Same Manner as a State" (TAS) for the purposes of administering water quality standards programs and issuing water quality certifications under the Clean Water Act (CWA).
- Mediation services were rendered in an effort to bring various stakeholders together on Vermont Wetlands regulatory issues with the ultimate goal of developing legislative reforms to wetlands regulation in Vermont. Previous efforts to pass meaningful reform legislation were frustrated by the differing views and positions of the various stakeholders (Environmental organizations, regulated entities and regulating agencies). Major obstacles to success included a long history of mistrust and inability to successfully negotiate and compromise among the parties.

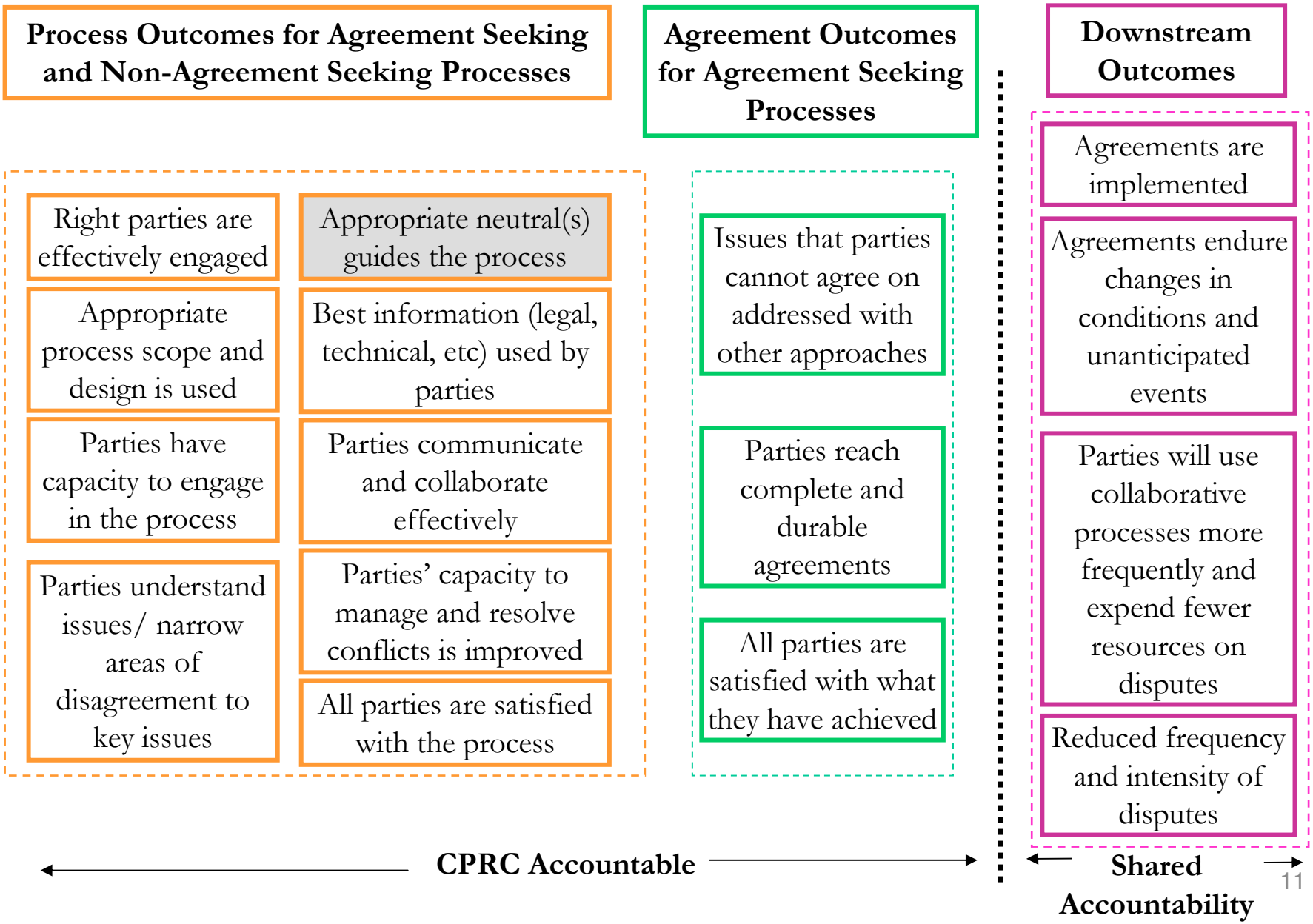
Step 1 Create a Program Logic

- Started with state agency using an outcome mapping approach to address the first two questions:
  1. What do we need to achieve?
  2. How will we recognize it?
  3. How are we doing now?
- Result was taken to federal agency who were initially resistant ("The federal world is much more complicated") then adopted the structure with minor revisions
- Then created measurement
  - Web survey to all parties in the process and mediators
  - Self administered and analyzed by agency
  - Periodic limited data sharing

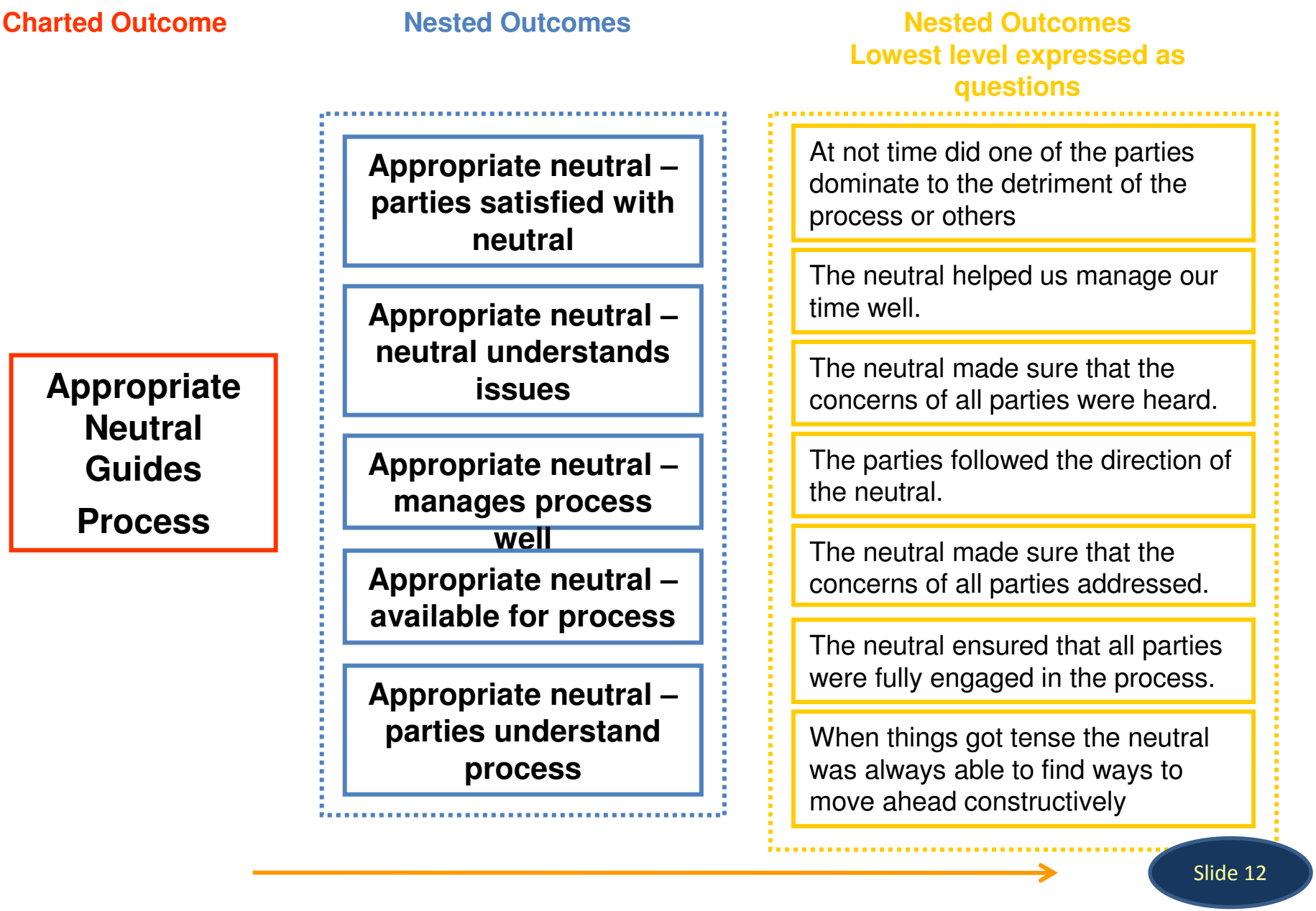
Outcomes Articulated at 3 Levels of Increasing Detail



Program Level Logic



Program Operational Logic



Quality of the Agreement

Agreement Outcomes for Agreement Seeking Processes

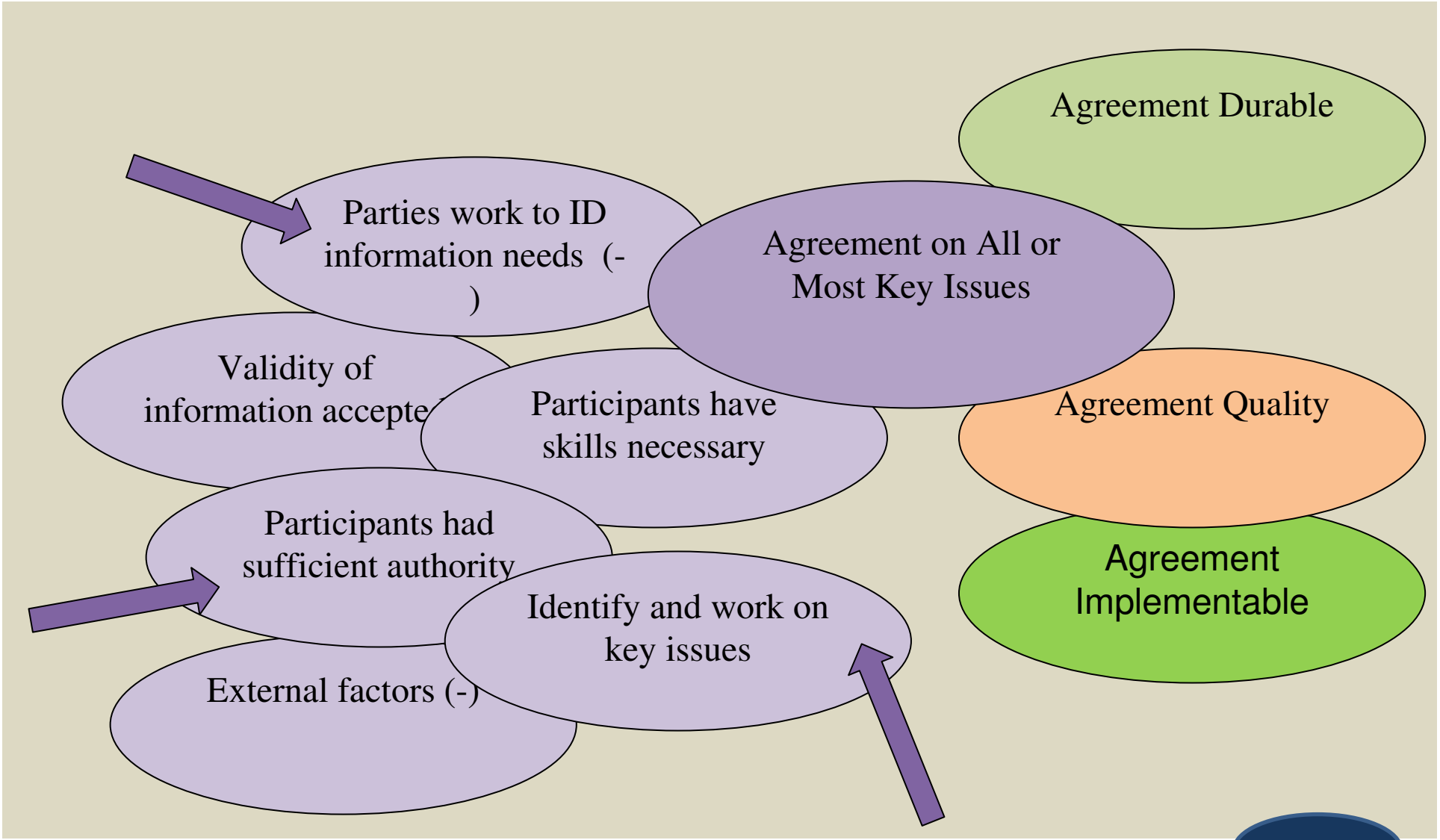
Issues that parties cannot agree on addressed with other approaches

Parties reach complete and durable agreements

All parties are satisfied with what they have achieved

Outcomes (0=not at all, 10=completely)	All CPRC (22)	Up stream (3)	Down stream (11)	Super fund (8)
Agreement is of high quality and the participants expect the agreement to last	7.1	7.6	6.7	7.2
Agreement has necessary elements for agreement	7.8	8.1	7.5	7.7
Agreement has necessary elements for agreement (neutral)	9.0	9.0	8.1	9.6
Agreement has necessary elements for implementation	5.9	6.6	5.1	6.3
Agreement is durable	6.2	6.7	5.9	6.1
Agreement meets more of parties' interests than without the collaborative process (neutral)	9.3	8.7	9.3	10.0

Example Advice to Agency - Levers to Improve Performance on Agreement Outcome



Used but not Useful

- The development, implementation and use of this evaluation system is fully and highly compliant with what we know about successful evaluation use.
  - Collaborative at all stages
  - Salient, credible, legitimate
  - Users have full ownership
- System is used aggressively and well regarded.
- Across all the agencies, there are two examples of use for improvement, termed instrumental use.
- System development and operations estimated to have cost USD 1.5 million.
- Conceptual use is high – first articulation of a program theory.
- Process use is high - same agencies who are not using process evaluation for improvement aggressively now use evaluation of other services and programs for improvement.
- Persuasive use is high – selective use to persuade external decision makers that ECR programs are effective.

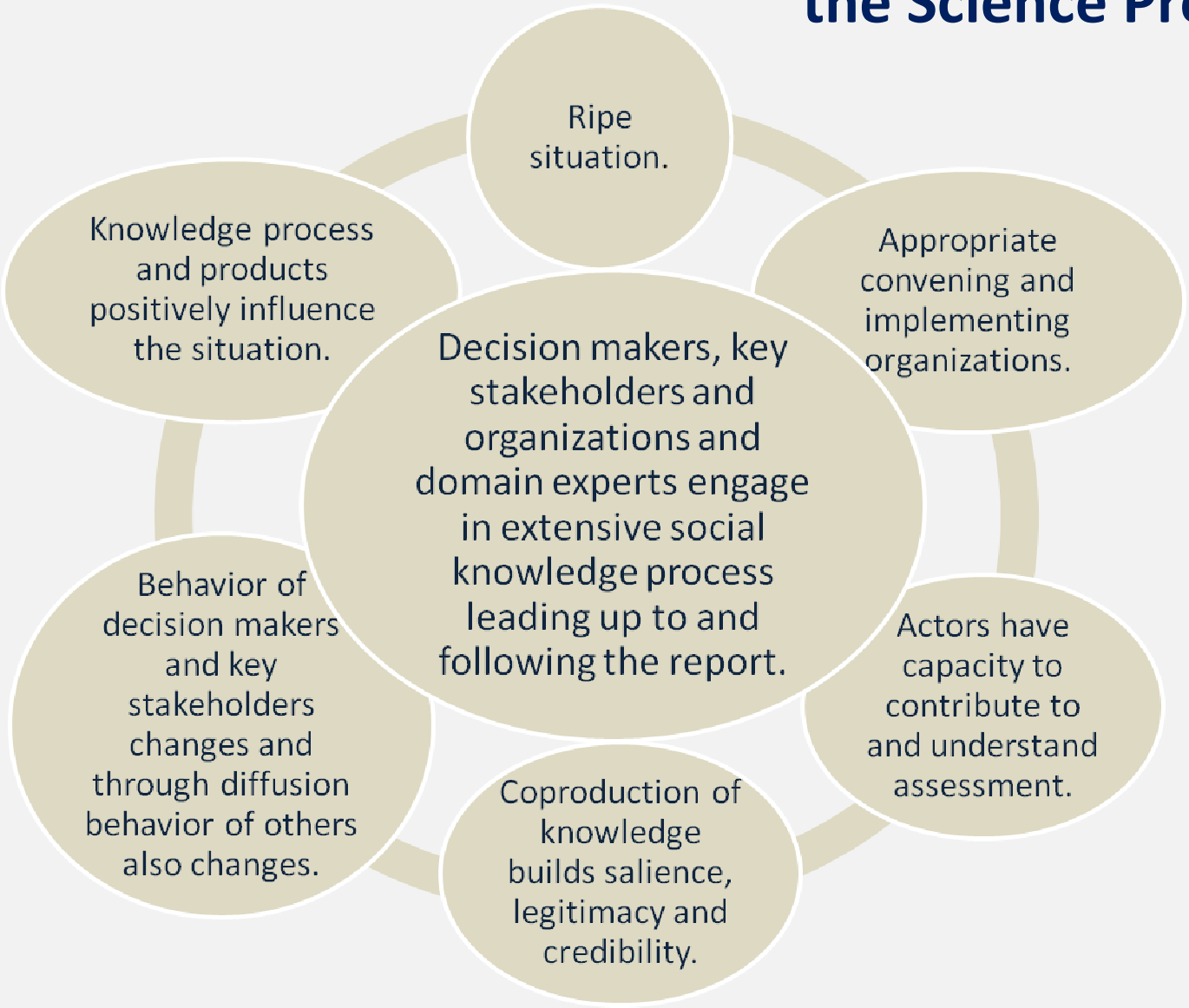
Accountability Trumps Improvement

- All of the agencies use information from the system to address their performance reporting requirements
- Revisions to the system have increased the amount of performance information
  - Revisions have not addressed needed improvements in program theory and related measurements
- Quality of performance information is very low
  - Example - parties compare collaborative process to another process they identify from a list
  - Parties from the same case rarely identify the same alternatives, often select contradictory and sometimes impossible alternatives
- Strong possibility that these performance questions will not be permitted in the next revision

Use at State and Federal Levels

- Two identified settings where the information has been used for improvement are state/provincial:
  - Oregon provided training for mediators focusing on areas of practice where improvement possibilities were identified.
  - Alberta has modified their program and provided advice to mediators.
- No identified use for improvement by federal agencies.

Draft LKwA Theory of Change for the Science Program



SEEER – Environmental and Economic Results

- SEEER was developed to compliment the process evaluation system by focusing directly on results attributable to the decision
  - Intent was to integrate the two systems once each working well
  - One of two lead federal agencies in the process evaluation is supporting SEEER, other lead agency regards SEEER as competition for its revisions to enhance accountability measures in the process evaluation

Marmot Dam Removal

- Proposal by the utility to decommission a hydro dam, or to license a new hydro dam, attracting or requiring participation from:
  - Federal, state and local government agencies who have a legal or regulatory interest,
  - Tribal governments who have legal and ancestral rights,
  - Environmental groups concerned about endangered species and other environmental issues likely to be affected,
  - Individuals and enterprises concerned with recreational uses (angling, white water boating),
  - Resource sectors and enterprises (agriculture, forestry) that might be affected.
- Utility has option of applying for a license directly to the Federal Energy Regulatory Commission and defending the likely litigation and administrative requirements that will follow, or engaging in a facilitated settlement process to address issues prior to application.

GE Pittsfield Superfund Site

- General Electric plant in Western Massachusetts had been manufacturing at the site since 1903 and produced insulators using PCBs from 1932 – 1977.
  - Waste was dumped on the site and in the Housatonic River that flowed through three states en route to the Atlantic. Soil from the dumpsites was donated by GE for community uses such as to build playing fields and for residential landscaping.
  - GE strongly resisted, spending as much as a half billion on external attorneys and science before agreeing to mediation,
  - Federal interests were represented at the highest level, highly unusual and likely due to the strong and visible participation of two Massachusetts federal senators, Kennedy and Kerry, the governor of Massachusetts and senior political and agency officials from the two other affected states.

Off Road Vehicle Use in National Seashores

Fire Island National Seashore	 Cape Cod National Seashore
Tension between vehicle use and ESA	Tension between vehicle use and ESA
10 – 15 nesting pairs of Piping Plovers	80 – 90 pairs of Piping Plovers
Dynamic dunes and shoreline	Dynamic dunes and shoreline
Everyday driving <u>needs</u> for residents, visitors and businesses	All driving <u>demand</u> is recreational
18 well established seasonal and year round communities in place when park created	Park adjoins communities with traditional use; Park lands sparsely and seasonally populated

Even Apparently Similar Cases Have Important Differences

What Makes SEEER Significant?

- First known systematic assessment of environmental effects of decisions from ECR
  - Most previous work focused on process elements and fidelity to best practice
- Important advances in methods:
  - Concept of negotiated alternative for use as comparison case in evaluations
  - Metrics for measurement of environmental effects
  - Triangulated assessments
  - Demonstrates participants can make valid and reliable judgments of effects under appropriate conditions

SEEER – Findings to Date

- Oregon cases selected to test methods and likely biased in favor of collaborative processes
- Across all cases, better environmental effects compared to the decision reached using the alternative decision process
  - Range from 10% to 25% better by case and setting
  - For some cases environmental gains are equal or favor the comparison decision process
  - Economic value of key effects in two fish cases estimated at USD 1.4 million
- US EPA cases required fewer person years to reach the decision compared to alternative (0.5 – 1.3, one case required more PY).
- Significant gains in social capital for some organizations
- Agreements are more adaptable to unforeseen circumstances

SEEER – Moving from Cases to Programs

- SEEER is well regarded in the ECR field, by managers of conflict resolution programs in federal environmental agencies, and by the mediators, program staff and parties who have participated.
- SEEER can be used to assess most site related environmental decisions, not just collaborative decisions
- Two main environmental agencies, US EPA and Interior want to use SEEER for programs
- Requires permission to survey more than nine non-federal employees from the Office of Management and Budget (OMB)
- OMB requires experimental designs which are not feasible and potentially not legal in these settings
  - Negotiations continue

Summary Observations

- Instrumental use is often very difficult
  - Federal agencies report they do not know what to bring changes
  - State and Provincial agencies use the advice and information
- Suggest to me that the federal agencies are not the right convening organizations
  - Do not have clear authority, a paradox since they have contracting leverage
  - Right convening organization might be consortium of agencies, mediators and party representatives – aligns with coproduction of all who can say “no”
- Clear warning that under current conditions there is a danger of improvement being forced out by accountability
- With experimental design bias this means that programs will not be as effective and will not show as well as they should if evaluated summatively