



# Evaluation Use and its Implications for Evaluation in South Africa

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 use of evaluation from a practical viewpoint, and the lessons from reflecting on my recent work,
- Propose a way of thinking about use that I find more succinct and salient than the very extensive discussion in the evaluation literature,
- Suggest some tendencies and habits in evaluation that undermine use.



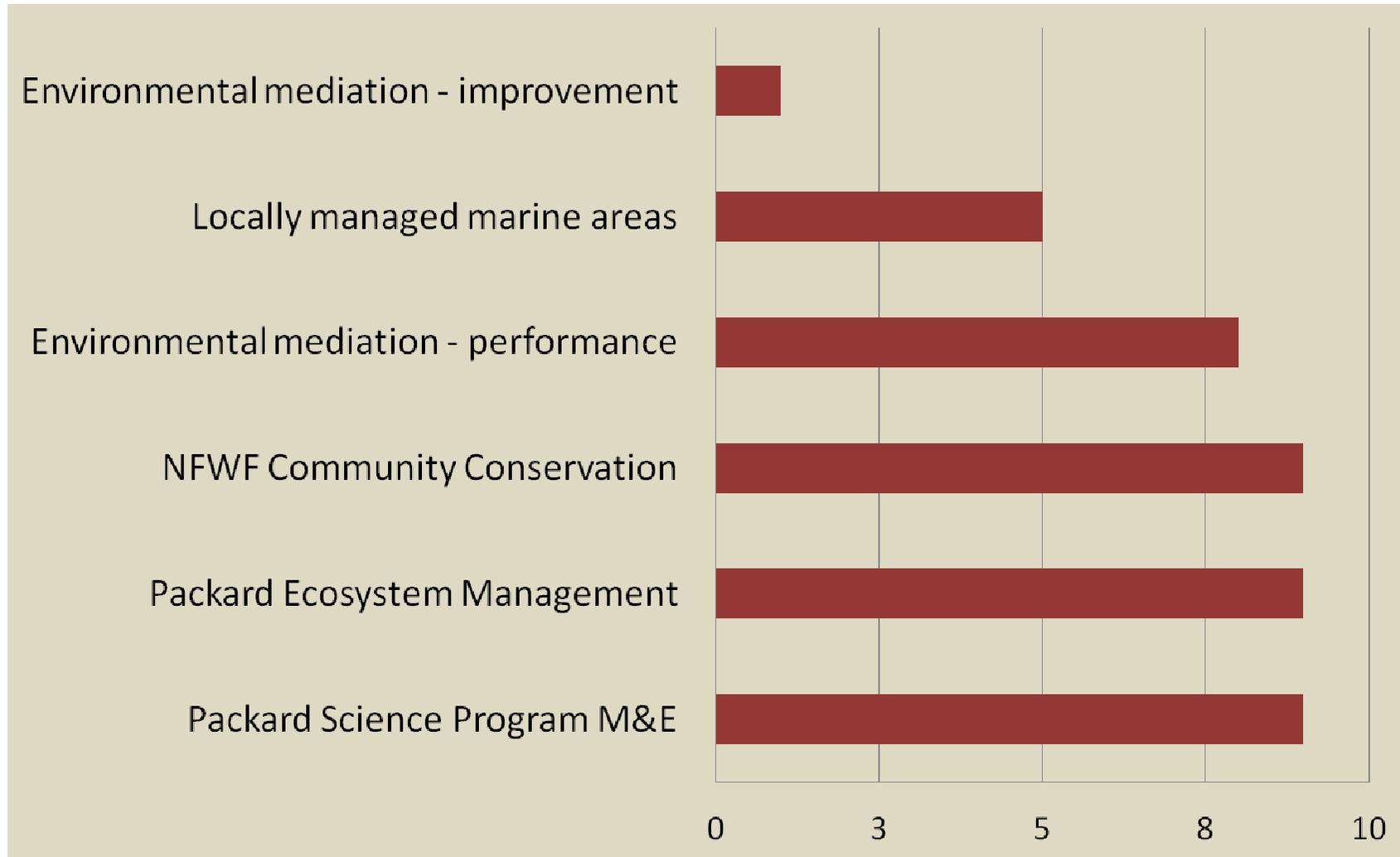
# Key Concepts

- Three types of evaluation:
  - **Summative** evaluation judge 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

# Background on My Reflective Review

- Over the past year I have been talking about use to programs with whom I have worked as an evaluator.
  - Review included a semi-structured initial interview followed by follow up discussions.
- Three types of evaluations: formative evaluations, formative evaluation systems and developmental evaluation.
  - Evaluation system is self administered and was the first rigorous evaluation of environmental conflict resolution.
- Reflection Process has been collaborative:
  - In 2008, federal and state programs using the evaluation formed a session to discuss use at the biannual Environmental Conflict Resolution Conference.
  - In 2009, representatives from one of the programs and I discussed use of their evaluation at a session at the annual Environmental Evaluators Network conference.

# Indicative Level of Instrumental Use of Evaluation Advice



Scale 0 (not at all) to 10 (fully)

# Example of Evaluation Advice and Program Response – NFWF Chesapeake Bay

## Advice

- Convert from community based in name to community based in reality
- Program acquire capacity in working with communities
- Provide TA to grantees on monitoring and collaborating with communities and use for post project assessments
- Ensure maintenance of actions
- Site visits to projects
- Change size of grants
- Simplify and improve admin, and provide better direction to grantees for reporting

## Response

- Site visit Fridays
- TA contracts
- More credible community efforts
- Grant program modified, higher ceiling and new smaller entry / planning grants
- Administrative changes adopted Foundation wide, not attributable to the evaluation

# Example of Evaluation Advice and Program Response – EBM Evaluation

## Advice

- Revise priorities and approach
- Support use driven science
- Projects need to change priorities, accelerate work with human systems
- Program and projects need more plausible logic
- Continue and enhance
  - Communities of Practice
  - Relationship grantmaking
  - Collaborations amongst grantees

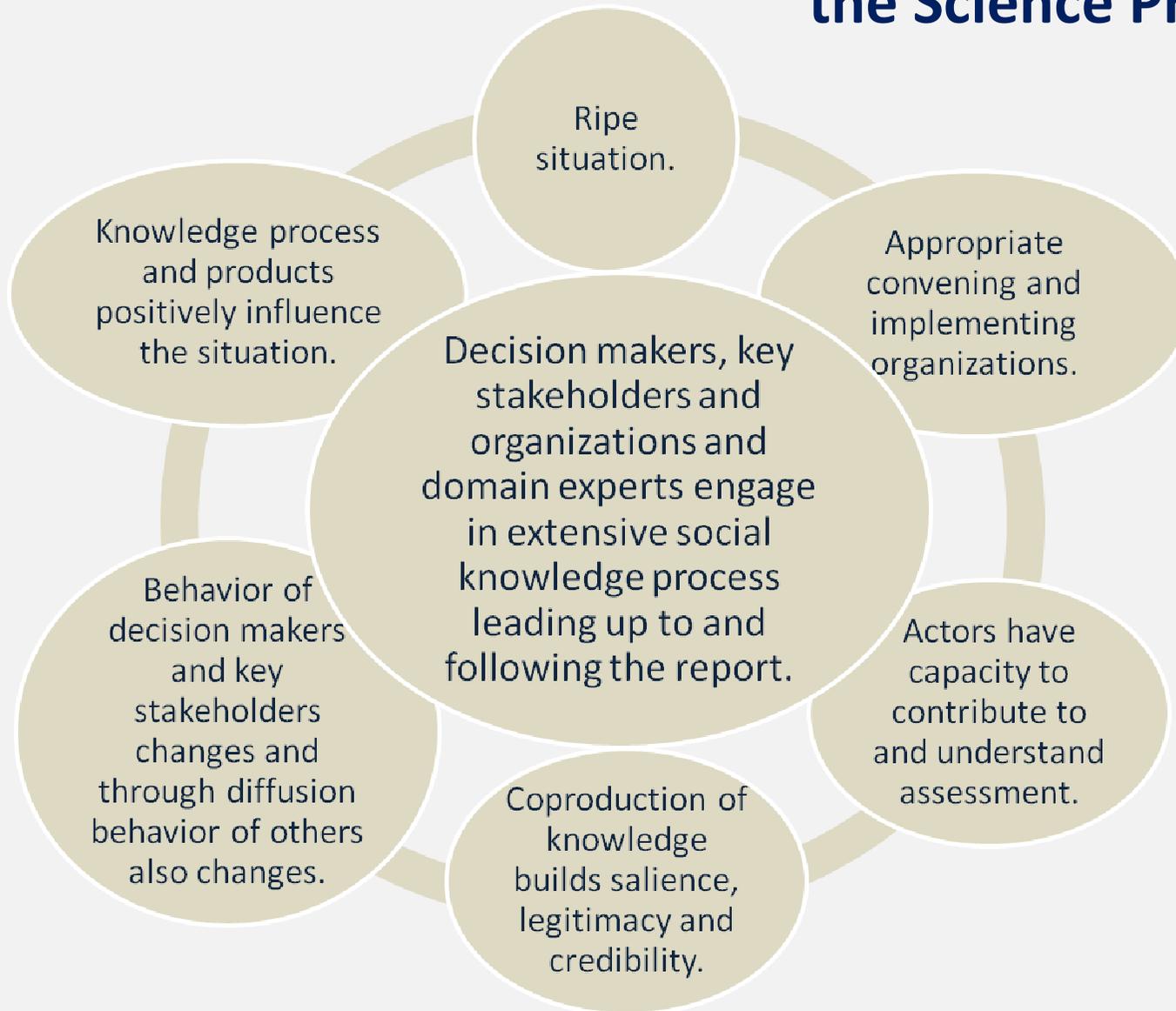
## Response

- Several were implemented before the evaluation was completed (e.g. program logic, use driven science, revising approach)
- All advice was accepted with pushback on some aspects
- Program was cut because of significant decline in endowment and absence of political capital for EBM in the organization

# Observations

- All of the projects were undertaken as collaborative / participatory approaches
- A key enabling factor was the ripeness of the setting for change
- Both evaluations were collaborative processes where the evaluator maintains control and responsibility for information and advice but the program participates in understanding key observations and framing options for improving
- For both, the report was not central to use
- For NFWF, the program took the lead in the strategy to gain buy-in from managers and Trustees

# Draft LKwA Theory of Change for the Science Program



# LKwA Fits Evaluation Use Examples

- Ripe Situation
  - Openings for change were created by changes in staffing and management, additional funding
  - Implementation of agreed changes
- Capacity
  - Programs and other decision makers were able to engage with the evaluators in the evaluation process, contribute value to exploring changes
- Coproduction
  - Engagement of programs contributed to salience and legitimacy
- Behavior Change and Diffusion to Grantees
  - By program using transparent collaborative approaches

# Evaluation System for Environmental Conflict Resolution Processes

## System

- First articulated a logic for the practice of ECR
- Then expressed the logic as observable outcomes
- Information from parties and mediator for each case
- Intent was to provide information to improve ECR practice
- Implemented by three leading US federal environmental agencies and several state conflict resolution agencies starting in 2002
- Still used and only source of systematic information about ECR processes
- Also now used by province of Alberta in Canada and the World Bank

## Use

- System implemented = use
- Substantial statistical work on data confirmed logic and utility of the methods
- No evidence that federal agencies have used information from the system for improvement
- However, it is a major source of performance information for the programs
  - System does not include program performance measures

## Used but not Useful

- Across all the agencies there are only two examples of use for improvement, both at state / provincial level
- System development and operations estimated to have cost well over USD 1.5 million.
- System is used aggressively and well regarded
- 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
  - Coproduction of evaluation system
  - Users have full ownership
- The same agencies who are not using process evaluation for improvement aggressively use similar evaluations of their other services and programs for improvement.

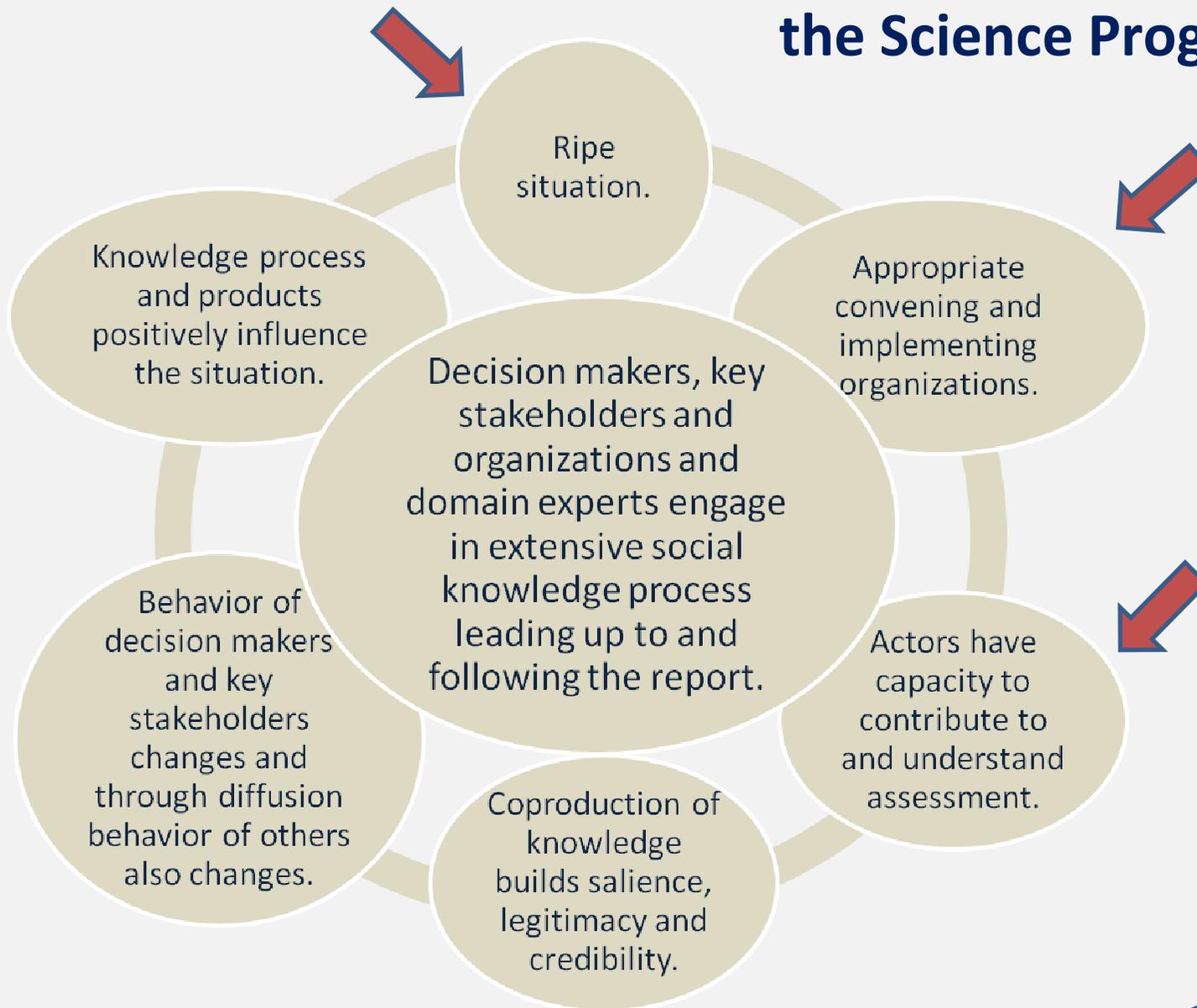
# Accountability Trumps Improvement

- All of the federal agencies use information from the system to address their performance reporting requirements
- Revisions to the system have slightly 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

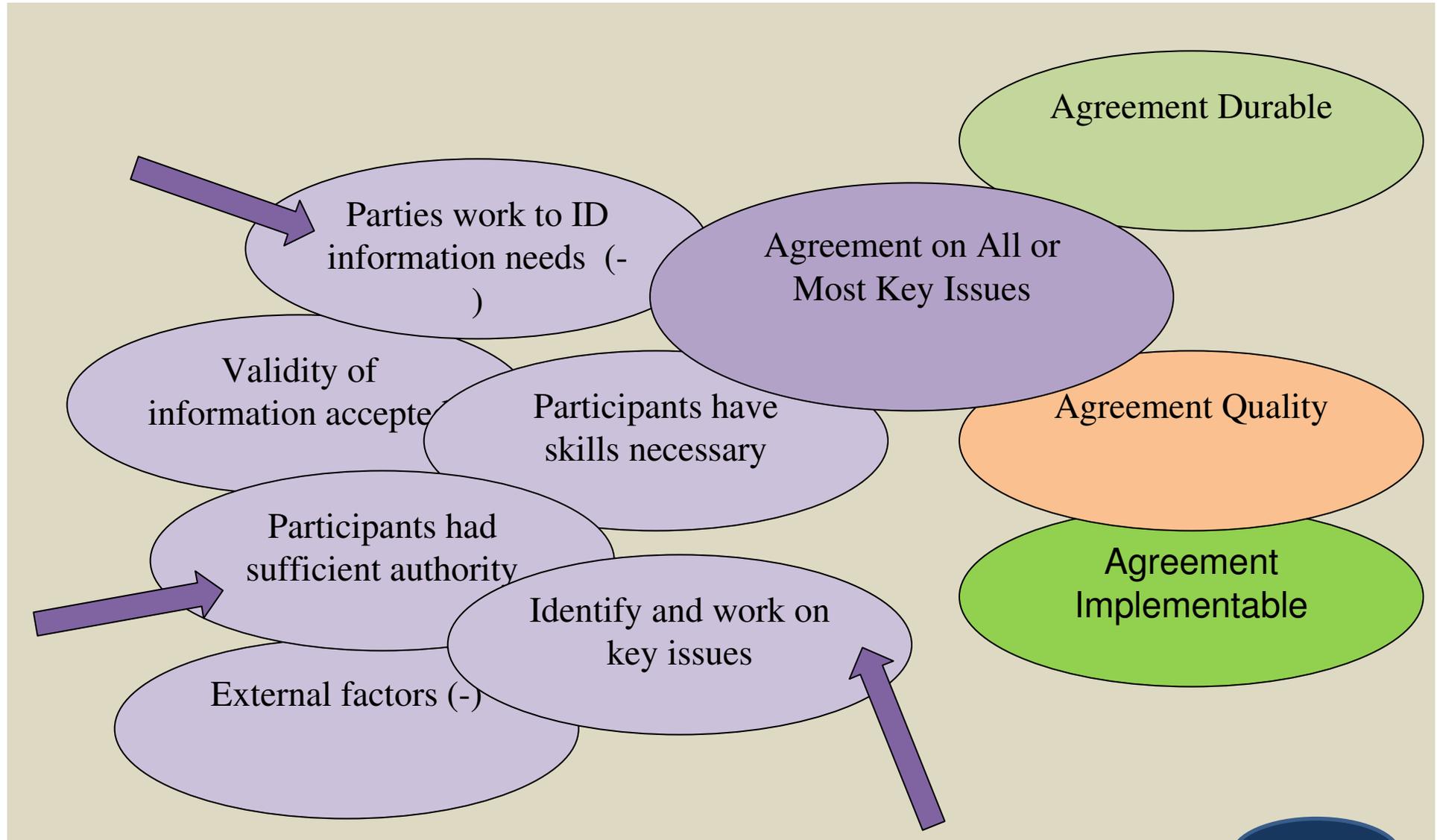
## Use at State and Federal Levels

- 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



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



## Summary

Evaluation use of formative evaluations is promoted when evaluation is undertaken as a social knowledge process, not as a technical or applied research undertaking where the presumption is that use comes from the report and will move positively with the quality of the evidence and analysis.

If accountability trumps improvement, resources for improvement will be reduced and the efficacy of evaluation will diminish – resulting in lower levels of evaluation use.

Using evaluation advice seems to be much harder than expected.