Non-Linear Impact Assessment: Challenges, Approaches and Tools

An Annotated Bibliography

Dynamical Systems Innovation Lab
Honolulu, Hawai'i
July 20-25, 2014

Briefing Paper

Group Members:
Diana Chigas and Tim Ehlinger (organizers)
Barbara Befani
Jayne Docherty
Jay Michaels
Richard Smith
David Stanfield
Tjip Walker
Ricardo Wilson-Grau
Ria Yoshida

Introduction to the Non-Linear Impact Assessment Theme

Learning and non-linear impact assessment is a fundamental issue to be addressed in the implementation of innovation that employs dynamical systems theory (DST). Complex social and social-ecological systems change in non-linear and unpredictable ways, and the knowledge about complex system dynamics and how to best affect outcomes emerges over time. All too often, monitoring and evaluation is based on pre-determined indicators (typically output metrics) that only measure attainment of and/or compliance with project goals. These metrics often provide little value for learning about the workings of the system in a way that can facilitate understanding of the effects of the work on the system and inform adaptive management to improve outcomes.

Non-linear impact assessment looks at good practice and innovation in tools and processes for measuring (evaluating) programmatic impacts in complex systems and for learning-enabled monitoring systems that can "learn fast" and systemically.

This theme is broad (i.e. many different meanings of "impact assessment"), and within our diverse group of academics, policy makers, practitioners and evaluators there are different perspectives on what is important and what the challenges are in non-linear impact assessment. Within the evaluation field, considerable work has already been done, and we did not want to duplicate that. Our group process began with a broad inquiry to elicit key challenges from a variety of perspectives and develop a picture of what the group members could and wanted to contribute:

1) What are the key challenges needs or questions about impact assessment in your field?
2) What are you working on (research, publications, practice or policy) in relation to the topic?
3) What other experience or research exists and do you think the group ought to consider in relation to the field?
4) Do you have insights or thoughts about a case study which could be examined/explored to be illustrative of this theme?
5) Are there ideas or innovations that you consider promising which we might further develop in our paper and/or in the Lab sessions?

We start from the premise that the lines between assessment, monitoring and evaluation cannot be hard and fast in complex systems, that there are a number of tools out there already (like outcome harvesting, which Ricardo introduced the group to in Sheboygan), and that we cannot produce a comprehensive or single treatment of "impact assessment" in a 20-page essay. After several conversations reviewing contributions of literature and ideas by group members, we found several pieces useful for providing a broad framework for organizing thinking on this topic as well as helping to sort through the myriad tools and methods that exist (and that we did not want to summarize in a list). This is the point of departure for this briefing, which is in the form of an (adapted) annotated bibliography, with a focus on some tools and approaches that seem to provide promise and a platform for ongoing inquiry. Some sections are summaries of literature contributed by group members; some sections were prepared by specific group members on specific topics and are so indicated.

We begin by elaborating on what is meant by "impact assessment" before describing key challenges for applying this idea to non-linear, complex systems. We then transition to discussion about how impact assessment might be implemented to evaluate programmatic impacts in complex systems with consideration of some specific, possible, tools. Finally, we conclude by identifying possible ways to continue developing this innovation.

What do we mean by "impact assessment"?

An initial challenge in considering non-linear impact assessment is understanding what is meant by "impact". In the dictionary (Oxford) *impoact* is defined broadly as "the effect or influence of one person, thing, or action, on another." This can mean different things to different people in different contexts and different interventions, including:

- The effects or results of an intervention (positive or negative), in other words, what has happened in the sphere of influence of the intervention as a result of the intervention?
- *Long term* improvement in people's lives or in the state of the environment, with the implication that "impacts" refer only to long-term, sustained changes. The OECD's Development Assistance Committee defines impact in that way as "Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended." (Glossary of Key Terms in Evaluation and Results Management, 2002).
- In peace and conflict contexts, impact often refers to the effects of an intervention on the drivers of conflict and peace, or changes in the broader conflict situation ("Peace Writ Large") to which the intervention aims to contribute, beyond effects on participants. The OECD DAC guidance, *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility*, notes that "[i]n fragile and conflict-affected contexts the criterion of impact is used to identify and evaluate the effects of the intervention on the key driving factors and actors of the conflict, as well as on broader development and statebuilding processes, as relevant."[1]
- In the evaluation profession, impact evaluation has a very specific meaning referring to causation. In this interpretation, impact evaluation is understood to mean assessment of the net effect of the intervention, as determined by what would have happened in the absence of the program. It tends to emphasize quantifying the effect. It focuses on establishing cause and effect and "must include a counterfactual or similar analysis to address attribution and establish causality."[2]
- In business and infrastructural projects, impact assessment is a prospective process, effectively a risk assessment, examining the scale, duration, severity and certainty/likelihood of impacts (mainly negative) of a project.

There is no "right" answer or definition of "impact". However, the varied definitions have different implications for what approaches are appropriate for assessment. Moreover, different stakeholders may have different interests and values about what definition should be privileged in a particular situation. Understanding and negotiating agreement about what is meant by impact and impact assessment among key stakeholders is thus important, as it not only informs what is valued as "success", but also strongly informs the shape, methods and use of the evaluation. In essence, this is a first step in what the boundary of the evaluation is – what is "in" and "out," timing, "place" (where an evaluation takes place), what is valued and not, which stakeholders are included and not, what issues are included and not and what tools and methods are appropriate and useful.

Key challenges of impact assessment

Many of the challenges of evaluating interventions that are complex or that are implemented in complex environments have been discussed extensively (*see, e.g.,* Chigas, Church & Corlazzoli 2014; Stern 2012; Funnell & Rogers 2011; Quinn Patton (2006); Williams & Iman (eds) 2006; Eoyang & Berkas 1999). We will not repeat these here. In our discussions, we have raised a number of challenges and needs for impact assessment relevant both to academics and practitioners and policy makers.

---

[1] OECD DAC (2012) *Evaluating Peacebuilding Activities in Settings of Conflict and* Fragility. Paris: OECD DAC, p. *67.*

[2] See *DfID Evaluation Policy*, para. 77, 79.

1. ***Understanding impacts across levels.*** Traditional boundaries between individual, interpersonal, group, inter-group, community and social processes still are dominant in relation to measurement and methodologies. There are different methods used to measure and assess outcomes at each of these levels. A systemic approach to impact assessment should be able to bridge the gap between levels of analysis. For practitioners, especially in peacebuilding, demonstrating how a program or intervention contributes to broader change (e.g., from individual to community level, or to address drivers of conflict) is important, both in practical terms to validate the interventions to donors and for learning about how to proceed strategically. Yet for assessment of systemic interventions, the links and interactions between processes and changes at the individual, family, community and national and even regional and global levels are important. What tools and methods can help practitioners assess impacts from a multi-level systems paradigm—to examine what is happening at different levels and across the levels? Work done in social neuroscience and by Lan Bui-Wrzosinska's has found that some subtle micro-indicators can be linked directly to key qualitative changes occurring within communities. Yet we still lack adequate theoretical, conceptual and methodological tools to link the levels and allow us to make solid inferences about macro trends from micro and meso-level indicators, and to link macro changes more clearly to the efforts of practitioners working at different levels to catalyze them.

2. ***Defining and measuring "success"/results.*** In complex situations analysts and practitioners do not have to, and indeed often cannot, plan predefined effects or results. Yet the current operating environment of most donors emphasizes the need to demonstrate impact or assess progress against pre-determined outputs and outcomes that can be measured quantitatively. A systems understanding of conflict suggests that this type of approach is unlikely to glean data that will be useful, and cannot capture emergent phenomena, or dynamic models very well. Challenges include:

- What methodologies are appropriate for assessing "results" that are not pre-determined, where programs (including goals) may evolve over time and in response to the context and differing perspectives exist on the problem?
- How can we develop indicators of system change—what would indicators of "system health" or dynamic indicator frameworks look like and how to develop them?
- How do we capture key feedbacks people pay attention to?

3. ***Assessing narratives.*** Language and metaphors are attractors around which whole systems organize. They are also part of the "glue" that connects multiple levels of a system and also the "grease" that makes the system run. Narratives might be a useful focus to provide insight into different levels of a social system and may be useful as dynamic indicators. How do we assess, measure and interpret changes in language and narratives as attractors? How can we identify the potential for and resistance to change using narratives?

4. ***Theories of change.*** How can we formulate and assess theories of change that capture the logic of social interventions and evolve over time?

5. ***Translating academic work for use in impact assessment.*** As we come from different fields—experimental, academic, practice, evaluation, policy, conflict, development, environment, psychology, etc.—one question that arises is how to bridge the academic-evaluation gap, specifically for impact assessment. How can we translate and use lab-related findings in the "external world" and how might they be useful (and what their limits might be)?

Process dimensions of impact assessment: how do we do it?

Process dimensions of systemic impact assessment include:

- What are appropriate approaches for evaluation in complex, non-linear systems?
- What quantitative and qualitative methods are appropriate and useful?
- What is the role or position of the evaluator in relation to the program and the team?
- What are implications for the evaluation process of the fact that evaluation itself is part of the system (e.g., conflict sensitivity of the evaluation, ethical concerns, uses and impacts of the evaluation itself, etc.)?

1. Planning evaluation of system change.

Hargreaves' (2010) guidance on planning evaluations of system change interventions provides a useful overarching framework for developing an approach for assessing impact systemically, and choosing appropriate evaluation methods and designs for the complexity of the situation and the intervention. Hargreaves lays out a three-step process analyzing the complexity of the situation or system, the nature and complexity of the intervention and the purposes, users and methods appropriate to the system, the intervention and the evaluation purposes.

Hargreaves suggests there are three elements of a systems change evaluation:

a) Understand the conditions and dynamics of the situation or system
   i. Identify the system's or situation's boundaries—what is "in" or "out" of the evaluation. What are appropriate boundaries for the system to be evaluated (determining the scope of the evaluation) so that they are not too broad and overwhelming but also not so narrow as to exclude important system perspectives and elements? And, what is the proper scale for the system or situation (level of analysis)?
   ii. Determine the complexity of the relationships. What are the connections or exchanges that occur within and across system levels? What are the structure and nature of the inter-relationships in the situation and the processes and dynamics between the elements of the structure? Here social network analysis, causal loop diagramming or drawing a rich picture can be useful in generating a systemic understanding of the situation.
   iii. Explore multiple perspectives. There may be a diversity of perspectives regarding the problem, goals and strategies for addressing the problems, and the intervention, among different stakeholders that are important to explore.

Hargreaves draws heavily on Bob Williams' work on systemic inquiry in the development of her planning framework. In *Systems Concepts in Evaluation: An Expert Anthology* (2007), *Systems Concepts in* Action (Williams & Hummelbrunner 2010) and his most recent book, *Wicked Solutions* (Williams & van t'Hof 2014), Williams identifies inter-relationships, perspectives and boundaries as three common threads to all systemic methods and approaches; these form the basis for developing an understanding of the situation or system in Hargreaves' model to determine the boundaries, scope and focus of the evaluation, and provide an understanding of the situation to assess change. These areas might also be areas for inquiry and critical examination in relation to the intervention itself and choices made by program staff as part of the evaluation process. Particularly in relation to exploration of perspectives and boundaries, choices made regarding which perspectives (framings of the problem and purposes) to privilege and which not; what the goals of the intervention are; what resources are made available to decision makers and what they do not have control over, what knowledge and expertise is provided and honored, by whom; how the intervention identifies and handles people who are harmed or marginalized by the intervention—set boundaries and establishes who and what is "in" and "out" of the intervention. These choices have ethical implications, implications for the ability of an intervention to achieve "success" and

potential negative conflict implications and should be examined—Williams would say critiqued—rigorously (Williams & van t'Hof 2014).

When applied to a complex system, some additional considerations might add to Hargreaves' general approach. These include:

First, understand the elements and dynamics of the intervention:

- Understand the dynamics of the intervention's governance (e.g., funding, management, decision making, structure, implementation, etc.). Complex efforts may involve collaboration among multiple actors in multiple sectors across multiple levels.
- Identify causal mechanisms (theories of change). In simple situations, linear logic models can be used, but more complex system interventions may require more complexity-based theories of change, and benefit from causal loop diagramming, ecosystem mapping or other visualizations, as well as identification of simple rules and processes of self-organization. It is important to identify misalignment between the intervention and the system it is seeking to change.
- Analyze the intervention's intended outcomes. These may be very well-defined, focused and measurable in simple situations. In complex situations, stakeholders may share a common vision but may not be able to predict specific outcomes to emerge from the intervention.

Second, determine the users, purposes and methods of the evaluation. The methods should fit the evaluation's purposes and be appropriate for the level of complexity of the intervention and the context.

A piece missing from this framework is the identification of pre-existing data and knowledge (quantitative and qualitative) about the situation and what gaps exist, as part of the process of understanding the conditions and dynamics of the situation. Pre-existing information may reveal problems invisible to program staff or evaluators, and gaps may reveal problems in the system itself. Pre-existing data might come from quantitative data (e.g. census data, economic indicators), written histories, archive data, or narratives.

2. Developmental evaluation (Diana Chigas).

Developmental evaluation is a term coined by Michael Quinn Patton to describe an evaluation process for social innovation and complexity-based interventions. It views program design and implementation, the purposes and uses of evaluation, and the role of the evaluation differently than traditional evaluation. *Developmental evaluation* refers to long-term, partnering relationships between evaluators and those engaged in innovative initiatives and development. It is designed for innovative and emergent programs – where a vision may exist, but for which it is impossible to set predetermined goals. Accordingly developmental evaluation is useful where the program, including the very goals and outcomes, are being adjusted continuously in responses to the changing context, new learning, new opportunities and constraints. The purpose of evaluation is not to judge success or failure, but to provide feedback, generate learning and affirm direction or support changes in direction when appropriate.

In contrast to traditional evaluation processes, the evaluator is part of a developmental initiative or innovation team whose members collaborate to conceptualize, design and test new approaches in a long-term, on-going process of continuous improvement, adaptation, and intentional change. The evaluator's primary function on the team is to promote evidence-based reflection, decision making and adaptation—informing team discussions with evaluative questions, data and logic, and to facilitate data-based assessments and decision-making in the unfolding program. The evaluator essentially accompanies the program, introduces feedback supported by data that helps to reality-test how things

are unfolding, discerns what directions hold promise and which ones should be abandoned, and supports consideration of what new approaches should be tried.

When using developmental evaluation. data-gathering is open-ended and frequent, questions and concerns are emergent, and trial and error is carefully mined for learning. Data-gathering also includes systematic and robust documentation of the innovation process; as the program design and implementation is often emergent, tracking the process supports accountability by documenting choice points, decisions made and allowing reflection on the implications of each decision and the cumulative effect of many smaller decisions. Developmental evaluation does not rely on any specific approach or method, but draws on many of the methods familiar in evaluation, such as surveys, interviews, quantitative studies, as well as more complexity-based tools (e.g., network analysis, agent-based modeling, etc.). Therefore, developmental evaluation may be especially useful when working with complex problems that change rapidly and in non-linear ways.

Patton (2006) summarizes the differences between traditional and developmental evaluation as follows:

TRADITIONAL EVALUATIONS...	COMPLEXITY-BASED, DEVELOPMENTAL EVALUATIONS...
RENDER DEFINITIVE JUDGMENTS OF SUCCESS OR FAILURE.	PROVIDE FEEDBACK, GENERATE LEARNINGS, SUPPORT DIRECTION OR AFFIRM CHANGES IN DIRECTION.
MEASURE SUCCESS AGAINST PREDETERMINED GOALS.	DEVELOP NEW MEASURES AND MONITORING MECHANISMS AS GOALS EMERGE & EVOLVE.
POSITION THE EVALUATOR OUTSIDE TO ASSURE INDEPENDENCE AND OBJECTIVITY	POSITION EVALUATION AS AN INTERNAL, TEAM FUNCTION INTEGRATED INTO ACTION AND ONGOING INTERPRETIVE PROCESSES.
DESIGN THE EVALUATION BASED ON LINEAR CAUSE-EFFECT LOGIC MODELS.	DESIGN THE EVALUATION TO CAPTURE SYSTEM DYNAMICS, INTERDEPENDENCIES, AND EMERGENT INTERCONNECTIONS.
AIM TO PRODUCE GENERALIZABLE FINDINGS ACROSS TIME AND SPACE.	AIM TO PRODUCE CONTEXT-SPECIFIC UNDERSTANDINGS THAT INFORM ONGOING INNOVATION.
ACCOUNTABILITY FOCUSED ON AND DIRECTED TO EXTERNAL AUTHORITIES AND FUNDERS.	ACCOUNTABILITY CENTERED ON THE INNOVATORS' DEEP SENSE OF FUNDAMENTAL VALUES AND COMMITMENTS.
ACCOUNTABILITY TO CONTROL AND LOCATE BLAME FOR FAILURES.	LEARNING TO RESPOND TO LACK OF CONTROL AND STAY IN TOUCH WITH WHAT'S UNFOLDING AND THEREBY RESPOND STRATEGICALLY.
EVALUATOR CONTROLS THE EVALUATION AND DETERMINES THE DESIGN BASED ON THE EVALUATOR'S PERSPECTIVE ON WHAT IS IMPORTANT.	EVALUATOR COLLABORATES IN THE CHANGE EFFORT TO DESIGN A PROCESS THAT MATCHES PHILOSOPHICALLY AND ORGANIZATIONALLY.
EVALUATION ENGENDERS FEAR OF FAILURE.	EVALUATION SUPPORTS HUNGER FOR LEARNING.

3. "Outside in" and "inside out" analysis: processes for evaluating impacts in peacebuilding interventions? (Diana Chigas)

Chigas, Church & Corlazzoli (2014) outlined a process for evaluating broader impacts of peacebuilding efforts, i.e. effects on drivers of conflict or peace. A distinguishing characteristic of evaluations of impacts

in peacebuilding is that the changes an intervention has produced or contributed to must relate to the dynamics or drivers of conflict and peace. When assessing impacts on "peace writ large"—or contributions to a broader peacebuilding dynamic—it is not enough to demonstrate that an intervention has contributed to changes in the *context or the individual*; the intervention must be *relevant* to those aspects of the context of conflict or fragility that directly shape or influence how the conflict develops and whether violence occurs (OECD DAC 2012). If not, then the intervention may do good and have impacts, but still not change any underlying dynamics or drivers of conflict, and therefore not have *peacebuilding* impacts.

Conflict analysis is therefore key to evaluation of impacts of peacebuilding programming and policies, along with an analysis of the *relevance* of the intervention to the conflict. Conflict analysis helps to understand key driving factors of conflict and peace and the evolution of those drivers and dynamics over the time period covered. Relevance analysis identifies whether the broader changes targeted by the intervention are connected to the needed changes specified in the conflict analysis and thus whether the intervention makes a logical contribution to meeting peacebuilding needs. Together these types of analysis help evaluators a) identify any change in the conflict situation to which the program may have contributed, and b) identify and unpack the hypotheses or theories of change of the intervention in relation to peacebuilding—i.e. the assumptions of how it will contribute to addressing the drivers of conflict—so they can be tested in the evaluation (Chigas, Church & Corlazzoli 2014).

Combining conflict analysis with an understanding of the theories or assumptions about how interventions are supposed to address conflict drivers allows impact evaluations to provide both 'inside out' and 'outside in' analysis. This is important for understanding non-linear cause and effect relationships as well as how constraints and opportunities, as well as effects catalyzed by an intervention in combination with other factors, have significant effects on the nature, scale and scope of 'impacts'. The 'outside in' analysis uses context and conflict analysis to gain an understanding of how a conflict has evolved, as well as what the constraints and opportunities for an intervention's influence on it are at any particular moment in time. The 'inside out' approach looks at the intervention itself, its activities, outputs and outcomes and the extent to which these are reflected in the working (rather than assumed) theory of change. By linking these two perspectives, evaluators can minimize issues of bias by placing the intervention in the larger context in which it is operating (and of which it is a part) and can identify the range of factors and dynamics, including the intervention, that have influenced the evolution of the conflict.

Tools and approaches

Hargreaves (2010) provides a variety of methods useful to evaluate systems that display a variety of levels and dynamics. She specifically proposes to match evaluation approaches and methods, both quantitative and qualitative, to the level of complexity of the system and the intervention (see Table 1). She uses the distinction between simple, complicated and complex systems that has become widely accepted in the field (see, e.g., Britt 2013, Funnell & Rogers 2011, Patton 2007) as a basis for choosing what methods and approaches are appropriate.[3] Traditional research methods, she notes, are best

[3] Simple, organized dynamics are characterized by fixed patterns of behavior, and predictable and known (often linear) cause-effect relationships resulting in high degrees of certainty and agreement about how to solve a problem. In complicated dynamics, cause-effect relationships are detectable or knowable, but may be separated over time and space, and while there may be agreement among stakeholders on how to solve a problem, it may not be so clear how to do it. In complex dynamics there is neither certainty

suited for simple, organized system dynamics, while others (such as causal loop simulation modeling, social network analyses, case studies) are better suited for complicated dynamics and still others (e.g., GIS spatial analysis, agent-based modeling, case studies, outcome harvesting, appreciative inquiry, etc.) designed for situations with complex adaptive dynamics.

nor agreement on how to solve the problem, and cause and effect are understandable, but not predictable in advance.

Table 1. System Dynamics and Evaluation Methods

System Dynamics	Quantitative Methods	Qualitative Methods
Random Dynamics: • Random activity—no discernible pattern • Unconnected collection of parts • No detectable cause-effect relationships • No purpose or direction—people act independently, react blindly, or avoid action • Turbulence—no stability or equilibrium • Answers are unknowable		Case studies, interviews, and focus groups, observation of activities Mapping of community assets Environmental scans, Needs assessments Situational analyses One–time rapid assessments
Simple Dynamics: • Stable, static pattern • Parts tightly connected • Predictable cause-effect relationships • System can be reduced to parts and processes and copied or replicated as best practices • Directive leadership, designed change • Answers are knowable, obvious, prescriptions for action	Randomized experiments Quasi-experimental comparisons Regression discontinuity analyses Hierarchical linear modeling Performance measurement and monitoring Program audits and inspections	Case studies, interviews, and focus groups Thematic content analyses Purposive sampling of relevant cases Reviews of program documentation Literature reviews Logic models
Complicated Dynamics: • Dynamic patterns of feedback loops with many coordinated, interrelated parts within and across system levels • Cause and effect separated by time and space • Self-correcting feedback loops maintain equilibrium • Self-reinforcing feedback loops disrupt equilibrium • Answers are knowable through expert analysis	Computer simulation models of stocks, flows, and feedback Computer simulation models of causal loops Social network analyses Pre-post measurements of change Interrupted time-series analyses Comparative measurement and monitoring	Case studies, interviews, and focus groups Thematic content analyses Reviews of program documentation Literature reviews Causal loop diagrams, participatory impact pathways analysis System mapping, ecosystem maps
Complex Dynamics: • Dynamic patterns—parts adapting, coevolving with each other and environment • Parts are massively entangled and interdependent; nested webs and networks • System change emerges through interactions among parts; cause and effect are known in retrospect • Equilibrium is in flux, sensitive to initial conditions • Parts self-organize, learn, and change	GIS spatial analysis Agent-based modeling Time trend analyses Observational or cross-sectional studies Retrospective analyses Adaptive learning measurement systems	Case studies, interviews, and focus groups Observation of activities Document reviews Outcome mapping, concept mapping Analyses of emergent systemwide patterns, tracking of events, encounters, and policy changes Use of simple rules and conditions of self-organization Soft systems methodology Appreciative inquiry, reflective practice

Sources: Holland 1995; Eoyang and Berkas 1998; Snowden and Boone 2007; Leischow et al. 2008; Parsons 2007; Bloom and Dees 2008; Douthwaite et al. 2008; Kramer et al. 2008; Patton 2008; U.S. GAO 2009; Hargreaves and Paulsell 2009; Goldsmith et al. 2010; Patton 2010b; and Williams and Hummelbrunner 2010.

A number of the tools and methods listed here will be familiar to researchers and practitioners, while others are not. Rather than describe all of them, we can highlight a few approaches and methods that the group has brought up (and in which the group members have some expertise and knowledge) to explore.

1. ***Narratives (Jay Michaels)***

Narratives can provide insight into different levels of a social system (e.g. small group versus large-scale culture), and can inform about social relationships at a community-level. Furthermore, narratives might be useful as dynamic indicators, especially when examining changes in narratives.

While there are many ways to define "narrative," one useful definition may be that a *narrative* is the set of explanations a person or group of people hold for some particular event. Construed in this context, narratives may reveal attitudes and beliefs and information about perceived victims, offenders, and transgressions in conflict situations. Furthermore, narratives may contain insight into emotions, which are especially neglected in conflict research (e.g. Nair, 2007).

The following summarizes some sources that address this idea of narratives and communication as related to individual and group dynamics.

a) Use of narratives to examine cultural differences and change. Some social psychologists (Chiu & Hong, 2007) suggest that *knowledge structures*, or the ways of understanding used to create a sense of experienced, grasped reality, may be insightful when considering different cultures and cultural change. The key to this idea is that people tend to operate with two forms of knowledge that may manifest in a narrative: procedural knowledge and declarative knowledge. Procedural knowledge refers to knowing *how* to achieve a goal. Declarative knowledge refers to the *what* – the characteristics of people, objects, or events.

For evaluation purposes, declarative knowledge might be useful as Chiu and Hong (2007) subdivide declarative knowledge into knowledge about people, events, and norms. Accordingly, their chapter may offer a framework for evaluating people's recollections of events and explanations for any given situation. Analysis could then focus on a) how people are portrayed in the narratives, b) how events are framed, and c) the behaviors expected to emerge from preexisting conditions (i.e., the norms). Such analysis could reveal accurate and faulty perspectives in addition to bias (e.g. stereotypes).

b) Meaning making is automatic. Meaning-making is emphasized throughout the book, *Handbook of the Psychology of Religion & Spirituality* (Paloutzian & Park, 2013). While different authors draw upon the automatic tendency people have to create meaning to examine religion and spirituality, the ideas are easily extended beyond a faith-focus.

Park's (2013) chapter on "Religion and Meaning" is insightful and emphasizes the idea that people tend to frame events in the most global perspective possible. That is, people prefer to understand their actions and the world from a broader, coherent perspective that simplifies and avoids excessive detail. While not discussed in this particular chapter, it is easy to see that this type of automatic worldview simplification can readily lead to stereotypical thinking and use of mental or decision-making errors (cognitive heuristics). And, this idea also seems linked to the attractor notion of conflict (as a conflict or problem worsens, there is a tendency for the situation to become simplified in the minds of those involved).

Park's (2013) chapter adds to these ideas noting that people first tend to use their global meaning system to understand specific situations. However, what's of most interest is the idea that people may *change* their preexisting meaning (i.e., narrative) when faced with stress. This idea goes along with those from *Action Identification Theory*, which posits that people prefer to understand themselves, others, and situations more globally. However, when faced with difficulty, people will tend to shift thinking to

instead focus on the details. This shift to the details permits people to find new ways of understanding that may promote problem solving (e.g. Michaels, Parkin, & Vallacher, 2013). These ideas are further linked to recent theories and studies about emotions that propose negative emotions alert people to situational problems so that they may mentally adapt to the situation (e.g. Isbell, Lair, & Rovenpor, 2013).

Could these ideas be informative when examining narratives about social problems? Perhaps *shifts* in a narrative from a focus on global ideas to specific details could signal system change?

c) Using Narratives to Examine Attractors & Change. As noted above, many psychology theories elaborate on how people frame their world depending on whether they are in a period of stability or change.

Action Identification theory is useful for examining implicit stability versus destabilization in how people explain their own actions as well as those of others. The core idea with action identification is that people automatically assume the most complex way of framing their actions, and that these "higher level" ways of thinking reflect greater stability. For example, if I were to think of my current action in writing this paper as "synthesizing knowledge," that is decisively higher level than "typing notes." Empirically, lower level frames tend to be far less stable than higher level frames. In an applied setting, the core idea of action identification theory is simply that assessing how people communicate about a situation (via narratives) can reveal the level of complexity and abstraction in their thought pattern. Importantly, situations where thoughts are based on complex, abstract notions are likely to remain stable whereas situations involving simpler, more concrete thoughts are likely to present greater fluidity. For evaluating any program, simpler patterns of thought are thus more likely to signal that change is in fact taking place.

As a tangible example, consider effective goal-setting and planning. It is well known that one of the most effective ways to accomplish difficult tasks is for an individual to break-down such tasks into a series of smaller, achievable, concrete steps. Therein, "writing a research article" could be broken down into "drafting an opening paragraph" and then "writing the first part of the literature review." The higher level frame ("writing a research article") is difficult to tackle; the lower level frames (e.g. "drafting an opening paragraph") become more concrete, easier to achieve, and promote progression.

d) Narratives have shown up in conflict research. There are many studies that have used narratives to examine conflicts and/or specific contexts of a given conflict. Some of these include…

Habashi (2008). Palestinian children crafting national identity. *Childhood, 15*, 12-29. – Discusses use of interviews to understand identity in relation to the Israeli-Palestinian conflict and how various levels of social reality (e.g. family unit to larger media) play a role with this. Declarative knowledge in terms of the person is emphasized, especially how the children interviewed frame their self-identity versus identity of others.

Sagy & Adawan (2006). Hope in the time of threat: The case of Israeli and Palestinian Youth. *American Journal of Orthopsychiatry, 1*, 128-133. This study used a more structured approach when it comes to narratives. The researchers used surveys that included examination of Israeli and Palestinian beliefs about the dispute's historical narratives to try to understand Israeli's and Palestinian's sense of hope.

Shalhoub-Kevorkian (2009). The political economy of childrens' trauma: A case study of house demolition in Palestine. *Feminism Psychology, 19*, 335-342. In contrast to the prior article that focused

on a highly structured method and large sample, this article demonstrates how specific and deep a narrative approach may go to understanding social problems. A main focus in this article is emphasis on how the narrative approach permits insight beyond the typical socio-political focus that emerges in many conflict analyses.

e) A "narrative" approach has been used to examine complexity in social interactions. There are two prominent lines of research that have used what could be considered a narrative approach to gain deep insight into complex social interactions. John Gottman's extensive work on marriage involved coding behavior and communication between husbands and wives. Gottman's work led to his developing mathematical models about love-relationships. In addition, Losada used a structured approach to analyze the interaction between people who were part of work groups (teams). From lab work and mathematical modeling, Losada was able to identify differences in high-performance versus low-performance teams. The precise ways both researchers examined human communication could provide insight into how to utilize narratives as a tool for evaluating programs and change in dynamic systems.

2. Dynamic Indicators and Social-Ecological frameworks (Tim Ehlinger)

USAID's Discussion Note on *Complexity-Aware Monitoring* (2013) proposes developing "sentinel indicators," or indicators of system health, a concept borrowed from ecology referring to easily communicated indicators that capture the essence of the process of change, as a way of monitoring systemic change. In ecology, a species may be designated as a sentinel of the overall health of an ecosystem. Identification of sentinel indicators begins with a holistic picture of the project and the system, mapping the intervention in relation to its context. "Sentinel indicators are placed at critical points in a system map to help monitor and inform the mutually influencing relationships between the program and its context" (Britt 2013: 8)—for example, leverage points.

In considering indicator frameworks, we need to reflect on the tensions and dilemmas of monitoring and evaluating complex systems with indicators. There is an inherent tension between indicator-based monitoring and complexity: indicators are based on what we expect might change, but complex aspects of a situation make it difficult to predict what will change and how (Britt 2013: 9). They may need to be supplemented by more open-ended inquiry with a range of stakeholders (stakeholder feedback) (see Golwyn & Chigas 2013). Sentinel indicators are also proxies, and therefore necessarily provide incomplete information about a social system; judgments based on a single indicator (or several) can be dangerous (Britt 2013: 8). Finally, some analysts have questioned the direct transfer of models from the natural sciences to social systems—and some have assessed application of complexity theory to social contexts "as just another manifestation of scientific reductionism"; it will be important to be cautious about integrating analogous processes from ecology or the natural sciences and be clear about what kind of complexity theory is being applied, why and how (Mowles 2014: 164).

Philosophical debates about the nature of knowledge, such as the merits of mechanistic reductionism vs. rational constructivism, inductive vs. deductive logic, etc. continue to play out over and over again in various literatures – catalyzed in some cases by the desire to draw meaning from the patterns in "big data" and use "predictive analytics" to elucidate mechanistic relationships in complex systems. For example, the perceived problem in many ecologic studies is that data are aggregated over and across groups rather than on individuals (Wakefield 2007). This can result in the false identification of relationships between purported causal factors and outcomes, especially if the loss of information due to aggregation generates bias toward geographic (or group) correlations as compared to an individual-level or agent-based model. The key to avoiding ecological bias is to ensure that a mixture of individual and ecologic data are used to cross-validate hypotheses. This idea may be extended to assessing social

systems by being mindful that evaluation should take place across levels and that data gathered from micro-levels may not provide deep insight into what happens at a macro-level, or vice-versa.

One promising approach is to apply the well-established frameworks of risk assessment and risk propagation used in public health and environmental engineering (sensu Novonty et al. 2005, Lawrence 2013) within the complex contexts of social-ecological systems (Gunderson and Holling, 2002, Ostrom 2009, Biggs et al. 2012). Complex, *social-ecological systems* (SESs) are composed of multiple subsystems and internal variables within these subsystems at multiple levels (see Figure). Ostrom's (2009) SES framework provides initial indicator domains across sub-systems, including *External Drivers:* (S) Social, economic and political settings and (ECO) Related ecosystems, and *System Dynamics:* (RS) Resource Systems, (RU) Resource Units, (GS) Governance Systems, (U) Users, in addition to (I) Interactions and (O) Outcomes.

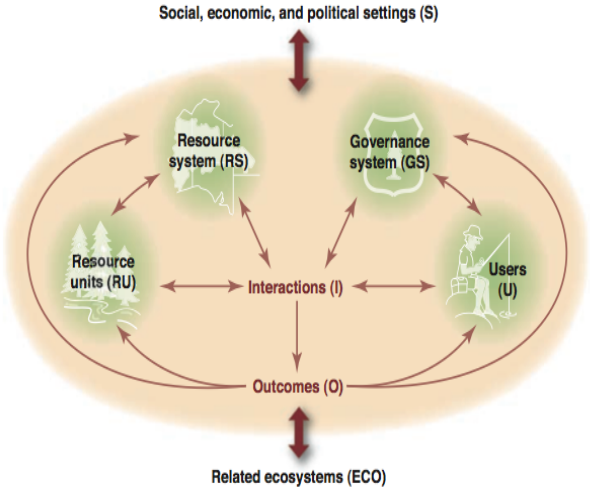


Fig. 1. The core subsystems in a framework for analyzing social-ecological systems.

Ostrom 2009

Table 1. Examples of second-level variables under first-level core subsystems (S, RS, GS, RU, U, I, O and ECO) in a framework for analyzing social-ecological systems. The framework does not list variables in an order of importance, because their importance varies in different studies. [Adapted from (*12*)]

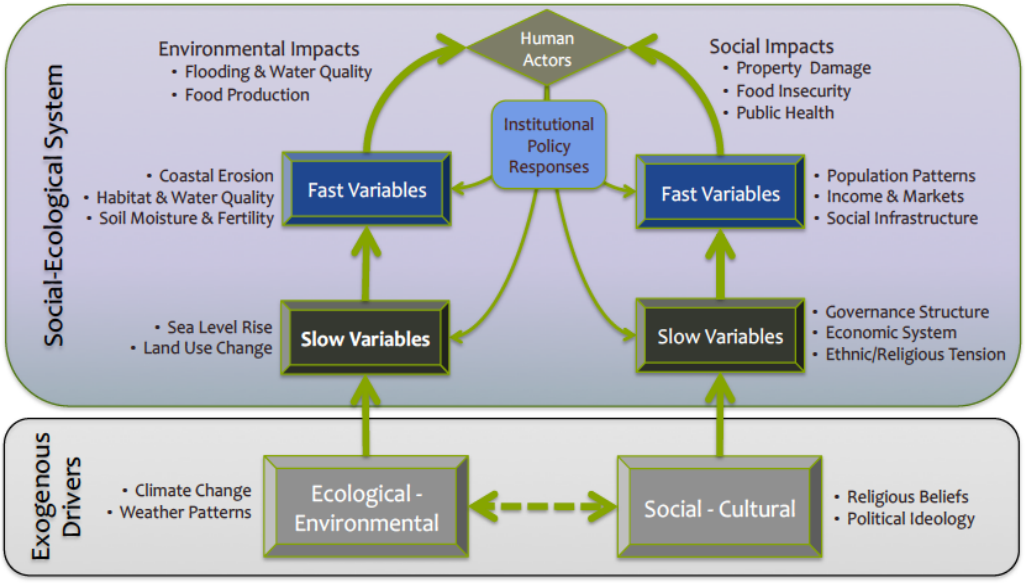
Social, economic, and political settings (S)	
S1 Economic development. S2 Demographic trends. S3 Political stability. S4 Government resource policies. S5 Market incentives. S6 Media organization.	
Resource systems (RS)	*Governance systems (GS)*
RS1 Sector (e.g., water, forests, pasture, fish)	GS1 Government organizations
RS2 Clarity of system boundaries	GS2 Nongovernment organizations
RS3 Size of resource system*	GS3 Network structure
RS4 Human-constructed facilities	GS4 Property-rights systems
RS5 Productivity of system*	GS5 Operational rules
RS6 Equilibrium properties	GS6 Collective-choice rules*
RS7 Predictability of system dynamics*	GS7 Constitutional rules
RS8 Storage characteristics	GS8 Monitoring and sanctioning processes
RS9 Location	
Resource units (RU)	*Users (U)*
RU1 Resource unit mobility*	U1 Number of users*
RU2 Growth or replacement rate	U2 Socioeconomic attributes of users
RU3 Interaction among resource units	U3 History of use
RU4 Economic value	U4 Location
RU5 Number of units	U5 Leadership/entrepreneurship*
RU6 Distinctive markings	U6 Norms/social capital*
RU7 Spatial and temporal distribution	U7 Knowledge of SES/mental models*
	U8 Importance of resource*
	U9 Technology used
Interactions (I) → outcomes (O)	
I1 Harvesting levels of diverse users	O1 Social performance measures (e.g., efficiency, equity, accountability, sustainability)
I2 Information sharing among users	
I3 Deliberation processes	
I4 Conflicts among users	O2 Ecological performance measures (e.g., overharvested, resilience, bio-diversity, sustainability)
I5 Investment activities	
I6 Lobbying activities	
I7 Self-organizing activities	O3 Externalities to other SESs
I8 Networking activities	
Related ecosystems (ECO)	
ECO1 Climate patterns. ECO2 Pollution patterns. ECO3 Flows into and out of focal SES.	

*Subset of variables found to be associated with self-organization.

Although the SES concepts were developed initially for problems associated with the ecological economics and stewardship of common-pool resources, they have been extended to apply more broadly across diverse systems (Chapin et al. 2009). SES sub-systems and component indicators also can be placed within spatial, temporal or community contexts, which provides the matrix for aggregation and statistical analysis of data metrics and indicators (e.g. see example figure below for climate change in the Black Sea Coastal Zone).

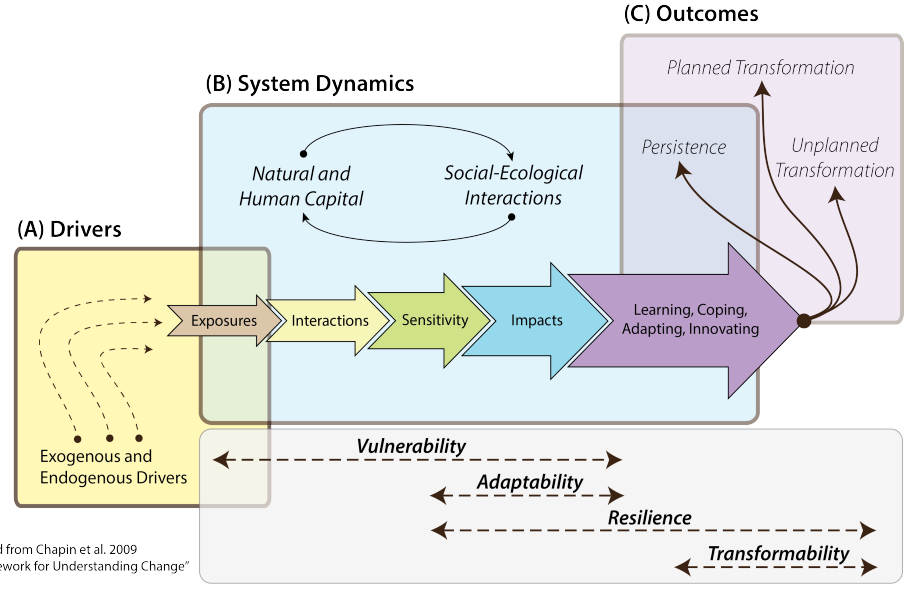
The concept of *Risk Propagation* provides an analytical framework through which indicators are mapped to stressors using a layered hierarchical model that assumes that the probabilistic relationships between

layers are modified due to various anthropogenic and natural impacts. Propagation Models typically consists of 4 layers; (1) *Root Stressors* that act on a global, regional or local scale; (2) *Drivers* of change that create exposure to risk factors; (3) *Risk Probabilities* associated with exposures. These are typically expressed by numerical probabilities of undesirable hazards; and (4) *Impact Endpoints* reflect measures of system-related goods and services of value to the public (Novotny et al. 2005).



Ehlinger (2014)

Such models can be adapted to examine the impacts of Stressors on components of a SES (or conversely the impact of interventions. In the heuristic model shown here, **Drivers (A)** are functionally linked to **Outcomes (C)** by the manner with which disturbances impact **System Dynamics (B)**. In this model, Risk propagates through the system *initiated* by exposures to the drivers, *modified* by interactions among exposures and exposed system components, *weighted* by differences in sensitivity, *measured* by quantitative and qualitative impacts on key indicators, and *expressed* in changes in the system based upon its the ability to learn, cope, adapt, and innovate. Each arrow may be comprised of multiple indicators, each of which will be influenced by probabilities of effects (risks) associated with each arrow. This process in turn influences the form of outcomes ranging from persistence to planned or unplanned



Modified from Chapin et al. 2009
"A Framework for Understanding Change"

transformations.

Another innovation of SES framework is the inclusion of **Community Characteristics (D)**, which include assessments of: (1) community vulnerability to changes (impacts) brought on by exposure to specific or unspecified drivers, (2) the adaptability of the community to maintain its ability to function in the face of disturbance, (3) the resilience of the community with regards to its ability to gather information and respond to influence outcomes, and (4) the transformability of the community. These can be attributed to both ecological and human communities that are part of the interaction within the SES. In addition, they can be used as a guide for examining the association linked traits (indicators) across different phases of the risk propagation sequence in the SES. This is particularly useful when examining suites of indicators in *ex post* evaluations and impact assessments. Advances in computational analytics using artificial neural networks and other non-linear modeling procedures have made it much easier to explore multivariate associations among indicators in complex systems.

For example, Dellinger et al. (2013, 2014) used a 2-step methodology using artificial neural networks (Estep 2006) to conduct community-level risk propagation assessment. Their approach first employs *self-organising feature maps* (SOM) to generate multidimensional clusters that visualize various outcome syndromes (e.g. ecological health, causes-of-death, and birth outcome metrics) and then utilizes *supervised learning and risk profiles* to identify key factors influencing dynamics of the system. The method has been successful in identifying exposure-outcome linkages at the county and watershed level for counties Wisconsin and for civil divisions in Romania (Dellinger et al. 2013 and 2014). The use of SOM visualizations combined with the risk profiling approach produces a systems-based mapping output that fosters discussion among the participating stakeholders. Risk profiles do not unequivocally demonstrate linear cause-effect relationships, but rather provide tools for identifying leverage points for targeted investigations and for risk-management prioritization, and are useful in situations where multiple stressors are implicated in the degradation of the system. The profiles can be manipulated to simulate how the system changes in response to changing individual stressors alone or in combination. This process not only assists in producing a more robust understanding of the feedbacks, but also is an effective learning-enabling tool that helps participant come to understand the situational-dependence, indirect effects and unanticipated consequences that single-issue interventions can have on the system. Supervised learning also has the ability to identify the key factors influencing cluster identity, and thereby assists in selecting sentinel indicators (Britt 2013). The prediction profiles can also detect and show distinct threshold effects and demonstrate the sensitivity of how changing one parameter may produce unexpected results because of how it changes the shapes of other relationships in the system.

3. Goal-free evaluation approaches

Outcome harvesting (developed by Ricardo Wilson-Grau) and Most Significant Change (MSC) (developed by Rick Davies and Jess Dart) are two approaches that do not measure progress toward predetermined outcomes but identify and collect evidence of what has been achieved. They are both, in principle, highly participatory. These approaches, along with stakeholder feedback mechanisms, are identified in another recent publication the group found useful, USAID's Discussion Note on *Complexity-Aware Monitoring* (2013) as particularly appropriate for complex situations, where cause and effect relationships are poorly understood and there is little agreement among stakeholders about how to solve the problem being addressed. Specifically, these approaches are able to address "blind spots" in performance monitoring: dealing with a broad range (including unintended) of outcomes of an intervention, identifying alternative causes from other actors and factors, and considering non-linear pathways of contribution (Britt 2013).

Outcome Harvesting (Wilson-Grau & Britt 2013; World Bank Institute 2014) is a utilization-focused, highly participatory tool that enables evaluators, grant makers, and managers to identify, formulate, verify, and make sense of outcomes they have influenced when relationships of cause-effect are unknown. It is inspired by Outcome mapping, which was developed at the International Development Research Centre (IDRC) in Canada, as a participatory approach for planning, monitoring and evaluation of interventions (see Earl, Carden & Smutylo 2001). The distinguishing feature of outcome mapping and outcome harvesting is that they focus on how an intervention has contributed to changes in behavior of and relationships among actors that it intends to influence and who are key to bringing about the macro-level impacts it intends to contribute to. This focus is designed to address the difficulty of attributing impact to a program, i.e. establishing a sustainable improvement in the environment or well-being of a large number of people, isolating factors that caused the changes and attributing them to a particular program. In complex environments, such attribution is nearly impossible, and thus outcome mapping focuses on results achieved "upstream" from impact. It does this by focusing on the changes that are clearly within a program's sphere of influence. "While, at first glance, this appears to suggest concentrating on easier, less important, short-term achievements, in fact it does the opposite. It focuses attention on incremental, often subtle changes, without which the large-scale, more prominent achievements in human well-being cannot be attained or sustained… The intended 'impact' of the program is its guiding light and directional beacon, a test of its relevance — it is not the yardstick against which performance is measured." (Earl, Carden & Smutylo 2001).

Unlike some evaluation methods, Outcome Harvesting does not measure progress towards predetermined outcomes or objectives, but rather collects evidence of what has been achieved, and works backward to determine whether and how the project or intervention contributed to the change.

There are six iterative steps of the process:

1. Design the harvest
2. Review documentation and draft outcomes
3. Engage with informants
4. Substantiate
5. Analyse and interpret the findings
6. Support use of findings

Outcome harvesting identifies useable questions, and then collects information on outcomes from documentation to develop 'outcome descriptions' documenting the change in the key actors in specific and measurable ways. The 'outcome harvesters' (evaluators) engage with the change agents to review

information, collect additional information, examine and substantiate the system coherence, and specify as well as inform on the plausibility of the explanations of how the intervention has contributed to outcomes. Triangulation of data sources, including collection of outside views from people familiar with the intervention but independent from it, is used. In addition, rigorous approaches such as General Elimination Method are used to rule out alternative explanations for the outcomes. A June 2014 compendium of ten cases (World Bank Institute 2014) presents the results of World Bank pilot learning to explore how outcome harvesting tools might help project teams and stakeholders to learn from complex and difficult to monitor development processes.

Most significant change (MSC) is a form of outcome monitoring that focuses on collecting and selecting stakeholder accounts of significant changes that have occurred and deciding which is the most significant and why. The approach emphasizes the need to respect stakeholders' judgments regarding changes an initiative has made to their lives or in their community. Although the approach includes validation of the stories provided through triangulation with other data sources, it is an inductive, goal-free methodwith no pre-determined notion of what impacts ought to have been achieved. Rather. MSC elicits stories (both positive and negative) of what impacts stakeholdersexperienced from an initiative. The central element of the approach is the systematic collection and selection of a purposive sample of significant change stories. Stories are collected from people most directly involved with an intervention (such as participants, field staff, affected community members) by asking a simple, open-ended question: 'what was the most significant change that took place for your community' (in a particular domain, such as relationships among people, over a particular period of time)? The most significant of the stories are selected through a multi-layered group process of review, selection—often (although not always) involving participants and community stakeholders.

MSC is a useful methodology for complex situations because it bases its judgments and analysis on the personal experiences of individual local stakeholders, helping program implementers to understand in participants' own words how they view their environment and the change processes that take place within it. In a complex conflict context, MSC also can help to shed light onto which program areas are most relevant to peace. MSC is generally not intended to be used as a stand-alone methodology for evaluating impact, but should be combined with other methods to investigate the reported changes and the causal links between the intervention and those impacts more deductively. In particular, a meta-analysis can also be used to identify trends that are crosscutting among a number of stories that should be subject to more thorough investigation.

***Systems mapping*.** While not sufficient in itself, systems mapping can be an important step in the assessment process. The mapping techniques discussed by Danny Burns and Rob Ricigliano may be highly useful. Accordingly, both Burns' (2007) *Systemic Action Research: A Strategy for Whole System Change* and Ricigliano's (2012) *Making Peace Last: A Toolbox for Sustainable Peacebuilding* are helpful resources for the mapping process. For our focus, the mapping approach could help to identify the system boundaries and help reveal issues of scale, especially if we are cognizant of the scale issue (e.g. micro versus macro level interactions, processes) when engaging in any mapping activities. Mapping is further useful for identifying various forms of feedback within the system. Understanding system feedback can provide insight into specific attractor dynamics at work to stabilize present conditions.

Mapping tools have been used in several evaluations of impacts of peacebuilding, to assess relevance, to situate the intervention within an evolving context, to synthesize multiple perspectives on how a situation has changed and the role of the intervention in that change, and to test and develop more systemic theories of change that take account of feedback mechanisms and system pushback (Chigas & Woodrow 2014)

Ideas for moving forward

1. Additional tools and approaches.

Our briefing paper has focused on broad processes and approaches and touched on a few tools the group has used or been thinking about, but there are many more that merit further exploration in relation to where, how and for what insights they might be used in impact assessment. UNDP's 2013 Discussion Paper, *Innovations in Monitoring and Evaluating Results* outlines a few, including technology-based tools, such as: crowdsourcing, mobile data collection, participatory statistics, and micro-narratives (big data story collection and analysis through Sensemaker software), among others. Social network analysis also holds promise, along with the CDE framework developed by Glenda Eoyang and agent-based modeling, and there are likely many more.

2. ***Narratives.***

We think that a narratives approach could be a useful foundation for our present work. There are enough studies and approaches available that narratives could be examined on a very small group level (e.g. interviews, writing) to large-scale societal levels (e.g. surveys). Furthermore, there are ways to analyze modern social media to get at key terminology and themes shared by groups who interact via Twitter, Facebook, or other social media. Maybe a narratives approach could provide insight into the patterns of exchange across different social levels? This would be useful to identify cohesion and connections within communities, and ultimately identify whether meaningful change is emerging in a system.

In sum, for our purposes, narratives could be used in the following way:

- Examine narratives (both preexisting and ones gathered in-person) to identify the system boundaries and structure – feedback and connections
- Examine information in narratives to discern whether different elements in the system are framed as low-level or high-level
- A connected network of elements having consistently high-level frames likely reflects an attractor
- A connected network of elements having consistently low-level frames likely reflects an area ripe for change.

These ideas are tentative and require further discussion, exploration, and ultimately validation.

3. Improving system "fitness".

Several frameworks take an approach of improving elements of system "fitness"—strengthening systems and relationships that can generate solutions, rather than on just the solutions themselves. The *"Local systems" framework* (USAID 2014) focuses on analyzing "5 Rs": resources (transformation of resources into outputs, roles, relationships, rules and results. The framework helps to identify and track strengths and weaknesses in existing local systems, provide a guide to systems-strengthening interventions, as well as potentially provide a framework for tracking system health or fitness. *"Collective impact"* might also

be explored as a framework for understanding system fitness. The collective impact framework, developed by John Kania and Mark Kramer of FSG and Harvard's Kennedy School (Hanleybrown, Kania & Kramer 2012, Kania & Kramer 2011, posits several conditions for social progress and collective impact of efforts, listed below:

Preconditions	**Collective Impact**
Sense of urgency	Crisis creates breaking point to convince people a new approach is needed? Opportunity (e.g., funding) to entice people?
Influential champion	Commands respect to bring CEO-level leaders together and keep engaged. Convening power.
Financial resources	2-3 years and one anchor
Conditions	
Common agenda	Shared vision for change including common understanding of problem and a joint approach to solving it
Shared measurement	Common set of measures to monitor performance/track progress, learn what is/is not working; collecting data and measuring results consistently across participants ensures efforts remain aligned and promotes mutual accountability.
Mutually reinforcing activities	Activities differentiated, but undertaken in ways that support and are coordinated with actions of others
Continuous communication	Consistent and open communication, open information flows from top down and bottom up, leading to development of trust and working relationships and resilience of actors in the system adapting their practice to changing conditions
Backbone support	Separate organization with staff and a specific set of skills to serve as backbone for entire initiative: strategic direction, facilitating dialogue between partners, managing data collection and analysis, handling communications, community outreach, mobilizing funding

With some work and adaptation, these might also be explored as elements of system fitness, along with others DST members might suggest.

Bibliography: Non-linear Impact Assessment

General

Eoyang, G. & Berkas, T.H. (1999) Evaluating performance in a complex adaptive system (CAS). In Lissack, M. and H. Gunz (Eds). *Managing complexity in organizations: A view in many directions.* Westport, CT: Quorum Books.

Funnell, S. & Rogers, P. (2011) *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco: Wiley.

Mowles, C. (2014) "Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship." *Evaluation*. Vol. 20(2): 160-175.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. & Befani, B. (2012) *Broadening the range of designs and methods for impact evaluation: Report of a study commissioned by the Department for International Development*. London: DFID.

Westley, F., Zimmerman, B. & Patton, M.Q. (2007) *Getting to Maybe: How the World is Changed*. Toronto: Vintage Canada.

Williams, B. & Hummelbrunner, R. (2011) *Systems Concepts in Action: A Practitioner's Toolkit*. Stanford: Stanford University Press.

Williams, B. & Iman, I. Eds. (2007) *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: EdgePress/American Evaluation Association.

Process Frameworks

Britt, H. (2013) "Complexity-Aware Monitoring." Discussion Note, Version 2.0, Monitoring & Evaluation Series, USAID. Washington, D.C.: USAID.

Chigas, D., Church, M. & Corlazzoli, V. (2014) *Evaluating Impacts of Peacebuilding Interventions: Approaches and methods, challenges and considerations*. Practice Product, Conflict, Crime and Violence Reduction Initiative, DFID. London: DFID.

Gamble, J. (2008) *A Developmental Evaluation Primer*. Montreal: J.W. McConnell Family Foundation.

Hargreaves, M. (2010) "Evaluating System Change: A Planning Guide." Methods Brief. Cambridge, MA: Mathematical Policy Research, Inc.

Patton, M.Q. (2011) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.

-- (2006) "Evaluation for the Way We Work." *Non-Profit Quarterly*. Spring 2006.

Williams, B. & van't Hof, S. (2014) *Wicked solutions: A systems approach to complex problems: A workbook*. Self-published.

Tools

United Nations Development Program (2013) *Discussion Paper: Innovations in Monitoring and Evaluation*. New York: UNDP.

Systems mapping

Burns, D. (2007) *Systemic Action Research: A Strategy for Whole System Change*. Bristol: The Policy Press.

Chigas, D. & Woodrow, P. (2014) "Systems Thinking in Peacebuilding Evaluation: Applications in Ghana, Guinea-Bissau and Kosovo" In Andersen, O., Bull, B. & Kennedy-Chouane, M. *Evaluation Methodologies for Aid in Conflict*. London: Routledge.

Ricigliano, R. (2012) *Making Peace Last: A Toolbox for Sustainable Peacebuilding*. Boulder, CO: Paradigm Publishers.

Narratives and Action Identification Theory

Chiu, C. & Hong, Y. Cultural processes. In A. Kruglanski & E. T. Higgins (Eds.). *Social Psychology: Handbook of Basic Principles* (2nd ed., pp. 785-804). New York: Guilford.

Gottman, J. M. (1993). The roles of conflict engagement, escalation, and avoidance in marital interaction: A longitudinal view of five types of couples. *Journal of Consulting & Clinical Psychology, 61*, 6-15.

Gottman, J., Swanson, C., & Murray, J. (1999). The mathematics of marriage: Dynamic mathematical nonlinear modeling of newlywed marital interaction. *Journal of Family Psychology, 13*, 3-19.

Gottman, J. M., Murray, J. D., Swanson, C. C., Tyson, R., & Swanson, K. R. (2005). *The mathematics of marriage: Dynamic nonlinear models*. Cambridge, MA: MIT Press.

Gottman, J.M., Swanson, C., & Swanson, K. (2002). A general systems theory of marriage: Non-linear difference equation modeling of marital interaction. *Personality & Social Psychology Review ,6,* 326-340.

Isbell, L. M., Lair, E. C., & Rovenpor, D. R. (2013). Affect-as-information about processing styles: A Cognitive malleability approach. *Social and Personality Psychology Compass, 7*, 93-114.

Losada, M. (1999). The complex dynamics of high performance teams. *Mathematical and Computer Modeling, 30*, 179-192.

Michaels, J. L., Parkin, S. S., & Vallacher, R. R. (2013). Destiny is in the details: Action identification in the construction and deconstruction of meaning. In J. A. Hicks & C. Routledge (Eds.), *The Experience of Meaning in Life* (pp. 103-116). New York: Springer.

Nair, N. (2007) Towards understanding the role of emotions in conflict: A review and future directions. International Journal of Conflict Management, 19(4), 359-381.

Paloutzian, R. F. & Park, C. L. (2013). *Handbook of the psychology of religion and spirituality*. New York: Guilford.

Park, C. L. (2013). Religion and meaning. In R. F. Paloutzian & C. L. Park (Eds.), *Handbook of the Psychology of Religion and Spirituality* (2nd ed., pp. 357-379). New York: Guilford.

Vallacher, R. R. & Wegner, D. M. (1985). *A theory of action identification*. Lawrence Erlbaum Associates.

Vallacher, R. R. & Wegner, D. M. (1987). What do people think they're doing: Action identification and human behavior. *Psychology Review, 94*, 3-15.

Dynamic Indicators

Biggs, R., M. Schlüter, D. Biggs, E. L. Bohensky, S. BurnSilver, G. Cundill, V. Dakos, T. M. Daw, L. S. Evans, K. Kotschy, A. M. Leitch, C. Meek, A. Quinlan, C. Raudsepp-Hearne, M. D. Robards, M. L. Schoon, L. Schultz, and P. C. West (2012) "Toward Principles for Enhancing the Resilience of Ecosystem Services." *Annual Review of Environment and Resources* 37:421–448.

Chapin, F. S., III, G. P. Kofinas, C. Folke, S. R. Carpenter, P. Olsson, N. Abel, R. Biggs, R. L. Naylor, E. Pinkerton, D. M. S. Smith, W. Steffen, B. Walker, and O. R. Young (2009) "Resilience-Based Stewardship: Strategies for Navigating Sustainable Pathways in a Changing World." In *Principles of Ecosystem Stewardship*. New York: Springer. Pages 319–337.

Dellinger, M., L. Tofan, and T. Ehlinger (2013) "Predictive Analytics And Pattern Visualisation For Human Health Risk Assessment." *Journal of Environmental Protection and Ecology* 13:2186–2197.

Dellinger, M., M. Carvan, R. Klingler, J. McGraw, and T. Ehlinger (2014) "An Exploratory Analysis of Stream Teratogenicity and Human Health Using Zebrafish Whole-Sediment Toxicity Test." *Challenges* 5:75–97.

Estep, M. 2006. *Self-Organizing Natural Intelligence*. New York: Springer.

Gunderson, L. H. (2002) *Panarchy: Understanding Transformations In Human And Natural Systems* - Books on Google Play.

Kofinas, G. P., and C. Folke. 2009. *Principles of Ecosystem Stewardship*. New York: Springer.

Novotny, V., A. Bartosova, N. O'Reilly, and T. EHLINGER (2005) "Unlocking the relationship of biotic waters to anthropogenic integrity of impaired stresses." *Water research* 39:184–198.

Ostrom, E. (2009) "A General Framework for Analyzing Sustainability of Social-Ecological Systems" (supplemental Information). *Science* 325:419–422.

Wakefield, J. (2008) "Ecologic Studies Revisited." *Annual Review of Public Health* 29:75–90.

Most Significant Change

Davies, R. & Dart, J. (2005) *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*. http://www.mande.co.uk/docs/MSCGuide.pdf.

Most Significant Change. Description and resources at http://betterevaluation.org/plan/approach/most_significant_change.

Outcome mapping and Outcome harvesting

Earl, S., Carden, F. & Smutylo, T. (2001) *Outcome Mapping: Building learning and reflection into development programs*. Ottawa: IDRC. http://www.idrc.ca/EN/Resources/Publications/Pages/IDRCBookDetails.aspx?PublicationID=121.

Smutylo, T. (2005) *Outcome mapping: A method for tracking behavioural changes in development programs*. ILAC Brief 7. http://www.outcomemapping.ca/resource/resource.php?id=182.

Earl, S. (2008) Introduction to Outcome Mapping. Three videos providing an introduction to the concepts of outcome mapping. https://www.youtube.com/watch?v=fPL_KEUawnc.

Wilson-Grau, R. & Britt, H. (2013) *Outcome Harvesting* (revised). Cairo: Ford Foundation. http://www.outcomemapping.ca/resource/resource.php?id=374.

World Bank Institute (2014) *Cases in Outcome Harvesting: Ten pilot experiences identify new learning from multi-stakeholder projects to improve results*. Washington, D.C.: World Bank. http://wbi.worldbank.org/wbi/Data/wbi/wbicms/files/drupal-acquia/wbi/Cases%20in%20Outcome%20Harvesting.pdf.

Outcome mapping. Description and resources at http://betterevaluation.org/plan/approach/outcome_mapping.

System fitness

USAID (2014) *Local Systems: A Framework for Supporting Sustained Development*. Washington, D.C.: USAID.

Hanleybrown, F., Kania, J. & Kramer, M. (2011) "Channeling Change: Making Collective Impact Work." *Stanford Innovation Social Review Blog.* January 26, 2012.http://www.ssireview.org/blog/entry/channeling_change_making_collective_impact_work.

Kania, J. & Kramer, M. (2013) "Embracing Emergence." *Stanford Social Innovation Review Blog*. January 21, 2013. http://www.ssireview.org/blog/entry/embracing_emergence_how_collective_impact_addresses_complexity.

-- (2011) "Collective Impact." *Stanford Social Innovation Review*. Vol. 9, No. 1.