

Using the Cynefin framework in evaluation planning: A Case Example

Heather Britt

7 February 2011

Outline

1. In a Causal Rut: The Need for System Thinking in Evaluation Planning
2. The Cynefin Framework in Evaluation Planning: Outline of an Approach
3. Overview of the Cynefin Framework
4. Where Cynefin Sits in the Evaluation Planning Process
5. Three Steps to Using Cynefin in Evaluation Planning
6. Case Example
7. Conclusion

I. In a Causal Rut: The Need for System Thinking in Evaluation Planning

Not too long ago, a client came to me and said: “We’re rolling out a training program for our staff and partners around the globe and we want to see how effective the training is. We also want to know what supports the use of these new skills in the workplace, so we can advocate for appropriate changes in the organization. Can you help?”

My client wanted to measure results, but she also wanted to understand their causes. Bookshelves are filled with guidance on measuring training results; there’s much less on capturing the organizational and other factors influencing those results. Evaluators tend more to control for outside factors than to make them a focus of exploration and inquiry. Proving a relationship between training and results can be very useful, but it doesn’t cover the universe of information needs and evaluation purposes.

Of course, evaluators have developed a number of ways of addressing causality. But as practitioners, we often get into a narrow rut -- viewing, describing and measuring causality in a single way. In doing so, we blind ourselves to the variety of dynamics possible in a situation.

As evaluators, we need to become more adept at thinking about multiple and interactive causes, describing and measuring them. Currently, evaluators lack tools that:

- Distinguish between the kinds of causality that may operate simultaneously in a situation;
- Support choices among measurement methods to capture different dynamics;
- Communicate those distinctions and choices effectively to clients.

If we are to answer client questions effectively, we must open our eyes to a variety of causal dynamics in evaluation situations. We need to incorporate those dynamics in feasible, affordable evaluations. And we need to communicate our thinking in everyday language that makes intuitive sense to our clients and evaluation stakeholders.

This paper represents a step along the way. It presents my experience in using a systems-thinking tool to plan an evaluation with a client.

II. The Cynefin Framework in Evaluation Planning: Outline of an Approach

Williams, Gujit and Rogers, in their presentation at a recent international conference on impact evaluation¹, suggest that the Cynefin framework can assist evaluators in identifying various causal dynamics. The Cynefin framework proposes that situations contain aspects that can be described as simple (known), complicated (knowable), complex or chaotic. Each of the four zones is distinguished by a different kind of causal relationship, a different dynamic. Using Cynefin to view a situation helps evaluators to see various types of causality, and counteracts tendencies to represent an evaluation situation as *either* a linear logic model *or* a complex adaptive system.

The framework also includes strategies for approaching and understanding causal dynamics in each zone. Once evaluators have mapped situation aspects into different dynamic zones, Cynefin provides guidance on measurement methods suited to capturing the dynamics of each zone. The framework also helps to communicate the reasoning behind those choices better with clients and to provide a basis to engage them in evaluation planning.

The implications for evaluation are this: Different dynamics in a situation require different evaluation approaches to discover and measure what is happening and why it is happening. In simple zone where cause-effect relationships are known, useful evaluation practice may be limited to tasks such as monitoring against targets and validating best practices. Complicated aspects are better addressed by linear logic modeling tools and standard social science research methods. Complex aspects require evaluation approaches that allow for learning in emergent situations (i.e., those in which outcomes cannot be fully predicted). Chaotic components are not well-suited for most evaluation approaches, as stakeholders are focusing on survival in crisis.

III. Overview of the Cynefin framework

Cynefin (pronounced /'kʌnɪvɪn/) is a Welsh word that literally means “habitat” or “place.” The fuller, richer meaning of the word is “place of your multiple belongings.”²

Wikipedia defines Cynefin as “a model used to describe problems, situations and systems. The model provides a taxonomy that guides what sort of explanations and/or solutions may apply.”³ David Snowden and Cynthia Kurtz, the framework’s creators, describe it as a device which helps people make sense of the complexities made visible by relaxation of assumptions about order, rational choice and intent⁴. In other words, when we give up insisting that all phenomena fit a rational model, we can see different organizing dynamics at work.

¹ “Thinking Systematically about Impact Evaluation of Programs and Policies with Simple, Complicated and Complex Aspects,” presentation by Bob Williams, Irene Gujit, and Patricia Rogers at the conference Perspectives on Impact Evaluation: Approaches to Assessing Development Effectiveness (1 April 2009), Cairo, Egypt.

² David Snowden, <http://www.youtube.com/watch?v=N7oz366X0-8>

³ <http://en.wikipedia.org/wiki/Cynefin>

⁴ Kurtz, C. F. and D. J. Snowden. 2003. “The new dynamics of strategy: Sense-making in a complex and complicated world.” IBM Systems Journal, vol 42, number 3, page 462.

In Williams view, the important thing about Cynefin is that it provides an alternative to approaching everything as if it were simple, or conversely treating everything as if it were complex⁵.

Evaluators using the Cynefin framework may assign aspects of situations to one of four different types or zones, based on their causal dynamics. In the **simple** zone, causal dynamics are well known. The right answer is common knowledge. Best practices have been identified.

Complicated aspects of a situation require time and energy to understand and measure. Experts would be expected to possess the relevant knowledge, and to be able to identify effective practices.

Complex aspects of a situation are ones that cannot be known or predicted ahead of time; patterns and cause-effect relationships emerge only retrospectively. **Chaotic** aspects of a situation are ones that are characterized by volatility, uncertainty and disagreement. There are no right answers. It's better to nudge the situation towards stability.

A situation may contain elements that demonstrate behaviors and relationships associated with one or all of the four dynamic types. And over time, the nature of these elements may shift from one category of the framework to another. Labeling an entire system as a single dynamic may oversimplify the situation and omit information critical to sound evaluation planning.

For each dynamic zone, the framework suggests a pattern of productive action and inquiry called sense-making. Each sense-making strategy is comprised of a "suite" of three individual actions which can be used to gather information, make assessments and support decisions – the core of evaluation.

The tables below describe the six types of sense-making, and their unique combinations relevant for each dynamic zone.

Types of Sense-Making⁶

Sense	Collect sufficient data to identify the characteristics of this aspect of a situation
Categorize	Identify where these characteristics fit within known world
Analyze	Rely on expert opinion and diverse stakeholder perspectives in order to identify cause-effect relationships & select appropriate response
Respond	Carry out the practice that has been proven most appropriate to that category (e.g. best, good, emergent or novel practice)
Probe	An experiment that makes patterns more visible and knowable by sensing.
Act	A strong intervention designed to shock a chaotic aspect of the situation back into some form of order

Sense-Making Strategies for Each Zone

Zone		Sense-Making Strategies
Simple/Known	<i>Everyone knows the right answer</i>	Sense, Categorize, Respond
Complicated/ Knowable	<i>What we need to spend time & energy finding out; an expert would know</i>	Sense, Analyze, Respond
Complex	<i>What we can pattern retrospectively</i>	Probe, Sense, Respond
Chaotic	<i>What we need to stabilize for patterns to emerge; there is no right answer</i>	Act, Sense, Respond

⁵ Williams, Bob. 2010. "Cynefin" in *Systems Concepts in Action* by Bob Williams and Richard Hummelbrunner. Palo Alto: Stanford University Press.

⁶ Information in this table is drawn from Kurtz & Snowden (2003) and Williams (2010).

The framework is frequently represented by the following diagram⁷.



Figure 1: Cynefin Framework

IV. Where Cynefin Sits in the Evaluation Planning Process

Evaluation planning includes four basic steps:

1. Identify evaluation users & purposes
2. Draft key evaluation questions
3. Select appropriate evaluation methods for each evaluation question
4. Craft coherent design and evaluation implementation plan

This case example describes an instrumental use of the framework to select evaluation methods and approaches for answering evaluation questions. Used in this way, Cynefin is introduced after evaluation users and purposes had been identified and the key evaluation questions had been defined. Key evaluation questions are derived directly from the evaluation purpose, and purpose and users inform subsequent methodological choices. The selected methods are then integrated into a coherent evaluation design.

In this case example, the client and I worked together to identify primary evaluation users and purposes and draft key evaluation questions. I then considered the causal dynamics related to each key evaluation question and mapped them on the Cynefin framework. I referred to the sense-making strategies for each zone in selecting methods for each evaluation question. Throughout the process, I used Cynefin as an aid to communicate and refine the evaluation design and evaluation questions in collaboration with my client. Finally, I integrated the selected methods into a coherent evaluation design.

This approach to evaluation planning with the Cynefin framework uses evaluation questions as the “unit of analysis.” Evaluation questions -- rather than some other category -- were used to break down the evaluation situation into elements for analysis. For each evaluation question, we considered the relevant aspects of the situation and mapped them on the framework, and selected methods. Evaluation questions were determined to be the appropriate “unit of analysis” for two reasons. The most compelling reason is the fit with evaluation planning tasks. Evaluators engaged in design approach each key evaluation question to determine the best way to gather information. Mapping evaluation questions ensures methods fit to the evaluation question. A second reason for

⁷ <http://en.wikipedia.org/wiki/Cynefin>

using evaluation questions as the “unit of analysis,” concerns “scale”. Mapping evaluation questions provides the detail necessary to assist with evaluation design.

Experienced systems thinking practitioners tell us that sensitizing devices like Cynefin inform their overall approach to an evaluation, and can be used at each step of the evaluation planning process (Hargreaves, 2010). For example, Cynefin is one of many tools and exercises that might be used to early in an evaluation process to articulate the evaluation purpose and shape questions. Evaluation situations are complex and composed of interlocking systems and subsystems. To develop useful evaluation questions and facilitate inquiry, evaluators use different approaches to set boundaries around what will be examined and what will be excluded. Systems thinkers often use core concepts such as boundaries, relationships and perspectives to accomplish this task.

V. Three Steps to Using Cynefin in Evaluation Planning

There are three main steps in using Cynefin to select methods for an evaluation design.

1. Assign each evaluation question to the right zone of the framework
2. Select evaluation methods to answer each question
3. Integrate the key questions and methods into a coherent evaluation design

Step 1: Assign each evaluation question to the right Cynefin zone

In this step, evaluation questions are analyzed using the Cynefin framework and assigned to one of the four domains. The evaluator may undertake this task, or s/he may broaden the discussion to include various stakeholders. Those with experience facilitating exercises with Cynefin for groups have valuable guidance to offer about ways to engage stakeholders productively in this process.

Let me stress: This is an initial determination. Things may change, and usually do. Early in the evaluation design stage our knowledge of the situation may be quite limited. Over time, elements may shift from one zone of the framework to another. In some cases, the assignment of an evaluation question is clear. In others it is an educated guess that is worth re-visiting as the evaluation progresses and more information is available.

Two questions aid in assigning evaluation questions to the domains of the Cynefin framework:

1. What is the nature of the relationship between elements within this situation?
2. How can we know the answer to this evaluation question?

Cynefin Question 1: What is the nature and strength of the relationship between elements within this situation?

One way to identify the Cynefin domains is through the relationship between elements in the situation under consideration. Each evaluation question defines an aspect of the situation that contains a number of elements. Snowden & Kurtz refer to the relationships between elements as “connection strengths”⁸. Strong connections between elements restrict freedom of movement, and increase stability in a domain. Weak or absent connections allow for more freedom of movement between elements and decrease stability. The nature and strength of connections are of central

⁸ Ibid, p.470

importance to evaluation questions addressing causality. Another way of asking the question is “What causal relations are operating in relation to this evaluation question?”

What do we know about the relationship between these elements related to the evaluation question? To answer this question, examine the evidence about the nature and strength of connections, and the type of causality at play in the situation. The nature and strength of connections operating in relation to the evaluation question indicate the Cynefin zone in which the evaluation question belongs.

Connection Strengths (Cynefin Framework)⁹

Question 2: What is the nature of the relationship between elements within this situation?	Cynefin Domain
<i>Weak connections between elements, but a strong link to a central control element.</i>	Simple/Known
<i>Strong connections between elements, but each element is still has strong links to a central controlling element.</i>	Complicated/ Knowable
<i>Strong connections between elements, but no central controlling element.</i>	Complex
<i>Weak connections between elements, and no central organizing core.</i>	Chaotic

Cynefin Question 2: How can we know the answer to this evaluation question?

Cynefin offers a second lens with which to view the evaluation question – the sense-making strategies. Now, we consider how information related to the evaluation question is best gathered and understood. Is it a situation in which everyone knows the right answer? Or will we need to consult an expert? Can we make valid predictions or is it best to pattern retrospectively? Is the situation so unstable that it’s not possible to answer the question? For each evaluation question, make a determination of how best to answer the question, and select the corresponding framework zone.

Question 1: How can we know the answer to this evaluation question?	Cynefin Domain & Sense-Making Strategy
<i>Everyone knows the right answer</i>	Simple/Known Sense, Categorize, Respond
<i>What we need to spend time & energy finding out; an expert would know</i>	Complicated/Knowable Sense, Analyze, Respond
<i>What we can pattern retrospectively</i>	Complex Probe, Sense, Respond
<i>What we need to stabilize for patterns to emerge; there is no right answer</i>	Chaotic Act, Sense, Respond

Examine each evaluation question using both connection strengths and sense-making strategies in order to assign it to the proper framework zone. Does considering the evaluation question through both lenses of the framework yield the same result? To make the final determination of where a specific evaluation question sits, compare the answers to 1) connection strengths, and 2) sense-making; then resolve any discrepancies.

There is a tendency for evaluation questions to fall into the two middle domains – complicated and complex. Simple questions may be appropriate for monitoring -- the collection, compilation and

⁹ Williams, Bob and Richard Hummelbrunner. 2010. "Cynefin" in *Systems Concepts in Action*. Palo Alto: Stanford University Press

analysis of “known” answers. Answering such questions is not to be overlooked, as often they support inquiry into complicated and complex questions. Questions that fall within the chaotic may be beyond the managerial reach and resources associated with most evaluations and M&E systems. The atmosphere of sink-or-swim is not conducive to inquiry.

Step 2: Select appropriate evaluation methods for each evaluation question

Decisions on methods are influenced by the limits of a practitioner’s knowledge and the current state of the evaluation field. The Cynefin framework can help refine the practitioner’s judgment and sharpen her awareness of her knowledge limits, and where they may need to be stretched.

In this step, the evaluator assigns appropriate methods to each evaluation question. The evaluator asks “How do we know?” again in more depth, this time using the sense-making strategies of each zone as a guide in selecting methods and approaches. In the simple zone, consider methods that involve assigning data to known categories (sense-categorize). In the complicated zone, methods that rely on expert knowledge to analyze information (sense-analyze) are most appropriate. In the complex zone, select methods that use retrospective patterning (probe-sense). The type of connection strengths related to the evaluation question may serve as a second lens in considering possible methods. Does the method under consideration operate well in the causal dynamics operating in relation to the evaluation question? Below is a selection of possible evaluation approaches and methods for each Cynefin zone¹⁰. This list is not intended to be exhaustive.

Cynefin Domain & Sense-Making Strategy		Appropriate Evaluation Approaches & Methods
Simple/Known Sense, Categorize, Respond	<i>Everyone knows the right answer</i>	Monitor against targets (quantitative and qualitative); collect self-evident facts or common knowledge; validate best practice; collect & share stakeholder views
Complicated/Knowable Sense, Analyze, Respond	<i>What we need to spend time & energy finding out; an expert would know</i>	Program-theory driven evaluation Standard social science methods, such as experimental & quasi-experimental designs.
Complex Probe, Sense, Respond	<i>What we can pattern retrospectively</i>	Experimental designs Facilitative & exploratory approaches to collect data on results, factors & relationships. Most Significant Change, Success Case Method, CE narrative methods IDRC’s Outcome Mapping, Checkland’s Rich Picturing Network analysis
Chaotic Act, Sense, Respond	<i>What we need to stabilize for patterns to emerge; there is no right answer</i>	Not a domain appropriate for most evaluation or research; act to nudge the situation towards simple or complex.

This step is deceptively simple. While theoretically it should be possible to identify evaluation methods which most closely match those means of sense-making, in practice identifying

¹⁰ Bob Williams, personal correspondence, September 2009.

measurement approaches that suit the dynamics is less than straightforward. Several evaluators working in the frontier territory between systems thinking and evaluation (Williams, Patton, and Hargreaves) have been addressing the question of which approaches and methods are most useful in the various domains. At this point, the discussion is evolving, and thinkers have not reached consensus on this topic.

Step 3: Integrate the selected methods into a coherent evaluation design

A set of evaluation questions and intended methods to answer them is a good start, but it does not constitute an evaluation design anymore than four tires, an engine and a steering wheel make a car. To get our evaluation car on the road, we have to integrate the various questions and methods into a coherent design that will provide accurate and high-quality information that satisfies the overall evaluation purpose, and the specific needs of various users, in an efficient, feasible and ethical manner. Final selection of evaluation methods should also take into account a number of relevant factors such as evaluation purpose, intended audience, organizational culture, and available resources.

A fully developed design will also require details on practical methods for sampling, data collection, data analysis, interpretation, and reporting. If the evaluation incorporates mixed methods, it will be necessary to explain how these methods complement each other, rather than duplicate or conflict with one another.

VI. Case Example

This case example describes work I completed as a consultant under contract for a large international development agency. The agency is undertaking efforts to mainstream gender¹¹ in its planning and operations. Training is one of the strategies being used to achieve gender mainstreaming. I was approached by a member of the unit responsible for developing, implementing and disseminating gender mainstreaming tools aimed at building capacity for gender analysis and planning for gender responsive programs.

My contact tasked me to design an evaluation framework to capture the outcomes of the current training programs, including the conditions and capacity for staff transfer of gender analysis skills to their daily work, and the impact of the training on organizational performance related to gender mainstreaming. This framework would serve as the basis for a system of on-going data collection, analysis, evaluation and learning.

I had limited access to the client, and conducted work on this consultancy off-site. I communicated primarily with a single representative and had minimal interaction with other key stakeholders in the agency. My contact consulted with colleagues and communicated feedback to me on their behalf. These circumstances are less than ideal for evaluation planning, but not uncommon in the

¹¹ "Mainstreaming a gender perspective is the process of assessing the implications for women and men of any planned action, including legislation, policies or programs, in any area and at all levels. It is a strategy for making the concerns and experiences of women as well as of men an integral part of the design, implementation, monitoring and evaluation of policies and programs in all political, economic and societal spheres, so that women and men benefit equally, and inequality is not perpetuated. The ultimate goal of mainstreaming is to achieve gender equality." United Nations. "[Report of the Economic and Social Council for 1997](http://www.un.org/documents/ga/docs/52/plenary/a52-3.htm)". A/52/3.18 September 1997. <http://www.un.org/documents/ga/docs/52/plenary/a52-3.htm>

arena of international development in which key stakeholders are spread around the globe, and evaluation funds are often tight and conserved for data collection.

Determining Users & Purposes of the Evaluation

If the circumstances had permitted, I would have preferred a more in-depth, participatory, and inclusive means of identifying stakeholders, stakeholder values, and purposes of the proposed evaluation system. Working within the limitations of the situation, I interviewed the client by telephone and collected further information by questionnaire and on-going correspondence. The information served to clarify the context, purpose and intended uses of the framework and the evaluation system to be developed based on that framework.

The intended users of the evaluation were identified as:

1. Unit charged with design and implementation of gender mainstreaming training
2. Unit supporting training and staff learning across the organization
3. Senior management across the organization mandated to implement gender mainstreaming

The agency has instituted evaluation mechanisms to measure progress towards its comprehensive gender mainstreaming goals. Findings from the proposed evaluation would provide a secondary source of information specific to gender analysis and planning. Thus, the overall purpose of the proposed evaluation framework was defined as determining whether and how training was contributing to gender mainstreaming in the organization.

Two specific, action-oriented objectives were identified for the evaluation:

1. demonstrate the effects of the training on changes in knowledge, attitudes and skills; and
2. identify individual, group and institutional successes and challenges in the transfer of gender analysis and planning skills to practice.

If the evaluation were to identify gaps in learning, the developers of the training hoped to improve the materials or methods. The client also planned to use evaluation findings regarding organizational factors influencing practice to close the gap between knowledge and implementation (the “know-do” gap) in the area of gender analysis and planning. They planned to lobby to enhance factors that facilitate practice and address inhibiting ones.

Drafting Key Evaluation Questions

Three key evaluation questions were drawn directly from the purposes identified.

Key Evaluation Questions

Q1. Does the training result in increased knowledge, skills and desired attitudes?

Q2. Are trainees practicing knowledge & skills taught in the training in their work settings?

Q3. What factors promote/inhibit the practice of skills in trainee work settings?

This phase of evaluation planning also included a review of the literature on monitoring and evaluation (M&E) of gender training to identify any existing evaluation frameworks that suited our purpose. We included the literature review at this stage because evaluation frameworks influence not only the choice of measurement methods, but also the types of questions that are asked. The client and I wanted to be fully informed of current theory and practice in monitoring and evaluation of gender training when we made our final selection of key evaluation questions.

Although modest in scope and bounded closely by its purpose, the review found little in the way of monitoring and evaluation of gender training. Specifically:

1. Limited attention has been paid to monitoring and evaluation of gender training.
2. The little monitoring and evaluation of gender training that has been undertaken covers trainee satisfaction, and does not address outcomes such as increased knowledge and awareness, improved skills, changed behaviors or higher level organizational changes.
3. There is no evidence of M&E systems which provide for the regular, on-going collection, analysis and use of M&E data to inform the design and implementation of gender training programs.

Because the review found few examples of evaluation or monitoring of gender training, the report was broadened to include capacity-building interventions more generally in international development. The expanded literature review found that the Kirkpatrick¹² framework is the most popular approach to the evaluation of capacity building.

I assessed the Kirkpatrick model (and several of its adaptations) in light of its ability to answer my client's questions. Kirkpatrick focuses attention on the training as the primary causal agent. This seemed appropriate for measuring learning, that is, in addressing our first key evaluation question about changes in knowledge, skills and attitudes resulting directly from training. But Kirkpatrick does not promote consideration of the other factors involved in translating knowledge and skills gained from training into organizational performance. Thus, it would be unsuitable for answering our third key evaluation question examining factors that promote or inhibit the practice of skills in trainee work settings.

The client and I discussed the findings of the literature review and our three key questions. We were becoming increasingly aware that each question involved different dynamics and would require different evaluation approaches. On the one hand, the links between training and learning were pretty clear, and popular measurement approaches might be used. However, Kirkpatrick would be inadequate to explain the complexities of translating training into practice. The client knew that while training might account for some changes in performance on the job, there are a number of other factors that influence performance on the job. Because the work settings of staff receiving the training are spread around the globe and vary considerably, there is no way to predict ahead of time the most salient factors influencing application of training knowledge and skills. In fact, identification of those factors is just what she wanted to get at.

The evaluation situation included components of various levels of complexity and predictability. I proposed that we use Cynefin to help us view the situation in a new light, and to design a framework that addressed all three questions. The client agreed.

The literature review and analysis of client information needs generated a number of key evaluation questions in addition to the three cited in this paper. The client and I reviewed the complete list and made a final selection of key evaluation questions taking into consideration the primary evaluation purpose, intended uses, available resources and organizational context.

¹² Kirkpatrick's framework includes four levels for evaluation of training programs: reaction (level 1), learning (level 2), application (level 3), and impact (level 4).

Introducing Cynefin to Clients

Introducing systems thinking approaches and tools to evaluation users should be undertaken thoughtfully. As evaluators, we may get carried away by recent breakthroughs in evaluation theory and practice. But for many evaluation users and stakeholders, evaluation is a mysterious and anxiety-producing activity¹. In addition, systems thinking concepts, approaches and tools have yet to gain widespread recognition in many professional arenas. The introduction of “strange” new terms and methods can further increase anxiety and undermine processes intended to increase evaluation use.

There any number of strategies that an evaluator may use to introduce systems thinking approaches and tools to evaluation stakeholders. Most of these are familiar to any practicing evaluator familiar with the educational and coaching aspects of our roles. For example, we should minimize technical terminology or avoid it altogether. It may not be necessary to name these tools as “systems thinking,” or fully explain their theoretical underpinnings.

No matter how sensitively we address the introduction of developments in evaluation, evaluation users and clients may reject them. In that case, we have several choices: 1) honor the client’s wish and carry out the evaluation relying on different approaches and tools, 2) withdraw from the evaluation if we feel the evaluation is irreparably compromised, or 3) use systems thinking as a sensitizing device for our own reflection and problem-solving, but abstain from using specific elements as requested. The last option may allow us to inform the evaluation or educate stakeholders about systems thinking through various means (such as the gradual introduction of new concepts and experiential learning), but we must be careful not to misrepresent our process.

Step 1: Assign each evaluation question to one of the four domains

I assigned the evaluation questions based on information I had gathered about the organizational context. This was an iterative process of deepening understanding of both the organizational context and the Cynefin framework. Let’s look at the process of assigning the three key evaluation questions as an illustration.

Q1. Does the training result in increased knowledge, skills and desired attitudes?

First, I asked: **What causal relations are operating here?** Managers and trainers within the organization have defined a specific set of skills and knowledge and designed training to teach them those to a large number of staff working in diverse organizational settings around the globe. The link between the training and the learning is intentional and clear. The training is the strong central controlling element. In relation to what they learn from the training, the participants have weak connections to one another. **Simple** causal relations are in operation here.

Then, I turned to the second question: **How can we know the answer to this question?** Which of the following options best represents the evaluation question: 1) the right answer is known, 2) expertise and effort are required, 3) only retrospective patterning is possible, or 4) there is no right answer. In a sense, the answer is “known” because each trainee knows whether she or he has learned the information taught in the training. It’s only a matter of asking each trainee. This confirms my original assessment that this evaluation question lies in the simple zone.

Q2. Are trainees practicing knowledge & skills taught in the training in their work settings?

First, I ask: **What causal relations are operating here?** What is the nature of the relationship between elements within this situation? Staff complete the training and return to their work settings around the globe. The distance between the staff & training grows wider and over time and distance, but still remains an important influence. The individual work settings also a strong effect on whether and how staff apply their new knowledge and skills. The elements relevant to this evaluation question demonstrate strong connections between elements (each trainee and elements of their work setting), while maintaining links to the central controlling element (the training). Therefore, this evaluation question is in the **Complicated/Knowable zone**.

Then, I turned to the second question: **How can we know the answer to this question?** Which of the following options best represents the evaluation question: 1) the right answer is known, 2) expertise and effort are required, 3) only retrospective patterning is possible, or 4) there is no right answer. The training taught a specific set of skills and we want to know whether those are being applied in their work. We are looking for a finite set skills, but the increased influence of the diverse work settings may result in variations in practice that are not immediately recognizable. Expertise and effort are required to answer this question. I'll stick with my complicated/knowable assessment.

Q3. What factors promote/inhibit the practice of skills in trainee work settings?

What causal relations are operating here? What is the nature of the relationship between elements within this situation? This evaluation question focuses on the interrelations between the trainees and their work settings as they attempt to apply skills. In this situation, the training serves to define the desired skill set, but its influence on behavior is not the focus of the evaluation question. Instead, this evaluation question concerns the relations between trainees and her/his work setting. Presumably, at least some of the trainees will try to apply the new knowledge and skills in their unit or department. How will the unit respond? What, if any, changes will occur in current procedures and practices? How will this affect interactions between colleagues? Cause and effect are intertwined, and the interaction of factors in workplace cannot be predicted. Another level of complexity is the diversity of work settings in this global organization. Offices, units and departments vary considerably across work functions, countries, cultures and languages to name just a few factors of diversity. The situation relevant to this evaluation question may be characterized by strong connections between elements (each trainee and elements of their work setting), but no central controlling element; this evaluation question falls in the **Complex** zone.

How can we know the answer to this question? Which of the following options best represents the evaluation question: 1) the right answer is known, 2) expertise and effort are required, 3) only retrospective patterning is possible, or 4) there is no right answer. The organization did not have a shared consensus on workplace factors influencing training as part of its common knowledge. Based on experience in the organization and our review of the literature, my client and I named a number of factors we thought likely to influencing the application of skills in the workplace. However, unlike the previous question concerning increased trainee knowledge and skills, we cannot predict all the answers that inquiry into organizational factors influencing use of skills would elicit. This eliminated expertise as the best path to answer this evaluation question. If we were to interview the trainees, their answers were likely to vary considerably and to reflect the diversity of organizational contexts around the globe. I did surmise that a pattern would emerge from analysis of responses from trainees (retrospective patterning), and this confirmed my original assessment that this evaluation question is complex.

Step 2: Select appropriate evaluation methods for each evaluation question

In this step, I referred to Cynefin's sense-making strategies for guidance on selecting methods for each key evaluation question. Each sense-making strategy is a suite of three actions which in combination make up the strategy. For each suite, the third action is "respond" which refers to the frameworks' emphasis on decision-making for action. In selecting the evaluation methods, gave priority to consideration of the first two actions.

Q1. Does the training result in increased knowledge, skills and desired attitudes?

This evaluation question is assigned to the Simple zone and the corresponding sense-making strategy is **Sense-Categorize-respond**. The desired training effects have been predetermined and represent the categories specific knowledge, skills and attitudes. "Sense" involves collecting

sufficient information to determine the presence or absence of each category. Methods using closed questions would be appropriate here. I select a pre/post test of the trainees.

Analysis of the simple connection strengths confirms this methods choice. The link between the cause (training) and effects (knowledge, skill and attitudes) is clear and the design should involve elimination of confounding factors at a basic level.

Q2. Are trainees practicing knowledge & skills taught in the training in their work settings?

This evaluation question is assigned to the Complicated/Knowable zone and the corresponding sense-making strategy is **Sense-Analyze-respond**. The influence of the training still allows us to predict some likely effects, but practice outside the carefully controlled training environment may also take on new and unpredictable forms. The boundaries between training-influenced practice and behaviors not related to training may not always be entirely clear. The appropriate measurement approach should allow for a degree of variation in desired practices both in terms of data collection and analysis. Review of the data will need to go beyond recognition of known categories; analysis will be required to determine whether reported behaviors demonstrate application of training. I chose to survey all trainees using both closed and open questions.

The connection strengths in the Complicated/Knowable zone are amendable to logic modeling to show how trainees interact with elements in the workplace in applying learning to their work.

Q3. What factors promote/inhibit the practice of skills in trainee work settings?

This evaluation question is assigned to the Complex zone and the corresponding sense-making strategy is **Probe-Sense-respond**. Retrospective patterning is the appropriate measurement approach for this zone. Making meaning of the data (sensing) follows collection of information that is not pre-categorized (probing). I chose to conduct interviews of success and non-success cases of trainees practicing skills in the workplace along the lines of the Success Case Method (SCM).

SCM was developed from Brinkerhoff's reformulation of the popular Kirkpatrick model for measuring training results. SCM is distinctive because it does not try to isolate the effect of training, but rather seeks to provide the various factors influencing application of training-related skills. SCM places training within a performance management system which includes the multitude of additional factors effecting performance.

I adapted the Success Case Method (SCM) to allow for both respondent-defined and pre-determined definitions of successful application of skills and performance¹³. Application of training can include specific attitudes, knowledge and skills. Success is defined as use of the training in a way that makes a significant difference in the trainees' work setting, usually in their unit,

¹³ SCM has a somewhat ambiguous stance towards pre-determining specific desired applications of training. The method uses a logic model to map desired applications and the impact of those applications for the organization as a whole. Yet, it acknowledges that such a model cannot capture the complexity of the situation:

"During the SCM inquiry, we may probe for and discover many applications of learning beyond what we listed on the impact model. This is well and good, and in fact is an aim of any SCM effort. But, to know where to begin to look and what to search for, the initial impact model is an indispensable guide."
(Brinkerhoff, 2010 p. 76)

The first step in the method – a survey of closed question items of all trainees – is constructed around the key behaviors defined by the model (Brinkerhoff, 2010 p. 92), therefore greatly restricting access to information regarding applications of training outside those narrow definitions.

department, or office. Respondents determine what a “significant difference” looks like in their specific context.

Specifically, I proposed to use interviews to collect descriptive reports from trainees on: a) the nature of the application of learning, b) the organizational impact of the application of learning, and c) the performance context factors that enabled/inhibited trainees to achieve results. The data will be analyzed using qualitative methods.

This method adaptation of SCM corresponds well to the Probe-Sense sense-making strategy recommended by Cynefin. The method does not attempt to predict the various factors that influence application of skills prior to data collection; these are determined through analysis of qualitative data. The method is also well-suited to capture complex casual relationships between performance factors and results for the organization.

Step 3: Integrating the key questions and methods into a coherent evaluation design

To develop a coherent evaluation design, I compiled the key evaluation questions and recommended methods into a single framework. Methods no longer stand-alone, but must work together in an integrated fashion. For example, the survey to answer evaluation question Q2, is now integrated with the interviews to answer Q3 according to SCM.

The framework also outlined measurement issues, as well as guidance on sampling, analysis, and data collection tools. As the client progresses in its plans for implementation of the training, the framework will be amended to develop a full evaluation plan. The table below summarizes the Cynefin category and suggested evaluation methods for each proposed evaluation question.

Proposed Evaluation Questions & Methods

Key Evaluation Questions	Suggested Evaluation Methods
Q1. Does the training result in increased knowledge, skills and desired attitudes?	Pre/post test (during training)
Q2. Are trainees practicing knowledge & skills taught in the training in their work settings?	Survey (a few months later – preliminary step in SCM)
Q3. What factors promote/inhibit the practice of skills in trainee work settings?	Adaptation of Success Case Method

VII. Conclusion

Despite the rise in popularity of systems thinking and increased awareness of the possibilities for its application to evaluation, the number of examples of the application of systems thinking tools in evaluation practice remains low. Interested evaluators will find very little in the way of guidance for using such concepts in everyday evaluation practice. I put forward this case example to encourage other early-adapters and to enrich the discussion concerning application of systems thinking in evaluation practice.

The compelling reason for using Cynefin or other systems thinking tools is that they improve our practice. This case documents three important benefits to evaluators:

1. Seeing various causal dynamics operating in a single evaluation situation;

2. Selecting approaches and methods appropriate to situation dynamics;
3. Communicating with stakeholders and engaging them in the evaluation design process.

Seeing causal dynamics at work

My evaluation practice had been strongly influenced by the program-theory driven approach. Drafting a logic model of the program theory was useful in suggesting evaluation questions related to learning gains and some changes in practice. However, logic models were less useful in guiding inquiry concerning the dynamics of changed practice across diverse work settings. I first tried to list possible factors influencing application of new skills and knowledge in the workplace, as well as the broader organization changes arising from the use of those new skills and knowledge. The client and I quickly deduced that we would not be able to predict all factors and their dynamics in offices across the globe. Moreover, since the organization is implementing a number of interventions to support gender mainstreaming, attributing changes to training alone would be reductive and inaccurate.

Critiques of linear modeling were not sufficient to release me from its influence. Despite our conscious recognition that the changes we were interested in went far beyond the training, Kirkpatrick-based models and program theory approaches to evaluation of training focused on the training as the primary factor influencing change.

Cynefin offered a paradigm shift and the tools to go with it. The framework's immediate appeal was its ability to distinguish between the simple, complicated and complex aspects of the situation. Early in the process, initial drafts of evaluation questions outlined broad areas of interest. Using the framework, I could see that different causal dynamics were operating in each of the areas of interest.

In early drafts of our evaluation framework, the following question featured prominently: *What have been the impacts for the organization as a result of training?* After viewing the situation through the lens of the Cynefin framework, I revised the evaluation question to: *What have been the impacts for the organization as a result of use of gender mainstreaming and analysis skills?* The second wording acknowledges that many other factors, in addition to training, influence the application of skills in the workplace.

Selecting approaches and methods

The second important benefit of using the Cynefin framework in evaluation planning is that it facilitates identification of appropriate approaches and methods. For the aspects of the situation that could be pre-determined (simple and complicated), logic modeling tools were helpful. For those aspects of the situation where the dynamics of the situation were entwined and emergent, the first step was to recognize this as an accurate description of that aspect of the situation, rather than a failure to accurately predict. Once the general evaluation approach was apparent, I could go on to recommend appropriate methods. For example, having embraced the complexity related to the organizational factors influencing application of new skill, I was now ready to propose evaluation methods that based on retrospective patterning.

Could I have designed the evaluation framework without the use of systems thinking tools? My design relies on Brinkerhoff's Success Case Method, so clearly there are other roads to this same destination. My argument for using Cynefin isn't that the resulting evaluation design is unique, but that it is an efficient means of selecting approaches and methods appropriate to the dynamics of each evaluation question.

Communicating with and engaging stakeholders

Cynefin also helped me to communicate with the client and engage her in the evaluation design process. When we failed to predict all factors influencing the application of training in the workplace, I was able to explain this as an accurate assessment of the dynamics at play, not a failure of the program theory or logic modeling. Rather than becoming frustrated or discouraged, she was energized and remained engaged in the evaluation design process. She saw how approaching each evaluation question differently would result in a more complete picture of the dynamics related to gender analysis and planning performance in her organization. She understood how surveys would work well for measuring learning in the training, but that qualitative analysis of interviews would uncover factors influencing application of new skills.

Looking ahead

Despite the benefits described above, a number of unresolved issues remain regarding application of Cynefin to evaluation practice. Interpreting sense-making strategies as evaluation approaches warrants more exploration. In the field, there is a notable absence of consensus about which methods are most appropriate for various dynamics. Systems thinkers working in evaluation don't agree about which evaluation methods work best in different causal dynamics. Despite these limitations, the Cynefin framework shows promise in addressing practical problems evaluators face in situations with multiple dynamics.

Acknowledgements

I gratefully acknowledge the encouragement and guidance provided by Dr. Bob Williams during the evaluation planning process described here, as well as his valuable comments on this paper. I would also like to thank Dr. Tina Christie and Dr. Stewart Donaldson of Claremont Graduate University who provided feedback on this paper under the auspices of the certificate program for advanced studies in evaluation. The final version of this paper benefited greatly from James Russell's editorial assistance.

References

- Brinkerhoff, Robert O. 2010. *Telling Training's Story: Evaluation Made Simple, Credible and Effective*. San Francisco: Berrett-Koehler Publishers, Inc.
- Brinkerhoff, Robert O. 2005. "The Success Case Method: A Strategic Evaluation Approach to Increasing the Value and Effect of Training." *Advances in Developing Human Resources* Vol. 7, No. 1 February 2005 86-101. Online version at: <http://adh.sagepub.com/cgi/content/abstract/7/1/86>
- Christie, C.A. & Alkin, M.C. 2008. Evaluation theory tree re-examined. *Studies in Educational Evaluation*. Volume 34, Issue 3, September 2008, Pages 131-135.
- Donaldson, Stewart I., Laura E. Gooler, and Michael Scriven. 2002. "Strategies for Managing Evaluation Anxiety: Toward a Psychology of Program Evaluation" in *American Journal of Evaluation*, Vol. 23, No. 3, 2002, pp. 261-273.
- Hargreaves, Margaret B. 2010A. "Evaluating System Change: A Planning Guide," *Mathematica Policy Research*, Inc.
- Hargreaves, Margaret B. 2010B. "Methods Guide for Evaluating System Interventions," *Mathematica Policy Research*, Inc.
- Kurtz, C. F. and D. J. Snowden. 2003. "The new dynamics of strategy: Sense-making in a complex and complicated world." *IBM Systems Journal*, vol 42, number 3, page 462-483.
- Patton, Michael Q. 2010. *Developmental Evaluation: Enhancing Social Innovation and Use in Complex Systems*. New York: Guilford Press.
- Rogers, Patricia. 2009. "Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation," *Journal of Development Effectiveness*, Volume 1, Issue 3 September 2009, pages 217 - 226.
- Rogers, Patricia. 2008. "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions," *Evaluation* 2008; 14; 29.
- Williams, Bob. 2010. "Cynefin" in *Systems Concepts in Action* by Bob Williams and Richard Hummelbrunner. Palo Alto: Stanford University Press. Forthcoming.
- Williams, Bob and Iraj Imam (Eds.). 2007. *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: EdgePress of Inverness.
- Westley, Frances, Brenda Zimmerman and Michael Quinn Patton. 2006. *Getting to Maybe: How the World is Changed*. Toronto: Random House Canada
- Zimmerman, Brenda and Sholom Glouberman. 2002. *Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like?* Discussion Paper No. 8. Ottawa: Commission on the Future of Health Care in Canada.