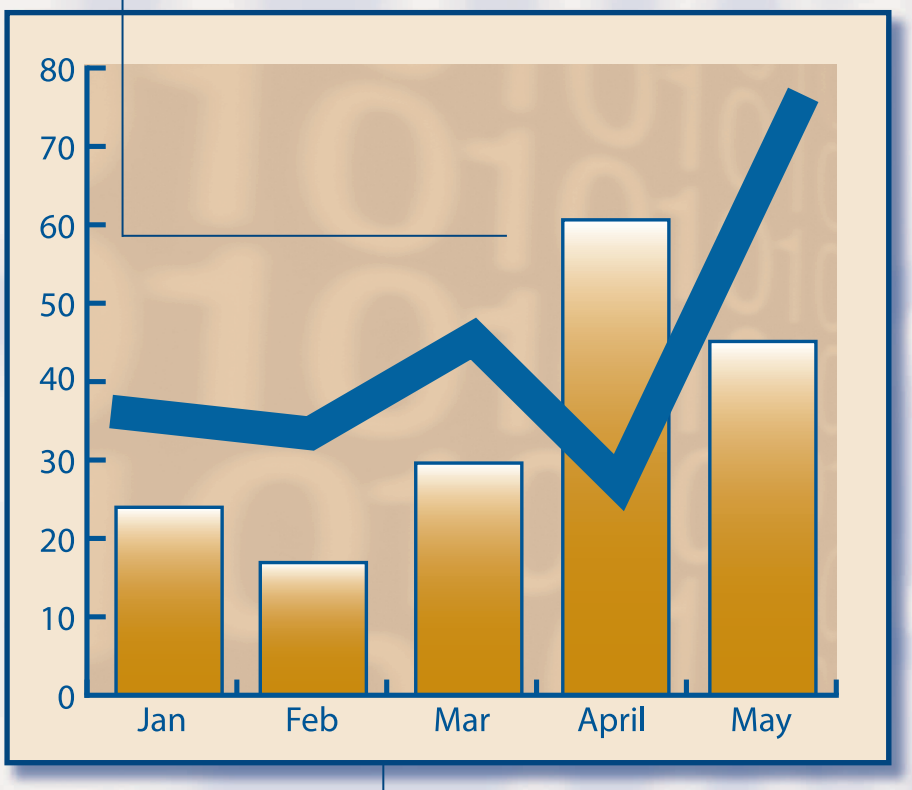
Designing Initiative Evaluation

A Systems-oriented Framework for Evaluating Social Change Efforts

W.K. KELLOGG FOUNDATION



September 2007

Foreword

The W. K. Kellogg Foundation, like many foundations, has made a serious commitment to evaluation and learning. However, as WKKF grantmaking goals and strategies have become more focused on creating sustainable change, we have found that traditional evaluation approaches have not been as useful as we would like. Program staff and grantees have told us that evaluation misses both "the big picture" and the subtle changes in relationships, resource allocations, and roles that lead to the bigger changes. In short, evaluation has not taken into account the complexity of social change efforts and the dynamic nature of the communities and systems in which we work.

As part of an overall effort to become a better learning organization, WKKF has been more thoughtfully applying systems-thinking approaches to our work. We offer this as our current best thinking on how the work of evaluation might be made more useful by using a systems-thinking framework. We hope that this document will provide WKKF program staff and our evaluators a common framework and language for conducting evaluations that add value to our change efforts, and that other funders and evaluators also will find this to be useful.

We would like to thank all of the WKKF staff who contributed by giving feedback on early drafts of the document: Gillian Barclay, Teresa Behrens, Tony Berkley, Caroline Carpenter, Ted Chen, Patti Grimes, Oran Hesterman, Karin, Ladley, Gail McClure, Tom Reis, Ricardo Salvador, Jocelyn Sargent, Ali Webb, Terry Wright, Huilan Yang, Alice Warner, and Dawn Winstone.

Sincerely,

James E. McHale
Senior Vice President for Programs

Table of Contents

General Orientation to Initiative Evaluation 1
- Introduction 1
- Cluster and Initiative Grantmaking at WKKF 1
- The Complementary Nature of Project and Initiative Evaluation 2
- Phases of Initiative Evaluation 4
- Systems Perspective and Initiative Evaluation 5

A Framework for Initiative Evaluation 7

Four Evaluation Designs for Initiative Evaluation 10
- *Overview* 10
  - Design Considerations and Foci 11
- *Exploratory* Evaluation Design 14
  - A. Introduction 14
  - B. Designing an *Exploratory Evaluation* 14
  - C. Planning and Engaging in Data Collection 16
  - D. Making Meaning from Data and Shaping Practice 17
- *Predictive Evaluation* Design 19
  - A. Introduction 19
  - B. Designing a *Predictive Evaluation* 20
  - C. Planning and Engaging in Data Collection 21
  - D. Making Meaning from Data and Shaping Practice 22
- *Self-organizing* Evaluation Design 25
  - A. Introduction 25
  - B. Designing a *Self-organizing* Evaluation 26
  - C. Planning and Engaging in Data Collection 28
  - D. Making Meaning from Data and Shaping Practice 29
- *Initiative Renewal* Evaluation Design 31
  - A. Introduction 31
  - B. Designing an *Initiative Renewal* Evaluation 32
  - C. Planning and Engaging in Data Collection 35
  - D. Making Meaning from Data and Shaping Practice 37
- Summary of Features of the Four Evaluation Designs 38

Summary 40

References 42

Appendix A 44

The Food and Society (FAS) Initiative: An Example Initiative 44

Overview

Purpose:

This document is designed for use by external evaluators who conduct initiative evaluations for the W.K. Kellogg Foundation (WKKF) – and, hopefully, other foundations and government agencies. It presents a systems-oriented framework and four general designs for initiative and cluster evaluation.

The designs are based on systems concepts related to change and the dynamics of systems. The focus is not on considering all ideas about systems that could be applied to initiative evaluation, rather on how different dynamics within systems can serve as the basis for initiative evaluation designs.

An Ongoing Conversation:

This document is WKKF's first attempt to articulate this framework and the evaluation designs. Further refinements and new understandings are anticipated based on the experience of initiative evaluators. Feedback is welcome and encouraged as part of an ongoing conversation about the evolving practice of initiative and cluster evaluation. Email comments to: evaluation@wkkf.org.

Additional Resources:

An extensive reference list of some of the many resources available about systems and systems change from various disciplines and perspectives is included in this document. (For resources addressing project-level evaluations see www.wkkf.org/evaluation.)

Acknowledgments

The following individuals contributed to the production of this publication: editorial direction, Teresa Behrens and Huilan Yang; primary author, Beverly Parsons; and editing, Linda Helstowski. Other significant contributors include: Glenda Eoyang, John Sherman, and Bob Williams.

The Lay of the Land: Initiative Evaluation

Introduction

What options exist for evaluating a large-scale initiative? What is the basis for selecting one evaluation design over another? Can differing approaches be used simultaneously to evaluate a complex initiative? How does a systems orientation contribute to initiative evaluation?

These are some of the many questions to consider in approaching evaluations of multi-level and multi-site bodies of work. To help answer such questions, this document outlines the approach the W. K. Kellogg Foundation (WKKF) is developing for cluster and initiative evaluation and four key evaluation designs based on looking at initiatives as complex systems.

The document assumes the reader has a general knowledge of evaluation practices and models as it explores and integrates current and emerging principles from varying perspectives with specific examples and learnings.

Cluster and Initiative Grantmaking at WKKF

Over the past three decades, WKKF has introduced two types of grantmaking that involve multiple projects, locations, and grantees: (1) "clusters" that typically focus on the exploration and/or development of new approaches to a program issue area; and (2) strategic "initiatives," designed to create systems changes that will lead to intended long-term, sustainable impact.

As WKKF began to increasingly support clusters and initiatives, it became clear that new opportunities for learning and evaluation arose beyond those available when looking at single projects; thus began cluster and initiative evaluation. (The evaluation designs for clusters and initiatives evaluation are similar. They are termed together as "initiative evaluation" in this document.)

The Complementary Nature of Project and Initiative Evaluation

Project evaluation and initiative evaluation serve different purposes and involve different stakeholders. A project evaluation focuses on the specific project of a grantee and serves the needs of the project leaders and participants. (For project evaluation information: www.wkkf.org/pubs/Tools/evaluation/Pub770.pdf.)

On the other hand, **an initiative evaluation looks across projects and their meaning within the initiative as a whole.** It serves the learning needs of WKKF, particularly program staff, as well as initiative grantees. It also informs the broader funder and practitioner fields in which WKKF works. Project evaluators are selected by each grantee whereas the initiative evaluator is selected by the WKKF initiative leaders.[1]

Initiative leaders are likely to use initiative evaluations to:

1. provide accountability to the WKKF Board for the expected benefits of the initiative [2];
2. gain a deeper understanding of the change theory under-girding their initiative to increase its practical utility;
3. assist in making adjustments in their funding strategy to better align with the evolving theory and conditions in the project sites;
4. support projects as they adjust their activities; and/or
5. determine how to sustain the beneficial aspects of the initiative beyond the period of its funding.

Initiative evaluation findings help project leaders put their work in perspective within the bigger picture addressed by the initiative. While project evaluations provide them with project-specific data, the initiative evaluation gives them unique information about relationships and patterns across the initiative that are not evident from a single project.

1 Project evaluations collectively can contribute to the initiative evaluation but are seldom the primary basis of an initiative evaluation.

2 Note that evaluation information is provided to the Board by the initiative leader, not the initiative evaluator.

Food and Society: An Example Initiative

WKKF has many initiatives under way at any point in time. The Food and Society initiative (FAS) was arbitrarily selected as an example to be used throughout this document to illustrate how evaluation activities vary by design and the situations to which they are being applied.

Launched in 2000, FAS is based on a vision of a future food system that provides all segments of society a safe and nutritious food supply grown in a manner that protects health and the environment and adds economic and social value to rural and urban communities.

FAS recently completed its first five years of operation and began a second five-year period. See a description of the FAS initiative in Appendix A.

Roles in FAS: Initiative and Project Evaluators

The initiative evaluator for FAS is the Minnesota-based Headwaters Group. The Headwaters Group has worked with the initiative since its inception. The lead evaluators had previous experience as directors of nonprofit organizations and grantmakers as well as evaluators.

The initiative evaluators helped project leaders select project evaluators if they hadn't already identified an evaluator. Some of the grantees selected their project evaluator before receiving funding while others waited until after they received their grant.

Furthermore, some project evaluators are working across several grantees. After five years, there were about 20 different project evaluators involved in FAS.

Phases of Initiative Evaluation

The evaluation process can be thought of as having three phases (see Figure 2):

1. designing the overall evaluation approach;
2. planning and conducting data collection; and
3. making meaning from the data and using what is learned from the evaluation.

The phases are iterative, often overlapping, and varying in length. They may repeat many times within an initiative evaluation, with each iteration influencing the next one. Although the phases repeat, the way each phase is carried out, how distinctive the phases are, and who is involved differ from one time to the next.

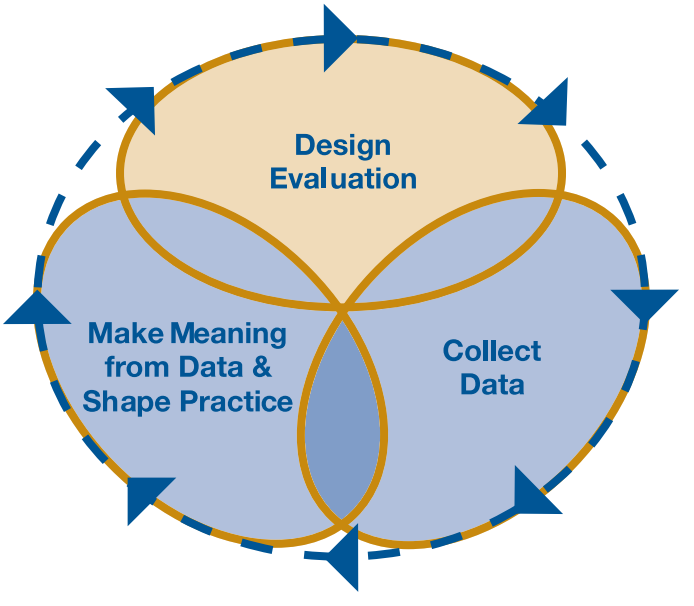


Figure 2. General Steps in Evaluation Process

Systems Perspective and Initiative Evaluation

Since the 1940s, researchers and practitioners in a variety of fields have used the concept of systems to inform their work.[3] This document does not consider the full range of ideas about systems that could be applied to initiative evaluation[4]. Rather it concentrates on how different dynamics within systems can serve as the basis for initiative evaluation designs. That said, there are a few general concepts about systems and their relationship to initiatives that are important to acknowledge.

- **Many different definitions, modeling techniques, practices, and methods have emerged** to try to address the systemic nature of human institutions and interactions. Related terms include, for example, systems dynamics, soft systems, causal loop diagramming, complex adaptive systems, and cybernetics. They are examples of theories, methods, and associated techniques that seek to describe the boundaries, diversity, and relationships that shape behavior of interconnected entities.
- **Each approach has its own particular emphasis and perspective that can inform an evaluator's work.**[5] The orientation used in this document draws especially on the concept of complex adaptive systems and emphasizes the variation in dynamics within systems.
- **Initiatives involve formal and informal social systems** such as the education system or a community. **Such systems are complex, dynamic, fluid, and flexible.** An initiative can be thought of as a complex system or highly entangled collection of systems. It may be thought of as systems within systems (nested systems),[6] networked systems[7], a web of multiply entangled systems, a complex, evolving, continually changing interconnected pattern of relationships, or (likely) some combination of all of these.
- **How one conceptualizes a system can have an impact on how one carries out an evaluation.** Some system theorists think of systems as reality and others view them as mental constructs – conceptual frameworks – that help understand the world even though they may not completely fit reality. It is similar to distinguishing the glasses one uses to view a landscape from the landscape itself. Wearing different glasses with different types of

3 See http://www.bobwilliams.co.nz for a brief description of many system concepts.

4 See Williams, B. and Imam, I, Systems Concepts in Evaluation: An Expert Anthology for examples of how other systems concepts may be useful in evaluation.

5 See Williams, B. and Imam, I (2007) for examples of systems. See Stacey (2007) for an elaboration of underlying concepts that shape views of systems, organizations, and change.

6 A nested structure is only one way of conceptualizing a system. An example of a nested hierarchical system would be community-based social service agencies which are within county governmental units, which are within state governmental units, which are within federal units. Another example would be hierarchical structures within a corporation.

7 Systems may be viewed as networks, for example, informal nonhierarchical connections among organizations or individuals as in partnerships.

lenses does not change the landscape itself but allows the possibility of experiencing the landscape in a different way. This document leaves it to the users to determine if they are viewing systems as reality and/or as mental constructs.

- **Initiative evaluators seldom, if ever, evaluate whole systems or the whole collections of systems of an initiative.** However, evaluators keep the "whole" in their peripheral vision as they focus on the parts that they are investigating.

- **Concepts of systems and system change are affected by deeply held – and often unarticulated or unrecognized – principles, values, beliefs, and assumptions.** These are the deep structures of social systems and are core aspects of initiatives. These core aspects underlie an initiative's theory of change.[8] Beliefs and assumptions about human nature, how people learn, how people interact with one another, and social phenomena may not be articulated in a theory of change. However, these beliefs and assumptions shape how one approaches the design of an initiative, involvement in it, and its evaluation.

8. A theory of change (most often depicted as a diagram accompanied by supporting narrative) presents the theory of how the initiative expects to achieve its intended outcomes and/or move in a desired direction. It highlights interconnections of system components or relationships among systems that are particularly relevant to the initiative. (See the FAS theory of change description and diagram in Appendix A for an example.)

A Framework for Initiative Evaluation

Initiative leaders and evaluators are concerned not only with the components of an initiative but also with the patterns of change of the systems within or affected by the initiative. **It is the change over time and across locations and context that is of special importance in an initiative evaluation.** Understanding the dynamics of interactions within the initiative is critical. Consequently, the evaluation designs presented in this document are based on variations in the dynamics of systems.

This section provides an overview of a theoretical framework for describing the dynamics of social systems that serves as the foundation for the four general evaluation designs for initiative evaluations at WKKF. They are based on a way of viewing social systems that draws on the work of Ralph Stacey (1996) and Brenda Zimmerman and colleagues (2001). The basic idea of this framework is that two factors can be used to describe the variation in dynamics of social systems (e.g., an initiative) and consequently how changes occur in systems. These factors are:

- **the degree of agreement** among those in a group, team, organization, community or other unit (about, for instance, the changes needed in a social system) and
- **the degree of certainty** about actions, conditions, or consequences of actions that exist or are likely at places in the system(s).

This framework does not replace the many evaluation planning, data collection, analysis, and communication techniques that evaluators use. Rather it provides a broad framework that allows the evaluator to select tools and techniques that especially illuminate change within complex systems such as initiatives. Figure 3[9] is a simplified way of visualizing the interactions of certainty and agreement within social systems. **Such interactions create three types of dynamics within a social system: unorganized, organized, and self-organizing.**

For purposes of this document, each of these dynamics is presented as a distinct aspect of a system (or collection of entangled systems). It is important to recognize, however, that **in a complex system or collection of systems, the dynamics are highly intertwined and that systems exist within a broader context** than is considered in the evaluation design.

9 This figure is a simplified depiction of the factors nesting different dynamics within social systems. It draws on the work of Ralph Stacey (1996) with adaptations by Zimmerman, Lindberg, & Plsek (2001) and Holladay (2003). This orientation to systems is one portion of the complex and rich wealth of information available about systems and other ways of viewing social processes and situations (see Stacey, 2007; Midgley, 2003; Johnson, 2001).

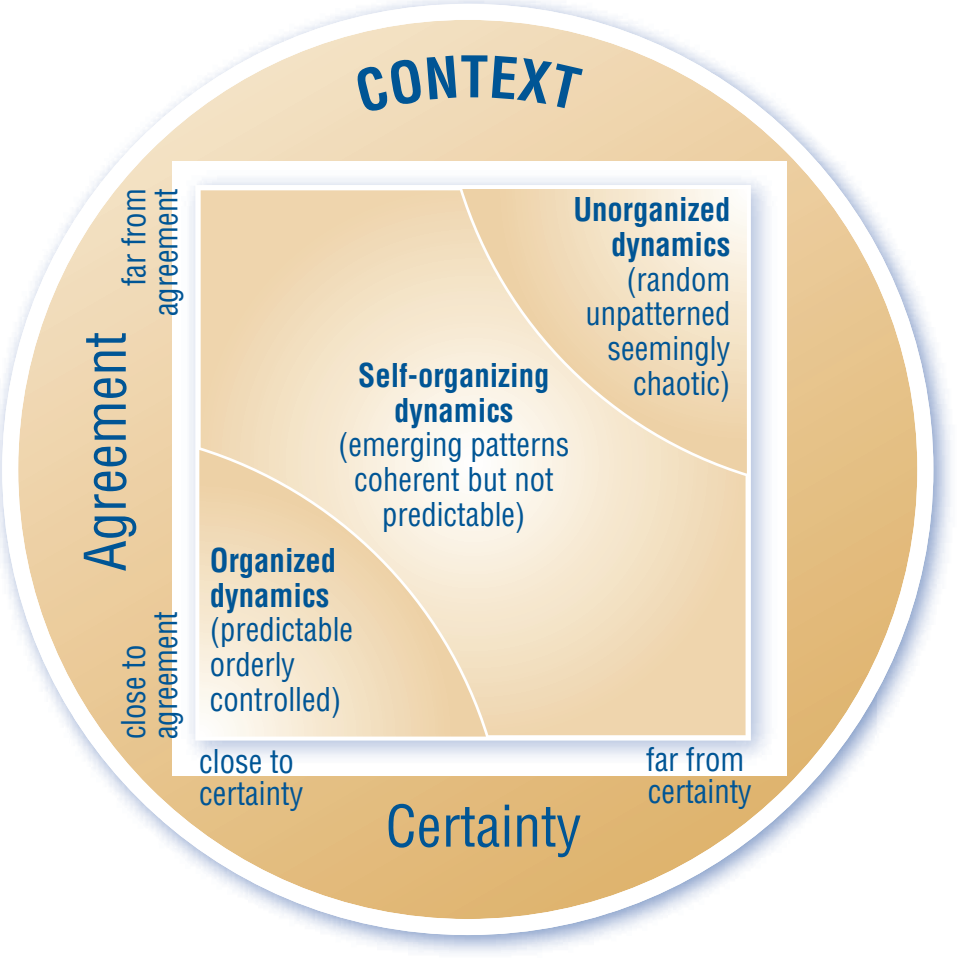


Figure 3. Three Dynamics of a Social System and its Context

Figure Description:

- At one end of the spectrum (upper right corner), where systems exhibit **both low certainty and low agreement is a random, unorganized dynamic.** There is considerable uncertainty and disagreement; actions appear random, unpredictable, unorganized, and without apparent patterns. In essence, the sense of system has disintegrated and change is unpredictable and incoherent.

- At the opposite end of the spectrum where the **levels of certainty and agreement are high is the stable, organized, predictable dynamics** of the system (lower left corner). Because this dynamic of the system is fairly orderly, organized, and agreed-upon, the outcome of an action can be predicted with fairly good certainty (cause-and-effect relationship).
- **Between these two ends of the spectrum is a special dynamic where the system is far from the equilibrium of either an organized state or the disintegration of an unorganized state** (middle section). This dynamic does not result from centralized control or intention. Rather, the system is said to **self-organize**. That is, identifiable patterns and principles emerge as the initiative proceeds. Agents in the systems mutually adjust without an overall design (although each agent may have intentions). Although behaviors or results are not predictable, they are influenced by the local action of agents operating as they deem appropriate or feasible.

Notably, the whole system or collection of systems with these multiple dynamics is embedded in a larger social context with many other systems that may be affecting the system(s) of interest (represented by the circle around the diagram in Figure 3). That is, the initiative is within a broader social context with its own dynamics and components that also need to be taken into account when evaluating an initiative.

Thus, **conceptualizations and theories about social systems, phenomena, and change shape both the operation of an initiative and its evaluation design.** The theories and concepts evolve and are more fully understood as the initiative progresses. The evaluation designs in this document are intended to be flexible enough to work with multiple theories of change so that any initiative can use one or more of these designs to craft an evaluation that is appropriate for its orientation.

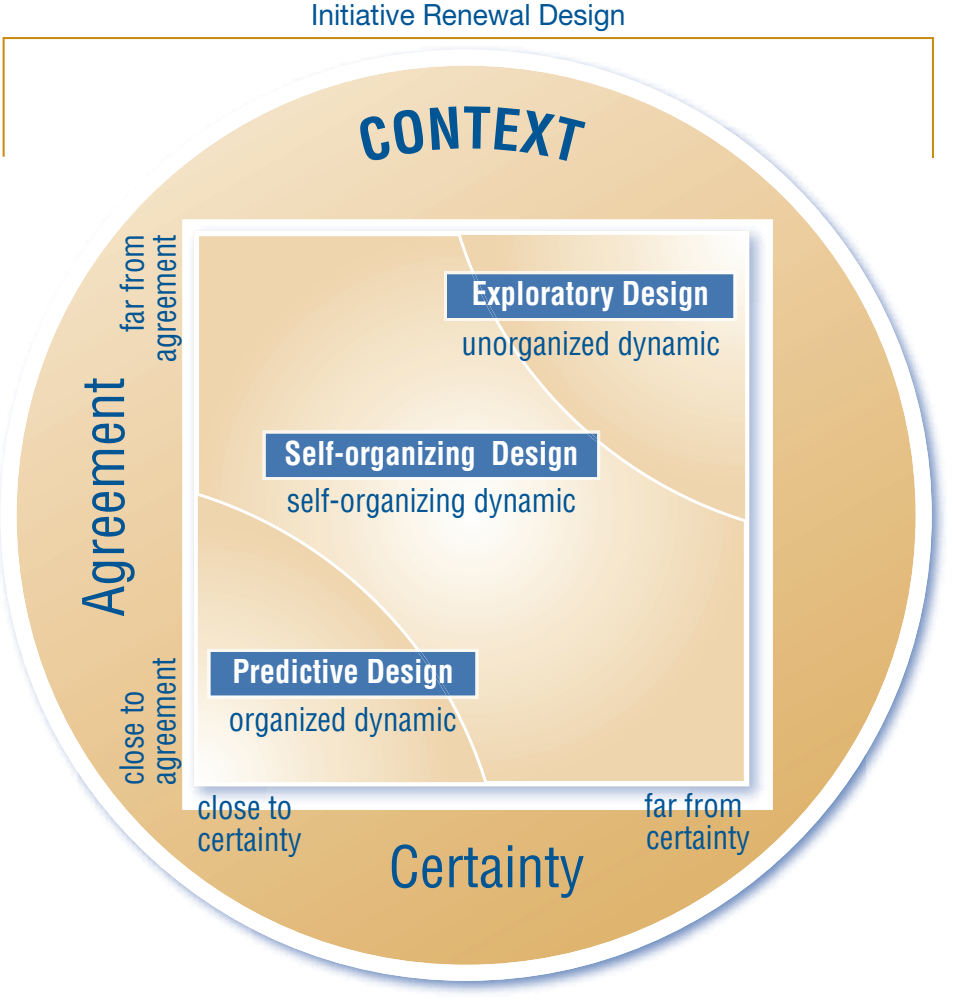
It is important to underscore that using this systems orientation does not mean that every aspect of the initiative is evaluated. **Only the aspects that need to be studied to help gain greater understanding of the initiative are evaluated.** Consideration of the dynamics of systems helps the initiative leaders and evaluators locate appropriate aspects of the initiative for evaluation. **Initiatives are about change. Change is shaped by dynamics.**

Four Evaluation Designs for Initiative Evaluation

Overview

Following are descriptions of four evaluation designs that build on the different dynamics of complex social systems (as outlined in earlier sections). These are: (**1**) ***Exploratory***, (2) ***Predictive***, (3) ***Self-organizing***, **and** (4) ***Initiative Renewal*** **evaluation designs.** The first three evaluation designs represent three ways of viewing the initiative with each one focusing on one of the dynamics discussed earlier - unorganized, organized, and self-organizing. The fourth views the dynamics of the initiative as a whole and its broader context.

Figure 4. Match of Initiative Evaluation Designs to Social System Domains and Context



Building on the earlier discussion (and Figure 3), the relationship of the evaluation designs to the dynamics of complex social systems is shown in Figure 4. The circle represents the broader context and the middle square represents the initiative with the three types of dynamics.

The *Exploratory* and *Predictive* designs have been the most commonly used types of evaluation designs in past WKKF initiatives and in the field of evaluation generally. The *Self-organizing* design brings in concepts that are fairly new to the evaluation field and opens up interesting possibilities for the initiative evaluator to develop a richer understanding of an initiative and enrich the evaluation's contribution to the initiative.

Design Considerations

An initiative may use a combination of these designs or a single design depending on the match of the (assumed or known) dynamics of the situation being studied with the underlying perspective of the design. **Shifting the evaluation design over time or using more than one of these designs concurrently is appropriate in an initiative evaluation.** More than one of these designs may be used simultaneously or sequentially during the initiative's evaluation because multiple dynamics may be operating in the initiative.

To effectively carry out a specific evaluation design, initiative evaluators and leaders consider the number of sites; the amount and type of variation among them; the speed and nature of changes that are occurring; and the expectations of the funders.

The initiative leaders and evaluators work together to first identify the dynamics of interest within the initiative and then select one or more evaluation designs congruent with those dynamics. As the initiative matures and as knowledge is gained, it may be appropriate to adjust the evaluation design or shift to one or more new designs to develop a deeper understanding of the initiative.

Shifts in designs may occur when new knowledge suggests new emphases; an initiative moves from one phase to another; new grantees join the initiative; a significant elaboration or refinement in the theory of change occurs; and/or the efficient use of evaluation resources calls for a new design. The shift may be from one of these four general designs to another or it may be a shift in emphasis or focus within a design type.

Case-In-Point

Food and Society (FAS)

Specific examples from the FAS initiative (see Appendix A) are used to help illustrate the evaluation designs as "Case-In-Point" sidebars throughout this document.

The first design – *Exploratory* design – uses examples of what FAS evaluators did during the first five years of the initiative. The next three designs – *Predictive, Self-organizing*, and *Initiative Renewal designs* – present possible approaches for the evaluation in future years of the initiative.

If and when an evaluation design shifts, it is important that initiative stakeholders understand the changes so they have appropriate expectations about their involvement in the evaluation, what is likely to be learned from the evaluation, and when evaluation findings will be available.

Design Foci

The different foci of the four evaluation designs allow each design to provide answers to different type of questions related to systems change (See Table 1 below). The initiative leaders and evaluators formulate specific questions related to one or more of these general aspects of systems change.

Table 1. Relationship of Evaluation Designs to Evaluation Focus

Initiative Evaluation Designs	Design Focus
Exploratory evaluation design (unorganized dynamics)	potentially important components and dynamics of change that are not yet delineated in the initiative's theory of change
Predictive evaluation design (organized dynamics)	cause and effect relationships between structured interventions of an initiative and the predicted outcomes/changes
Self-organizing evaluation design (self-organizing dynamics)	the patterns of change emerging from self-organizing dynamics within the initiative
Initiative Renewal evaluation design (initiative in context)	the interplay of multiple dynamics of change within the initiative and with its context that enrich its theory of change and have implications for the sustainability of the initiative

Design Distinctions

The four designs provide different conceptual frameworks that allow experience of the initiative in different ways. By aligning the evaluation design with the different dynamics of systems change, the evaluation can be more tailored to the nature of change within an initiative, resulting in evaluations that are timely, targeted, efficient, and useful. More than one of the designs may well be used at any given point in time – and multiple designs may be used through the lifecycle of the initiative – because multiple dynamics and change processes are likely to be occurring.

As noted earlier, the three variations in dynamics – unorganized, organized, and self-organizing – extensively overlap and intertwine. Thus, just as there are not rigid boundaries between the dynamics within the systems, **the evaluation designs are not rigidly distinct.** Viewing each separately can provide different insights into the nature of the work that can help stakeholders make choices about how to manage and/or participate in the work to move toward the desired situation.

Design Descriptions

For each of the four designs, the following sections provide a brief orientation to the design and consider the design in terms of the three phases of evaluation: (1) designing the evaluation, (2) data collection, and (3) making meaning and shaping practice.

The designs are presented in the order of moving from one end of the spectrum of certainty and agreement to the other end and then back to the middle (from *Exploratory* to *Predictive* to *Self-organizing*). Thus, the dynamics and consequent evaluation designs related to the two ends of the spectrum (as illustrated in Figure 4) which are most familiar are described first and are intended to illustrate the contrast with the *Self-organizing* design, which is less familiar. Then, the fourth design encompasses the other three in relationship to the larger context of the initiative.

I. *Exploratory* Evaluation Design

Case-In-Point

FAS *Exploratory* Evaluation Part 1 – General Information

The *Exploratory* evaluation design was the method of choice for the first few years of the FAS initiative.

The initiative evaluators designed an *Exploratory* evaluation for each of the five major outcome areas of the initiative: 1) policy, 2) scholarship, 3) farms/acreage, 4) markets, and 5) partners.

This design kept them open to a wide range of possibilities and gave them time to get to know the stakeholders and understand the range of dynamics and situations that might be relevant to the initiative.

A. Introduction

The *Exploratory* evaluation design is used to investigate the unorganized dynamics of an initiative. When an initiative begins, there may be little agreement among stakeholders about how the multiple systems involved in the initiative operate and little research that shows how the multiple parts of the initiative affect one another. The systems themselves may be undergoing major change, resulting in considerable uncertainty.

As an initiative develops, leaders also may discover new areas of interest that are not well understood. And, as time goes on, initiative leaders continue to have areas of interest to explore that are not identified in their original theory of change for various reasons, including lack of agreement or certainty about its relevance. In the *Exploratory* design, the evaluator engages in the evaluation from either an insider or an outsider perspective.

B. Designing an *Exploratory* Evaluation

Evaluation Questions Addressed

The *Exploratory* evaluation design is used to identify potentially important components and dynamics of change that are not yet delineated in the initiative's theory of change. The FAS evaluators used four broad questions to guide their *Exploratory* evaluation of the initiative (see Sidebar for further detail):

1. What happened?
2. What difference did it make?
3. What has been learned?
4. How will it inform the future?

In using these questions, they paid particular attention to what was unexpected or surprising. In this design, specific outcomes are not necessarily used, and no activities or processes of the initiative necessarily are assumed to be preferable to others.

Situation Where This Design is Used

The *Exploratory* design is used to look at aspects of the initiative that appear disorderly, unorganized, or not well understood. Such areas are unpatterned, surprising, and unpredictable. They may exist before the initiative starts or may be created intentionally or unintentionally during the initiative. For example, some systems within an initiative may be overly controlled and part of the initiative's task may be to remove the control or shift to different controls.

The transition may create an unorganized dynamic. By such expansion of the unorganized dynamic, initiative leaders may stimulate creativity and opportunities for new relationships and patterns to form.

The *Exploratory* evaluation is designed to see what insights can be gained about the areas where the complexity of the initiative is not yet understood or articulated.

Results from this design are likely to enrich the theory of change by, for example, indicating where boundaries or relationships can be established or encouraged that will help support the desired direction or outcomes of the initiative. For example, FAS added the Ten Percent Goal[10] as a result, in part, of what was learned through the evaluation. (The Ten Percent Goal is framed as a controlled dynamic, for which a *Predictive* evaluation design is appropriate.)

Case-In-Point

FAS *Exploratory* Evaluation
Part 2 – Evaluation Design

The reauthorization of the 2003 federal Farm Bill related to the FAS policy outcome was one area of focus for the evaluation. Building from their broad questions (see text), the evaluators gathered data about these four policy-related questions:

1. What outreach efforts did FAS grantees engage in to educate decision makers and activists?
2. What differences did the educational activities make in terms of legislative content, subsequent implementation of the Farm Bill, and partnership processes?
3. What have we learned about informing public policy?
4. How will the learnings inform the future work of the initiative?

The use of the *Exploratory* design in FAS resulted in considerable learning about federal policies that affect the initiative. For example, it became clear that:

- greater collaboration among grantees striving to inform the farm bill was necessary;
- greater diversity of voices participating in policy advocacy was needed; and
- federal policies beyond the Farm Bill would need to be considered within the initiative.

continued on next page

10 The initiative's Ten Percent Goal is to stimulate a shift that will result in at least ten percent of the U.S. food system functioning in alignment with the FAS vision. The vision is: "We envision a future food system that provides all communities access to a safe and healthy food supply grown in a manner that protects the environment and adds social and economic value to both urban and rural communities."

C. Planning and Engaging in Data Collection

Part 2 – Evaluation Design continued

When addressing the work related to markets, evaluators identified a gap between the vision of grantees wanting to influence markets and their capacity to implement that vision. They learned about business practices that grantees could use to significantly contribute to supporting businesses that enact the FAS vision.

When the initiative was designed, the leaders had not yet formulated theories about these issues nor did they have the necessary information to show how important the issues would be for achieving the FAS vision.

As time went on and these particular issues were explored more fully, new understandings (such as those noted above) surfaced that helped shape the next phase of the initiative.

After the initiative had been in place for several years, leaders recognized opportunities to significantly enhance the initiative based on what was learned, including adding the Ten Percent Goal. (The Ten Percent Goal is framed as a controlled dynamic, for which a *Predictive* evaluation design is appropriate.)

The *Exploratory* design often uses data gathering **tools and methods that are relatively unstructured and may shift rapidly** as information is gathered and new insights are gained. In contrast to other designs, it is not clear where or how to look for patterns (as is done in the *Self-organizing* design) or what predictable relationships may be worthy of attention (as addressed in a *Predictive* design).

Qualitative methods are commonly used in the *Exploratory* design for on-site data collection (e.g., interviews, focus groups, site visits). Because it is hard to know what the scope of the investigation needs to be at the outset, it often works well to use divergent approaches, such as "snowball" interview techniques (i.e., asking each interviewee who else needs to be interviewed).

Also, **it is useful to identify distinct perspectives about various situations.** In the FAS initiative some people may be focused on food quality while others are focused on protecting the environment or building locally owned food companies; some may be pursuing policy change while others are focused on markets or communications. When exploring a wide range of diverse perspectives, evaluators can enhance their understanding by gathering data about the beneficiaries as well as the victims of various aspects of the initiative.[11]

In many ways, **this kind of evaluation calls for the skills of an investigative journalist**. Those who handle these kinds of evaluations bring an open mind and a willingness to try different conceptual frameworks. They explore and inquire into areas that appear to be useful. If they uncover information that they are interested in pursuing, they follow the leads they have discovered. They may move from one conceptual framework to another in response to the inquiry.

11 Critical systems heuristics analysis may be a useful approach in this regard. See http://www.bobwilliams.co.nz .

An *Exploratory* design may also include environmental scans, review of related research, and study of existing data and analyses that provide clues to what is happening. Another approach is to gather opinions and perspectives from experts in different fields. The FAS initiative evaluators gathered considerable data from large-scale databases to explore options and ideas in regard to each of the goal areas of the initiative. For example, they used governmental databases to determine the percent of agricultural lands managed in ways supportive of FAS values. All in all, they explored a wide range of issues that might reveal dynamics of importance in systems change.

Case-In-Point

FAS *Exploratory* Evaluation Part 3 – Data Collection

The initiative evaluators gathered data over several years regarding the federal Farm Bill. They focused on educational activities provided to activists, policymakers, and their staff. They gathered data through these multiple means:

- survey of FAS grantees about their educational activities for policymakers and staff related to the Farm Bill;
- review of grantee annual reports; and
- interviews with key activist groups and federal policymakers and their staff about the value of various educational activities of grantees and others.

Evaluators asked interviewees for the names of others to interview to expand understanding of the situation.

D. Making Meaning from Data and Shaping Practice

In an *Exploratory* evaluation design, **the process of making meaning from the data and using the results is often interconnected**. Typically it is very helpful to involve a number of people with different perspectives in analyzing qualitative data.

Ongoing conversations between the initiative evaluators and the initiative leaders and/or grantees also are beneficial to test ideas and jointly formulate alternative ways of viewing the data.

In so doing, they generate ideas of where and how to encourage more coherence, certainty, and agreement through self-organizing strategies or through planned/controlled actions with defined intended outcomes.

Drawing on research in various fields to create conceptual frameworks for analyzing the data in multiple ways may also be useful. Grounded theory and logical-deductive theory both can be useful. These approaches can help identify opportunities to build certainty and agreement within the initiative, moving from seeing the situation as an unorganized dynamic to locating changes that may be closer to a self-organizing or predictable dynamic.

When using an *Exploratory* evaluation design, it is important to keep in mind that **the themes and observed relationships may be idiosyncratic and need to be tested further** with additional designs. A danger with *Exploratory* designs is over-interpreting the findings or seeing what one expects to see. It is important for evaluators to be aware of their own assumptions and try to set them aside as they look at the full range of data.

It often is useful to follow *Exploratory* designs with *Predictive* and/or *Self-organizing* designs related to the aspects of the initiative that have been informed by the *Exploratory* evaluation (see following sections).

In terms of reporting, frequent and informal memos and conversations often constitute a major part of an *Exploratory* evaluation, followed by periodic (e.g., every year or two) synthesizing reports. The written report serves as the basis for conversations among stakeholders as they continue to build conceptual frameworks and theories to leverage the unknown aspects of the initiative.

Case-In-Point

FAS *Exploratory* Evaluation
Part 4 – Making Meaning and Shaping Practice

Making Meaning:

The policy-area initiative evaluators summarized their data from each of their data-gathering processes (see previous box).

- They looked for themes regarding each of their guiding questions.
- They used data triangulation (i.e., a variety of sources and types of data) to help strengthen their understandings regarding each of the guiding questions.
- They used both qualitative (e.g., descriptions of educational activities) and quantitative data (e.g., numbers of grantees undertaking certain types of activities).

The evaluators and initiative leaders had monthly phone calls and semi-annual three-day retreats to discuss their findings and jointly interpret data and determine additional data collection. The evaluators prepared short memos and other periodic summaries of their data for use in these meetings.

At the end of four years, the initiative evaluators prepared a culminating evaluation report for each of the five outcome areas. It provided a summary of responses to each of the four guiding questions.

Shaping Practice:

The initiative leaders used the federal policy-related data to help determine if they had an appropriate mix of grantees within the FAS initiative who could inform federal policy. They also used the findings to identify other federal policies that may be relevant to the initiative.

II. *Predictive* Evaluation Design

A. Introduction

In addition to aspects of an initiative with unorganized dynamics (studied through the *Exploratory* design), most initiatives have activities that are carefully planned to lead to pre-defined outcomes. **It is a cause-and-effect dynamic**. Here the *Predictive* evaluation design is applicable. It focuses on the link between planned features of the initiative (e.g. actions, events, relationships, concepts, and/or values) and intended outcomes/results.

The evaluation looks at the situation through the lens of predictability and thus if the implemented plans lead to the intended outcomes. The evaluation is conducted from an outsider perspective. **Results from this evaluation design often are seen as important for accountability purposes** – a way to show concrete evidence of expected outcomes from the investment of the initiative.

This design is useful for looking at common outcomes across individuals, groups, organizations, and/or larger systems in fairly stable situations. It may take several years for the outcomes of an intervention[12] to be evidenced. Thus such designs may require data collection over extended periods of time.

Case-In-Point

FAS *Predictive* Evaluation
Part 1 – General Information

In 2006, FAS formulated a measurable goal: to stimulate a shift that will result in at least 10 percent of the U.S. food system functioning in alignment with the initiative's triple bottom line vision – a food system that:

a. provides all communities access to a safe and healthy food supply,

b. is grown in a manner that protects the environment, and

c. adds social and economic value to both urban and rural communities.

FAS wants grantees to show how they are contributing to this Ten Percent goal, as well as wants to see what grantee actions and conditions are particularly associated with achievement of the goal.

The initiative leaders and evaluators are considering using a *Predictive* evaluation design for the market-based outcome.

During the first five years of the initiative, the market-based *Exploratory* evaluation made such "link" testing possible. It identified specific grantee actions such as certain financial management, planning, marketing, and other business practices that are likely to be significantly related to contributing to achievement of the goal.

By building in realistic expectations at the beginning of the evaluation concerning when the findings will be available, leaders ensure that the communication plan for the initiative meshes well with the evaluation. Combining project-specific *Predictive* designs with some cross-project (initiative level) *Predictive* designs can produce an ongoing series of results that help serve as benchmarks for both funders and grantees.

12 An "intervention" refers to an intentional action that a person or organization takes as part of the initiative.

Case-In-Point

FAS *Predictive* Evaluation Part 2 – Evaluation Design

(Design under consideration by FAS)

There is considerable variation among the FAS grantees in their use or readiness to use a selected set of business practices identified through previous research/evaluation as likely contributors to the Ten Percent Goal.

One possible *Predictive* evaluation for use during the second five years is that the initiative evaluators may identify grantees who do and those who do not have the identified business practices operational within their organization. The process would include:

- They set up their *Predictive* evaluation design to answer the question: Do organizations that use certain business practices make a greater contribution to the Ten Percent Goal than organizations that do not use the practices? They compare the two groups in the extent to which they contribute to the Ten Percent Goal locally.
- They use a rating scale to determine which grantees have the business practices in place at a certain threshold level or higher.
- They select two groups of about 20 grantee organizations where the business practices are above the threshold level and ones whose practices in place are below the threshold. This allows them to compare the achievement of the goal by these two groups.

Through this means, they can gain insight into the extent to which various business practices need to be present to contribute in a meaningful way to the goal.

B. Designing a *Predictive* Evaluation

Evaluation Questions Addressed

The focus of the *Predictive* design is the cause-and-effect relationships between structured interventions of an initiative and the predicted outcomes/changes. If multiple types of interventions and outcomes comprise the initiative, as is the case with the FAS initiative, there may be more than one evaluation question and design of this type.

Each evaluation question may have its own evaluation design with multiple sites involved in each evaluative study. The evaluation questions address some type of predictable link. For example, the FAS evaluation might address the question: What business practices help an organization contribute to the Ten Percent Goal?

Situations Where this Design is Used

This design is used in situations that are (or are considered to be) fairly predictable, controllable, knowable, and causally determined. The situations have a fairly well-defined timeframe and a specified location or number of locations. These situations usually are characterized by a fairly high degree of intentionality about the link between:

a. the intervention (e.g., policies, practices, and conditions), and

b. one or more desired outcomes.

This is in contrast to the amorphous, open-ended situations investigated through the *Exploratory* or *Self-organizing* evaluation designs.

When the dynamic of the system is fairly stable, there is a relatively consistent relationship between action and outcome. Thus, the focus of the evaluation can be on the substance of the initiative. *Predictive* evaluations may be appropriate in hierarchically structured systems such education, social

services, and some businesses where such structures are designed as control mechanisms. When the *Predictive* evaluation design focuses on policy, it often requires gathering data over multiple years and many sites with intentional variation in site characteristics.

Planned actions within less hierarchical systems/situations (e.g., neighborhood associations and communities) also may be designed to lead to specified outcomes. However, the relationships between actions and outcomes may be weaker since leaders of these systems have fewer ways to exert control than leaders in a formal hierarchical system.

The *Predictive* design often provides results for accountability and can help leaders see what is within their power to influence. Initiative leaders often face a tension between accountability for specific programs/activities on the one hand and fostering complex and long-term systems change on the other. Sometimes it is appropriate to ease the tension by studying small controllable parts of complex systems to learn what works. Then these parts can be linked together being " aware that new interconnections may bring about unpredicted, emerging behaviors" (Zimmerman, Lindberg, & Plsek, 2001, p. 40).

Case-In-Point

FAS *Predictive* Evaluation Part 3 – Data Collection

(Design under consideration by FAS)

To carry out the *Predictive* design, the FAS initiative evaluators need common measures of the (a) business practices of the grantees and (b) level of accomplishment of the Ten Percent Goal. They use the rating scale described in the *Predictive* Part 2 sidebar to identify the level of relevant business practices.

Determining the level of contribution to the Ten Percent Goal is complex. The initiative evaluators work with each grantee to determine what the Ten Percent Goal means in this situation since the grantees are working in different aspects of the food system, e.g., some with fruit and vegetable farmers and others with farmers involved in meat production.

After they have a definition for each grantee of what achievement of the Ten Percent Goal means in their situation, they identify four levels of contribution from very low to high. This provides them with a common scale (i.e., a rubric) that gives them a common metric across sites.

C. Planning and Engaging in Data Collection

To answer the evaluation questions guiding the *Predictive* design, **evaluators collect data about one or more interventions and their intended outcomes** that are identified in the initiative's theory of change. They also need information on initial and ongoing conditions to aid interpretation. Qualitative data from key parties in the sites can help to explain the results of the quantitative analyses of relationships between the intervention and outcomes.

Some theories of change may have multiple interventions and outcomes (including intermediate as well as longer-term outcomes). The intermediate outcomes/outputs become part of the intervention to achieve the longer-term outcomes. Often the evaluator works with the participating grantees as well as initiative leaders to identify concrete indicators of the outcomes/outputs (e.g., specific learning outcomes; numbers of people served; number of dollars generated; amount or type of products produced; and/or land reclaimed for a given purpose).

When using the *Predictive* design, evaluators must keep in mind that **certain statistical analysis methods require a minimum number of sites or other units of analysis to produce reliable data**. This is an important issue to consider in site selection for the initiative as a whole or for the evaluation.

How much variation the design can reliably handle in the data collection methods and/or the type of data collected from site to site must also be considered. Evaluators use common measures of the outcomes and intervention or a means to convert varied measures to a common metric.[13] Similar data, a large number of sites, and attention to context make it technically easier to establish the links between the intervention and the outcomes/outputs. **Using common data across sites makes data aggregation easier.** (Of course, the data collection tool needs to appropriately capture what is happening in each site and not distort the picture of the results through forcing inappropriate similarities in data collection).

A logic model can be helpful in the *Predictive* design.[14] Logic models visually depict the relationship between planned project activities; desired outputs; and short-term, intermediate, and long-term outcomes/impact in a linear fashion. Although the aspects of an initiative that fit the criteria of a fairly stable system with strong agreement and certainty (as discussed earlier) may be a small part of the overall initiative, they can be of considerable importance for both practical and political reasons.

Evaluators can use a detailed logic model to determine what aspects of the intervention and the outputs and outcomes to measure and what data to collect regarding the existence and strength of these interventions. Project leaders can use the logic model as a management tool.

If the relationships are nonlinear, theory-of-change diagrams showing causal loops may be more useful than linear logic models. This is the case in the FAS initiative. (See theory of change diagram in Appendix A.)

The initiative and project evaluators frequently work together to determine the measures and coordinate data collection processes. They attend to the need for standardized data across many sites along with attention to context.

---

13 See Parsons, 2002, pp. 39-48 for a way to convert uncommon measures to a common metric.

14 See the W.K. Kellogg Foundation Logic Model Development Guide for more information on logic models.

D. Making Meaning from Data and Shaping Practice

The analyses for making meaning of *Predictive* design data seek to establish links between the intervention and the outcomes with varying levels of generality and causality. **A wide variety of analytic methods exist for the *Predictive* design.** Evaluators use statistical methods such as t-tests, ANOVA, and correlations to look at linear relationships. The analyses may include taking into account different initial conditions (e.g., differences in state policies or economic conditions in the communities served) that may exist among subgroups of the sites involved in the evaluation.

Inferential statistics may be used to establish links and generalize to a larger population when there is an assumption of a fairly linear relationship between the intervention and the outcomes (i.e., cause-and-effect or correlation). Other methods may be used to look at predictable non-linear relationships.[15] Qualitative data may also be used as a means of building the relationship (Patton, 2002).

Once the data are analyzed, evaluators interact with initiative leaders, project leaders, and/or participants to review the findings and help provide interpretive insights to make meaning of the analyses. Such interaction typically increases the utility of the evaluation findings.

Case-In-Point

FAS *Predictive* Evaluation
Part 4 – Making Meaning and Shaping Practices

(Design under consideration by FAS)

Making Meaning:

The initiative evaluators calculate correlation coefficients to look at the relationship between goal attainment and use of selected business practices. This method assumes a linear relationship between goal attainment and business practices. They also use what they learn while developing the rubrics to help them interpret the statistical findings.

Shaping Practice:

The evaluators prepare PowerPoint slides of preliminary analyses to use in conversations with grantees at a networking conference and in meetings with initiative leaders. After conversations they provide written reports with visuals that can be used by stakeholders as they continue their work.

The initiative leaders use the findings to identify technical assistance to provide their grantees regarding business practices and various ways to contribute to the Ten Percent Goal. They also include the information in their report to the WKKF board as part of their explanation for how they refined their selection of grantees.

15 Nonlinear causal relationships may be looked at through system dynamics or other formal methods of modeling. See http://www.bobwilliams.co.nz for references on systems dynamics modeling as a means of looking at non-linear relationships.

Reports of *Predictive* evaluations often are fairly formal written documents with tables of quantitative data and statistics that show the relationship between the outcomes and the intervention. **These reports show the analyses conducted along with possible interpretations of the findings.** They are prepared for the initiative leaders, the grantees involved in the evaluation, and other grantees who are moving in this direction. The information from a *Predictive* evaluation often is of interest to the WKKF Board as well as the grantee sites for accountability purposes.

The reports from this type of design are intended to help the initiative leaders see what aspects of the intervention and/or settings they or their grantees can and/or should control. Controllable components of an initiative that are likely to lead to some desired outcomes can motivate participants to continue.

The initiative leaders use the findings to refine their ways of providing technical assistance to grantees, selecting grantees, operating networking conferences, and working with the WKKF board in refining the initiative design and funding allocations. Grantees are expected to use the findings to refine their operating practices. Sites outside of the initiative that are considering implementing a specific intervention also may use the information.

III. *Self-organizing* Evaluation Design

A. Introduction

A less frequently used and understood design is the *Self-organizing* evaluation design. This is less frequently used because the design itself is not understood or because its importance in the change process is not understood.

Self-organization is defined as the process **"whereby new emergent structures, patterns and properties arise without being externally imposed on the system"** (Zimmerman, Lindberg, & Plsek, 2001, p. 270). No leader controls what happens. Instead of actions being controlled or controllable by the initiative or project leaders, self-organization occurs as actions emerge from the group or individuals who are involved in the initiative.

Although self-organization is often associated with work within communities, networks, and/or partnerships, it also occurs in formal organizational structures and other relationships where agreement and certainty are moderate to low.

In complex situations/systems, predictability (certainty) is limited in many situations. Instead many people are influencing the work with only a moderate level of agreement. Thus, it is important to understand self-organizing patterns.

Case-In-Point

FAS *Self-organizing* Evaluation Example Part 1 – General Information

(Design under consideration by FAS)

The initiative evaluators and leaders decide to look at self-organizing patterns by using the FAS conferences as an entry point. The leaders set up the conferences to encourage self-organizing.

Since 2001, WKKF has convened an annual FAS conference to inform the field and be an incubator of, and catalyst for, networking. About 550 participants attended the 2007 conference.

Informed by evaluation, WKKF has altered the format of the conference over time so participants increasingly identify and discuss the issues they deem most critical. The conference provides structured and unstructured sessions. The leaders chose the Open Space meeting format for the 2007 conference.

The initiative evaluators want to investigate self-organizing patterns among grantees. They decide to use the conference to identify people to serve as data sources (informants) and to identify important emerging and evolving topics.

They use a *Self-organizing* evaluation design to determine what patterns exist and emerge for their informants. The findings are expected to give clues to important self-organizing patterns within the initiative (including, but not limited to, patterns within the conference).

In complex self-organizing systems, a new order or pattern can emerge with no preplanning. The entangled complex systems don't necessarily move to being either stable or unstable. Rather, they may well be continually in a state of disequilibria (a moderate to low level of certainty and agreement), often characterized by contradiction and contention. Cooperation may coexist with competition, independence with interdependence on a long-term basis. (Stacey, 1996).

To understand what is meant by self-organizing, an Open Space meeting format provides a useful small-scale example (www.openspaceworld.org or Owen, 1997). Usually someone facilitates an Open Space meeting, but no official leader determines the content and process of the meeting. Participants are invited to co-create the agenda and identify and host their own discussion groups.

Spaces within the room are designated for discussions. Participants are free to move from group to group at any time guided by their own interests. Someone in each group records the conversation so it can be shared with the larger group.

The groups do not necessarily try to reach agreement. Rather the conversations inform each person's thinking as participants keep adjusting their individual behavior and thinking in response to those around them.

The *Self-organizing* evaluation design is congruent with the aspects of the initiative that have such self-organizing dynamics (as described earlier). A primary purpose for addressing such areas of an initiative is to understand the forces at play that are influenced but are not (or cannot) be planned or controlled by leaders/managers.

An initiative evaluator is looking for general patterns of similarities, differences, and relationships over time and locations that provide insights into ways those involved adapt to one another and local conditions to lead **or not lead in the desired direction.** Evaluators conduct the evaluation from an insider perspective.

Case-In-Point

FAS *Self-organizing* Evaluation Part 2 – Evaluation Design

(Design is under consideration by FAS)

The initiative evaluators and leaders decide to address the following evaluation questions:

- How do the conferences play into self-organizing that occurs before and after the conferences?
- How do topics of conversation and self-organizing processes shape patterns across the initiative?
- What patterns emerge in relationships?
- What conscious and unconscious boundaries seem to be influencing the self-organizing patterns?
- What types of diversity among the participants shape the self-organizing patterns?

To help answer these questions, the evaluators invite the participants who pose conference discussion topics to serve as ongoing data sources for the remainder the initiative.

Each year they add other participants who pose Open Space discussion topics. This gives them an evolving group of participants to help understand the self-organizing patterns.

B. Designing a Self-organizing Evaluation

Evaluation Questions Addressed

The evaluation questions that guide a *Self-organizing* design focus on investigation of patterns of change emerging from self-organizing dynamics within the initiative. For example, the evaluation may ask: What patterns of change in relationships occur over time? Answers to such questions help the initiative and project leaders see the limits of what they can plan and/or control.

The *Self-organizing* design **helps the initiative leaders enrich their theory of change** in the areas where a *Predictive* model is not appropriate.

Situations Where This Design Is Used

Self-organizing systems have a large and diverse number of agents that interact in adaptive and nonlinear ways. In a densely intertwined web of interacting agents (e.g., subgroups, individuals), each agent is responding to other agents and the environment as a whole. Each agent is continually adapting to its situation in the context of its relationships with other agents based on feedback, values, boundaries, relationships, and other conditions (Zimmerman, Lindberg, & Plsek, 2001; Stacey, 2000, Eoyang, 2001).

Although those involved in a self-organizing situation may have a desired direction that is generally agreed-upon, their relative valuing of the direction (versus their other values) shapes movement in that supposedly desired direction. Movement is shaped by a complex interplay of the self-motivation and independent and interdependent actions of many parties. **Actions and patterns start to emerge over time as people keep adjusting to their local situation and to one another's actions** – somewhat like drivers adjust to one another on a busy highway. (*See Johnson, 2001, for a very readable description of self-organizing systems and emerging patterns. See Midgley, 2003, and Stacey, 2007, for more technical and detailed descriptions of ways of thinking about self-organization.*)

When looking for patterns or changes in patterns, it is useful to attend to three features of systems: boundaries, relationships, and diversity.[16] Consider the boundaries (e.g., organizational roles) within which individuals or groups are operating consciously or unconsciously; the nature of the relationships they are developing; and/or the amount and type of diversity among them or affecting them.

Case-In-Point

FAS *Self-organizing* Evaluation Part 3 – Data Collection

(Design is under consideration by FAS)

The initiative evaluators gather several types of data over three years to understand the self-organizing patterns. At the conference, they track the topics of, and level of involvement in, the participant-determined Open Space discussion groups to determine shifts in interests and the interactions that affect the shifts. They also gather data from their identified informants before and after the conference (see Self-organizing Part 2 sidebar).

They ask their informants about actions they took since their last reporting, new ways of thinking that have arisen, and contacts they had with others related to the topics. At various times over three years evaluators ask, for example, about the type of diversity in the relationships, and what boundaries exist in and among relationships.

They ask them about shifts in their primary interests from one time to the next and the aspects of the theory of change they are most involved in. **Relationships are at the core of the patterns they are seeking to understand.**

16 The three categories of system characteristics affecting a situation – relationships, boundaries, and diversity – derive from the work of Eoyang (2001) with modifications by Williams and Imam (2007) and concepts from Stacey (2007). Stacey emphasizes the importance of diversity and relationships.

Initiative evaluators look for diversity that seems to stimulate or dampen the energy and movement within the intertwined systems. Differences in perspectives, for example, may create contradictions in systems that unpredictably lead to innovation on one hand or lack of productivity on the other.

The initiative leaders and evaluators work together to identify potentially important self-organizing dynamics within and among groups involved in the initiative. They choose units and topics for investigation where they think feedback about emerging patterns will be useful to those who are involved. Units of study may include neighborhoods; grantee sites; professional organizations; special interest groups; meeting attendees; subgroups within sites; connected entities across sites; and/or various types of collaborations or networking among groups.

C. Planning and Engaging in Data Collection

The initiative evaluators collect data across time and locations that help to identify patterns of self-organization. They consider interactions at local levels and the emergence of new patterns that can spontaneously arrive. They look for the evolution of patterns over time and space that are relevant to how the units of analysis are moving in ways that relate to the intentions of the initiative. **Data collection needs to be relatively frequent** since patterns can change in unexpected ways when seemingly small events trigger a cascade of new patterns.[17]

For example, some patterns are shaped by rules or behavior that are operating in a situation. These rules may be conscious or unconscious, intentional or unintentional. **Evaluators may look for apparent rules that shape the patterns.** Dominant-subordinate "rules" of behavior, for example, may create patterns within hierarchical organizations. Such patterns also may occur with social groups based on gender or race.

Other patterns are shaped by transient conditions. Whereas initiative evaluators in a *Predictive* evaluation may include a focus on the conditions that predict outcomes, in a *Self-organizing* evaluation they focus on the characteristics and conditions (e.g., types of feedback, relationships) that influence the speed or coherence of emerging patterns over time and locations realizing that the outcomes of those fluid patterns are quite unpredictable.

Data Collection Techniques

Qualitative data collection methods are commonly used in the *Self-organizing* design. Interviews, focus groups, tracking agendas or blogs, storytelling, and site visits are but a few examples. Journaling by key players, periodic reflective interviews, and observations of key exchanges also have their place. Changes in quantitative data may also exhibit patterns shaped by self-organizing action. **Data collection and analysis are often very connected in this design.** For example, shared reflection and dialogue can serve as both data collection and analysis techniques.

---

17 Notes that a difference between the *Exploratory* and Self-organizing situations is that in the self-organizing situations a new set of patterns begins to form and needs to be followed over time.

Certain patterns may only be evident by looking across several sites that are at different stages of development. Since patterns develop over time, historical data can help show patterns that can be analyzed for underlying rules/behaviors that possibly could be shifted by the participants to create new patterns. Patterns are shaped by boundaries, diversity, and relationships that create meaning across time and space (Eoyang, 2001).

Listening to stories can be an effective technique for understanding the patterns and dynamics in a situation. By interviewing those with a long history and deep familiarity with the situation (wise practitioners), evaluators gain an historical perspective and insight into motivations. Some informants may be especially good at reflecting on patterns in the situation while others can provide concrete data on happenings/events. The stories of "newcomers" can help depict the current reality. Through these stories, evaluators identify patterns of boundaries, relationships, and diversity that matter to those who are involved and affected.

D. Making Meaning from Data and Shaping Practice

The analysis strives to identify and understand emerging patterns. Many types of patterns can be identified. Broad patterns of behavior may be identified by contrasting them to patterns in nature. For example, a natural life cycle would be birth, maturity, destruction, and renewal. A pattern might be that certain work of the grantees stops at the maturity stage and is not renewed because those involved resist the destruction phase.

Patterns of interaction within and across groups illustrate another type of pattern. A stretch and fold pattern (like kneading bread) can be seen in human interactions. Some conversation may stretch the participants by building tension, revealing differences, incorporating diversity, and bringing in new energy. Communication may be difficult during stretching, but growth and opportunity arise. Folding reveals common values and releases the tension of the stretch. New connections are consolidated and change settles in. A movement back and forth between stretch and fold may help describe relationships within or among groups. *(See Zimmerman, Lindberg, & Plsek (2001) and Eoyang (1997, 2007) for more information on patterns.)*

It also is useful to look at similarities and differences in patterns within different parts of a system. For example, the analysis might look at how dominant/subordinate patterns compare within principal-teacher, teacher-teacher, and teacher-student relationships. What does this imply about how coherent a system is?

The data analysis seeks patterns that help to deepen understanding about what is happening in the self-organizing units and how it relates to the intentions of the initiative. For example, there may be patterns of interactions that indicate how people within partnerships shift from a pattern of

dominant and subordinate roles to more equal contributions and responsibilities. When people become aware of the conscious or unconscious rules by which they operate, they may change their behaviors to more equitable interactions which, in turn, change the patterns.

The purpose of studying these situations is not necessarily to help those involved in change efforts to put specific conditions in place that are likely to lead to certain outcomes. Rather, the evaluation may well aim to help those involved in situations to handle their anxiety as they question long-standing beliefs, assumptions, or interpretations; stimulate conversations and spark diversity; and ponder and test alternative positions (Stacey, 2000). Self-organizing patterns can spark ideas of how leaders might encourage shared principles (also referred to as simple rules) that may influence patterns.

As patterns shift (or if shifts in patterns are attempted), the *Self-organizing* design attends to how those involved recognize that something new is emerging that is worth trying to stabilize and incorporate into their systems or recognize that it is not controllable. Evaluators also seek to identify patterns that run counter to the desired direction of the initiative to help leaders better understand a wide range of dynamics involved in the initiative.

The understandings that come from an evaluation with a *Self-organizing* design are primarily fed back to those involved in the self-organizing units (as illustrated in the previous Sidebar). With this knowledge they can develop a deeper understanding of what to pay attention to as they continue to adjust to one another and are able to feel comfortable with continued paradox and uncertainty within the initiative.

Case-In-Point

FAS *Self-organizing* Evaluation
Part 4 – Making Meaning and Shaping Practices

(Design is under consideration by FAS. This is a hypothetical example of what the findings might be.)

Making Meaning:

The self-organizing patterns the evaluators identify from their data are multiple. For example:

1. Participants from underserved rural areas begin working with those from low-income urban areas facing similar "lack-of-access to locally grown, healthy foods" issues;
2. Participants active in one aspect of the market change to begin working with participants active in other aspects of the local food supply chain;
3. State-based groups working on various policy efforts recognize their commonalities, and begin exchanging strategies and discussing collective efforts around federal policies. As the participants develop trusting relationships with others, they begin together to generate new ideas about how to enact the FAS vision.

Shaping Practice:

The initiative evaluators provide feedback at later conferences about patterns they are seeing. A round of Open Space conversations focuses on the emerging patterns. The participants discuss how to amplify desired patterns and dampen less desirable ones.

All parties – grantees, initiative leaders, initiative evaluators, project evaluators, and others – talk about ways they have or might adjust their work based on the data and the feedback sessions. For example, the initiative evaluators change some of the questions they use to learn more about the role of technology in building the relationships among participants between meetings.

The information also can be used to identify what rules are operating consciously or unconsciously in the system (e.g., dominant-subordinate relationship) and what changes in operating assumptions might break old patterns and/or establish new desired ones. The findings also are useful for determining what types of feedback processes are especially important for the initiative. A deepening understanding may even result in a new definition of the overall purposes and goals of the initiative.

Initiative leaders use results of *Self-organizing* evaluations differently than results from a *Predictive* design. The data from a *Predictive* design suggest factors and actions to plan or control. The initiative leaders may use the understanding from a *Self-organizing* design to attend to their own ways of relating to others or to encourage the use of new ways of thinking or behaving in the collective action of participants. They do this knowing that the participants' consideration of these ideas may result in new and unexpected patterns.

In the FAS example, the initiative leaders might bring more grantees to the conferences to expand the differences or create some subgroups within this larger group. This adjustment can reduce or highlight certain differences that might stimulate movement toward the goal or build the energy in the group to move forward more quickly.

IV. *Initiative Renewal* Evaluation Design

A. Introduction

This evaluation design emphasizes the continual renewal and sustainability of an initiative as it increases its flexibility in some areas; becomes more predictable and controlled in other areas; and undergoes major transformations, produces results, and builds unexpected connections and conditions.

The evaluation encourages big-picture adjustment and longevity as the initiative's strategies (and, perhaps, the direction) change over time. **This design takes into account the multiple dynamics in operation** in the initiative as well as the larger context.

This type of evaluation helps initiative leaders, evaluators, and grantees periodically reflect on their overall progress and determine if redesign of aspects of the initiative, its evaluation, and/or its theory of change is needed. In conducting the evaluation, evaluators move back and forth between an insider and an outsider perspective.

B. Designing an *Initiative Renewal* Evaluation

Case-In-Point

FAS *Initiative Renewal* Evaluation
Part 1 – General Information

(Design is under consideration by FAS for 2009)

2009 A.D. – Now, three years since FAS articulated a theory of change (see Appendix A), it is time to step back and reflect on progress, the theory, and the initiative evaluation designs. Important changes are occurring within the initiative and larger social context (e.g., the growing attention to global warming and the demand for ethanol and its impact on commodity prices).

The initiative leaders and evaluators use an Initiative Renewal evaluation design to identify:

a. possible adjustments in the theory of change (e.g., how to include the interplay of self-organizing and planned change) and

b. the nature of a long-term monitoring system for movement toward the FAS vision.

Although work is under way on measures of the theory's components and dynamics, the monitoring system is not yet usable for *Predictive* evaluations.

Evaluation Questions Addressed

The evaluation questions addressed by an *Initiative Renewal* evaluation design focus on the interplay of multiple dynamics of change in the initiative and with its context. The questions are ones such as: What has been learned that can enrich the initiative's theory of change? How can the initiative be sustained? What modifications are needed in the intentions and strategies of the initiative? How does the evaluation design(s) need to be adjusted?

Situations Where This Design Is Used

The *Initiative Renewal* design focuses on how to renew the initiative through intentional changes in the initiative, its theory of change, and/or its evaluations. For example, in regard to evaluation, if certain aspects of an initiative are stabilizing in desirable, planned, and predictable ways, it may be useful to institute a new *Predictive* design to provide a clearer picture of the extent and ways in which the initiative is meeting planned outcomes.

If the unorganized dynamic seems to be expanding, either because of environmental shifts or initiative effects, it may be useful to emphasize an *Exploratory* design for that portion. If new patterns are emerging, changes may be needed in the *Self-organizing* evaluation design(s). This design is appropriate after significant shifts have occurred in the initiative itself or as evaluations of parts of the initiative are completed.

Once an initiative has been in operation for some time and initiative leaders and participants have received feedback from evaluations and other sources, they develop an increasingly sophisticated understanding of the mix of organized, self-organizing, and unorganized dynamics within the complex web of interacting systems of the initiative. They begin to see ways to change boundaries, relationships, and diversity of many types to stimulate different dynamics and bring coherence where needed.

Initiatives typically involve a major shift in perspective or philosophy. An *Initiative Renewal* evaluation design encourages **collective reflection** on whether and why those shifts are occurring and on the implications for the design of the

initiative and its evaluation. Shifts in initiative philosophy can sometimes stimulate rapid and transformative change (Bak, 1999). These points of transformation happen unexpectedly and are referred to as tipping points.[18] Prior to the tipping point, people struggle to change their perspectives. After the tipping point, the new patterns spread more easily. Patterns that exist prior to and after the tipping point are different. By recognizing these shifts, the evaluator can adapt to unexpected change and keep the evaluation focused on the patterns, activities, and results of most importance at any point in time.

This design also is useful when the results, activities, and/or theory of the initiative have become too ambiguous or too rigid. If the initiative is becoming too ambiguous, it may be helpful for the initiative to use more clear processes and outcomes that can be evaluated with the *Predictive* design. If things are getting too rigid, a more emergent approach may be needed to explore new, out-of-the-box possibilities.

By putting a greater emphasis on a *Self-organizing* or *Exploratory* design, the evaluation can help open the space for creative thinking and relating by stakeholders. This can build up the energy in the system to keep participants engaged. After the initiative has been in existence for quite a few years or as it is coming to an end, an *Initiative Renewal* evaluation design can be used to consider how evaluation capacity and feedback mechanisms can be sustained or how the initiative can be concluded or transformed into more informal contextualized activities.

Case-In-Point

FAS *Initiative Renewal* Evaluation Part 2 – Evaluation Design

(Design is under consideration by FAS)

The initiative evaluators and leaders decide to convene an Initiative Renewal evaluation summit. The summit participants are teams from grantee sites; experts in youth leadership, market-based strategies, policy, and communication; project evaluators who work with multiple grantees; evaluators of other large initiatives; and other relevant consultants and intermediaries.

Project/grantee leaders are asked to include team members who are insightful about the bigger context of their work and the relationships depicted in the theory of change.

Prior to the summit, participants receive summaries of evaluation findings and other data about components of the theory of change.

The findings include results from the *Predictive* and *Self-organizing* evaluation designs (described above) as well as from several new *Exploratory* evaluations.

For example, a policy-focused *Exploratory* evaluation addresses the questions: What public (local, state, and federal) policies appear to favor Good Food-focused community-owned and community-based enterprises? What seem to be the drivers in the creation and implementation of such policies?

A market-oriented *Exploratory* evaluation addresses the question: What is stimulating demand for Good Food, i.e., what are the values-based and policy drivers influencing demand – health, localness, etc.?

18 See Gladwell (2002).

Considerations in Planning the *Initiative Renewal* Evaluation

Planning an *Initiative Renewal* evaluation involves **getting multiple perspectives on the focus of the initiative and using cumulative data about the initiative's theory of change.** The *Initiative Renewal* evaluation may be facilitated by a mixed team of initiative leaders, initiative evaluators, grantees, and others who have not been previously involved.

The evaluation is likely to involve convening one or more meetings after data from other evaluation designs are available. The design/planning activities involve determining what evaluation findings and other information to provide participants about their roles and the processes of their interactions.

A "summit" style meeting of participants with multiple perspectives on key questions about the initiative is a likely evaluation process. Other means of bringing together multiple perspectives to look at findings from multiple evaluative studies might include a Delphi process,[19] and/or Internet-facilitated meetings/conversations. Participants might include grantees; policymakers; project evaluators; intermediaries; experts in various disciplines (e.g., disciplines ranging from scientific disciplines to communications); evaluators who have not been part of the work thus far; and other relevant parties. They would review what shifts have occurred in the initiative, whether the accumulating data support the current theory of change, and how the evaluation designs individually and collectively can be realigned to best support renewal of the initiative.

If a summit is used, it might include large-group processes such as Open Space Technology (Owen, 1997); World Café (Brown, 2005); Appreciative Inquiry (Watkins & Mohr, 2001); and Future Search (Weisbord, Janoff, & Weisbord, 2000). The summit preparatory activities depend on the familiarity of the participants with the initiative. The type of data and findings reviewed by the group and the type of questions asked could cover a wide spectrum.

The follow-up activities are shaped by the nature of the questions being asked, the complexity of the initiative, future funding plans, and other such factors.

19 A Delphi process involves giving respondents the results from one round of data collection in preparation for a second round of data collection.

C. Planning and Engaging in Data Collection

During this phase, evaluators conduct meta-analysis of data from large-scale databases about initiatives to generate new data. **They draw on multiple research and theory bases.**

The data collection during a summit or other setting **focuses on key issues about the initiative as a whole within its larger context as well as patterns, dynamics, accomplishments, and actions within the initiative.** The data collection and meaning making are closely linked. Interactive large-group processes where data are generated, synthesized, and interpreted progressively in the same setting are useful.

For example, a World Café format might be used. In this format, participants talk in groups of four around a small table with a paper tablecloth for writing notes. The groups focus on predetermined questions. After 20 or 30 minutes, three people leave the table to move to other tables. The fourth serves as the host as new people come to the table. The host summarizes the previous conversation and the new group picks up from there. After a few (three or so) rounds, the full group engages in a process to summarize the conversations.

On the following page, Table 2 gives examples of questions that participants in an *Initiative Renewal* evaluation might be asked after having reviewed data about the initiative. The questions cover topics that are important in rethinking the theory of change, evaluation designs, and initiative strategies and outcomes.

Case-In-Point

FAS *Initiative Renewal* Evaluation
Part 3 – Data Collection

(Design is under consideration by FAS)

The initiative leaders and evaluators set the context and review evaluation highlights and other data framed around the theory of change. Then the evaluation summit participants engage in a series of small group conversations using a World Café format (see text for description).

The first set of conversations is designed to stretch people's thinking about the initiative. The conversations focus on topics about boundaries, relationships, and diversity. (Topics are selected from the list in Table 2.)

Later in the day, a second series of conversations directly address the components and dynamics of the theory of change.

In these conversations, participants apply the ideas generated in the first series of conversations to the specifics of the theory of change and sustainability of the initiative. They also consider implications for the long-term monitoring system.

Topics	*Initiative Renewal* Evaluation: Data Collection and Analysis Questions
Boundaries	• What boundaries are changing or need to change between the initiative and its context? • What is defining the boundaries of the predictable, self-organizing, and unorganized dynamics of the work? How have the boundaries shifted over the life of the initiative? • What boundary shifts would bring new creative or stabilizing energy into the initiative?
Relationships	• What do data from the various evaluation designs and other sources show about the relationships among the components of the theory of change? • What relationships are producing meaningful results or conditions? • What relationships are languishing? Are these relationships needed or should they be left to languish? • What new relationships are needed and to what end? • What feedback loops are built into the system? • How does the initiative evaluation complement other feedback loops? • What roles are feedback loops playing? How are feedback patterns changing over time?
Diversity/ Differences/ Energy	• Is there sufficient diversity in the self-organizing units to stimulate creative ways of addressing challenging issues? • What are the important power dynamics in the initiative or between the initiative and the context? • What new definitions of diversity may help provide the basis for stimulating movement in the desired direction through self-organizing means? • Where is energy being created? Where is it stagnating? • Where could scattered energy be consolidated and directed toward the desired end?
Perspectives	• Are new perspectives (purposes) arising within the initiative? • Are changes occurring in whose expertise is valued? • What shifts are occurring in the acceptance and spread of the new philosophies of the initiative? • Are stakeholders not considering an important perspective?
Unanticipated Consequences and Concerns	• What, if any, unanticipated consequences are beginning to accumulate that need to be addressed? Are they desired consequences? Who is being harmed/victimized by these unanticipated consequences? • What didn't happen that you expected to happen? What did happen that you did not expect to happen? • What is concerning key stakeholders? Are key stakeholders getting tired? What are they complaining about? What is the volume of concerns? How varied are they? Are all voices given space to raise their concerns? • What new conditions are created as a result of both the anticipated and unanticipated consequences? Do the new conditions increase/decrease the predictable, self-organizing, and/or unorganized territories of the system?
Focus, Outcomes, and Resources	• Do the identified outcomes remain as priorities? • Have the criteria for success changed? • What new desired outcomes or directions have been identified? • What are the few important things to focus on for the next phase of the evaluation? • How stable is the funding for the work to continue toward the goal or in a desired manner? • What shifts are occurring in the resources needed to continue to move toward the goal?

Table 2. Examples of Data Collection and Analysis Questions for an Initiative Renewal Evaluation

D. Making Meaning from Data and Shaping Practice

As noted above, the data gathering and making-meaning tasks may be intertwined in the *Initiative Renewal* design. **The evaluation focuses on broad shifts across the whole domain of the initiative and/or its link to context as well as entangled parts and systems** within the initiative. A summit or other structure can serve as both the data collection and the making-meaning session and include discussion of existing analyses. Group processes such as the *World Café, Open Space Technology, Appreciative Inquiry, or Future Search* may be appropriate.

The report of an *Initiative Renewal* evaluation is likely to provide multiple perspectives and ideas for changes to make in the initiative, its theory of change, and its evaluation along with the rationale for changes. Graphic recording of the meetings may be an effective way to summarize some of the results of the *Initiative Renewal* evaluation to share with others.

The initiative leaders use the results to rebalance emphases in the overall initiative; stimulate and/or track self-organizing units within the initiative; redesign their theory of change; and/or build new relationships. **The initiative leaders and evaluators work together to redesign the evaluation plan.** It may include directing attention to new aspects of the initiative.

Case-In-Point

FAS *Initiative Renewal* Evaluation Part 4 – Making Meaning and Shaping Practices

(Design is under consideration by FAS)

Making Meaning:

The meaning-making activities are interspersed with the data collection throughout the summit. For example, during the summary session for each World Café conversation series, participants identify major themes and patterns concerning the theory of change and monitoring system.

After the summit, a smaller group further analyzes the information to determine key considerations for addressing the overall purposes of the *Initiative Renewal* evaluation.

Shaping Practices:

Initiative leaders and a consultant prepare an updated theory of change and strategy paper for guiding the next phase of the initiative.

Initiative evaluators develop an evaluation plan for the future. Another consultant refines the design of the long-term monitoring system.

These documents then serve as the basis for the initiative leaders to make decisions about grantmaking, technical assistance to grantees, and other practical decisions. Similarly they guide the practical decisions of the initiative evaluators.

Summary of Features of the Four Evaluation Designs

To recap, this guide has outlined theoretical and practical aspects of four evaluation designs. Each design has a different focus, with that focus being congruent with the different dynamics of the initiative being considered.

- **The *Exploratory* evaluation design** investigates the unorganized dynamics of an initiative. It focuses on potentially important components and dynamics of change that are not yet delineated in the initiative's theory of change.
- **The *Predictive* evaluation design** is used where the dynamics are organized. It studies cause-and-effect relationships between the structured interventions of an initiative and the predicted outcomes/changes.
- **The *Self-organizing* evaluation design** addresses initiatives with self-organizing dynamics, looking for the patterns of change emerging from these self-organizing dynamics within the initiative.
- **The *Initiative Renewal* evaluation design** reviews the whole initiative in its context. It looks at the interplay of multiple dynamics of change within the initiative and its context that enrich the initiative's theory of change and implications for sustaining the intentions of the initiative.

The following chart (see Table 3) provides a summary of how the phases of the evaluation process vary among the evaluation designs. These are not rigid categorizations but rather general tendencies among the designs.

.

Evaluation Phases	Features Within Each Type of Evaluation Design			
	Exploratory	***Predictive***	***Self-organizing***	***Initiative Renewal***
Designing the Evaluation Clarify theory of change or its status; identify meaningful units within initiative for evaluation; match characteristics of initiative units and evaluation orientations.	• Identify aspects of the initiative for which a theory of change does not yet exist, places where predictable relationships might be hidden, and/or where little/no agreement exists about how to move in a desired direction. • Conduct evaluation from either insider or outsider perspective.	• Identify specific desired outcomes of selected activities of initiative. • Identify intended, predictable links between selected activities and outcomes. • Clarify key features of activity serving as intervention. • Determine sites where predictable relationship is meaningfully investigated. • Conduct evaluation from outsider perspective.	• Identify aspects of the initiative where self-organizing patterns are likely to be a major force in shaping initiative. • Identify units of self-organizing by considering boundaries, diversity, perspectives, and relationships. • Conduct evaluation from insider perspective.	• Consider the whole initiative, its parts, and its context. • Identify multiple perspectives that can provide insights on dynamics within and around the initiative that seem related to revitalizing the initiative. • Move back and forth between insider and outsider perspective in conducting evaluation.
Planning and Engaging in Data Collection Gather data with attention to aligning data collection and analysis methods.	• Obtain data from site participants, specialists, cutting edge thinkers, activists, and researchers. • Use mostly, but not exclusively, qualitative measures/methods and existing data bases. • Often use very open-ended data collection tools. • Consider focusing on distinct perspectives to understand and contrast them.	• Identify/develop measures of intended outcomes (usually consistent measures across locations). • Collect data regarding initial conditions, interventions, and outcomes in a standardized manner. • Primarily use data sources from within participating sites. • Use mostly, but not exclusively, quantitative measures.	• Use participants engaged in self-organizing processes as data sources. • Primarily use qualitative measures/methods (e.g., individual and group interviews, group engagement processes, questionnaires).	• Obtain data from multiple perspectives and aspects of the initiative (e.g., composite of evaluation data from other designs). • Draw on multiple research and theory bases. • Attend to boundaries, relationships, diversity, values, and perspectives. • Conduct meta-analysis of data from large-scale databases about initiatives to generate new data. • Use group engagement processes to collect and synthesize data.
Making Meaning and Shaping Practice Data analysis, synthesis, interpretation, and use of results to enhance the initiative.	• Use theories from multiple fields. • Engage multiple stakeholders, theorists in interpreting the data. • Present results as tentative ideas to test out. • Disseminate results primarily to those immediately involved in initiative. • Often involve extensive interaction between evaluator and user. • Use results to shape further conversations and/or to shape pilot studies that employ either a *Predictive* or *Self-organizing* design.	• Often use quantitative statistical analyses. • Choose analysis methods that correspond to assumptions about linear and/or nonlinear relationships. • May make interpretations about generalizations and size of effect. • Write report as stand-alone document. • Often involve limited in-person interactions between evaluator and user. • Use report as basis for policy action, adoption of practices, studies by others, and accountability for results.	• Use qualitative analysis and synthesis methods largely with emphasis on finding/following patterns over time or space. • Involve participants in interpretation, (i.e., use interactive methods of meaning making). • Blend interpretation and meaning making with use and sharing. • Determine findings that are appropriate for shaping initiative's general practices, although most findings are for local use of participants.	• Often use group processes for analysis and meaning making. • Pay attention to change over time and space. • Contrast findings to relevant theories about stages and/or patterns of change. • Emphasize implications for initiative's activities, theory of change, and adjustments in evaluation designs. • Focus on broad shifts across whole terrain of the initiative and/or its link to context as well as entangled parts and systems within initiative.

Table 3. Summary of Differences Among Phases of Evaluation for Four Evaluation Designs

Summary

The goal of this document is to help WKKF initiative evaluators expand their thinking about approaches to initiative evaluation. It considers four evaluation designs that relate to differences in the dynamics of social systems within initiatives. The document provides general information about using the designs. Furthermore, it is intended to open conversation about these designs and will be revised and updated based on the evaluators' experience in using them. The designs are complementary and can be employed simultaneously to investigate different aspects of an initiative.

The conceptual framework serving as the basis for the four evaluation designs considers how the amount of *certainty* and *agreement* in situations shapes the dynamics of complex systems such as WKKF initiatives. These factors create three types of dynamics within social systems: unorganized (low certainty and low agreement); organized (high certainty and high agreement); and self-organizing (moderate certainty and moderate agreement). It is important to remember that, in a complex social system or collection of systems, the dynamics of unorganized, organized, and self-organizing are intertwined. However, this document focuses on these dynamics separately in the designs while keeping the others in our peripheral vision.

The system dynamics undergirding the ***Exploratory* design** (unorganized dynamics) and ***Predictive* design** (organized dynamics) are currently the most familiar and, consequently, these two types of designs are commonly used within WFFK initiative evaluations. The significance of the system dynamics underlying the ***Self-organizing* design** is relatively new to the evaluation field, but the forces that arise from the actions of those involved in the initiative independently of control by initiative or project leaders can have surprising implications for the initiative. The design opens up possibilities for describing and understanding these emergent situations so that leaders, funders, and participants can respond to them.

The ***Initiative Renewal*** design provides a means of looking at the initiative as a whole within its context while not attempting to evaluate all aspects of the initiative. It stresses the importance of periodically stepping back and reviewing the initiative, its theory of change, its evaluation, and its shifting context to determine where adjustments are needed.

Some of the same data-gathering techniques (e.g., interviews, case studies, questionnaires) are used to look at structures with unorganized, organized, and self-organizing dynamics. However, quantitative data tends to be used more heavily in *Predictive* designs while qualitative data is used more frequently in the *Exploratory* and *Self-organizing* designs. Other differences among the designs lie in the specific data gathered; the premises underlying how the data are analyzed; and the uses of the results.

Evaluators use the *Exploratory* design when the dynamics of the situation are so random and unorganized that they cannot focus on patterns or predictable causal relationships. In these situations they do not have a conceptual framework for understanding the situation. Initiative evaluators use a *Predictive* design when the dynamics of the situation are stable enough to focus on intended outcomes and how planned activities contribute to the accomplishment of these intended outcomes.

Evaluators use the *Self-organizing* conceptual framework where certainty or agreement is moderate. They watch for evidence of self-organizing behavior and look for patterns that indicate new connections are being made and/or creative possibilities are emerging.

In deciding how to strategize the evaluation design, it's important for evaluators and initiative leaders to remember that they can choose different designs for different aspects of the initiative, and they do not necessarily have to evaluate the initiative as a whole.

WKKF is committed to ensuring real change and, to that end, encouraging evaluation designs that are forward-thinking and useful in different and changing contexts. It is hoped that this document not only helps initiative evaluators see ways to discover and encourage new possibilities through the application of these evaluation designs, but also encourages them to **share their discoveries with one another and with WKKF** so this document can be updated and refined.

References

Brown, J. (2005). *The world cafe: Shaping our futures through conversations that matter.* San Francisco: Berrett-Koehler.

Checkland, P. (1999). *Systems thinking, systems practice.* New York: Wiley.

Eoyang, G. (1997). *Coping with chaos: Seven simple tools.* Circle Pines, MN: Lagumo.

Eoyang, G. (2001). *Conditions for self-organizing in human systems.* Unpublished doctoral dissertation, Union Institute and University, Cincinnati, Ohio.

Eoyang, G. (2007). *Attractors.* Circle Pines, MN: Human Systems Dynamics Institute. Vol. 4.3.

Gladwell, M. (2002). The tipping point: *How little things can make a big difference.* Boston, MA: Back Bay.

Holladay, R. (2003). "Using Human Systems Dynamics in Supervisory Relationships. "In G. Eoyang (ed.), *Voices from the field: An introduction to human systems dynamics*. Circle Pines, MN: Human Systems Dynamics Institute.

Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities and software*. New York: Penguin Books.

Kachigan, S.K. (1986). *Statistical analysis: An interdisciplinary introduction to univariate & multivariate methods*. New York: Radius Press.

Midgley, G. (ed.) (2003). *Systems thinking.* Thousand Oaks, CA: Sage.

Mohr, B. & Watkins, J. (2002). *The Essentials of appreciative inquiry: A roadmap for creating positive futures.* Waltham, MA: Pegasus Communications.

Olson, E. & Eoyang, G. (2001). *Facilitating organization change.* San Francisco: Jossey-Bass/Pfeiffer.

Owen, H. (1997). *Open space technology: A user's guide.* San Francisco: Berrett-Koehler.

Parsons, B. (2002). *Evaluative inquiry: Using evaluation to promote student success.* Thousand Oaks, CA: Corwin.

Parsons, B. (2007). "The State of Methods and Tools for Social Systems Change." *In American Journal of Community Psychology.* 39(3), 405-409. New York: SpringerLink.

Patton, M. (1997). *Utilization-focused evaluation.* Thousand Oaks, CA: Sage.

Patton, M. (2002). *Qualitative research and evaluation methods.* Thousand Oaks, CA: Sage.

Senge, P. (1990). *The fifth discipline.* New York: Doubleday/Currency.

Senge, P., Kleiner, A., Roberts, C., Ross, R., & Smith, B. (1994). *The fifth discipline fieldbook.* New York: Doubleday.

Stacey, R. (1996). *Strategic management and organisational dynamics.* 2nd edition. London: Pitman Publishing.

Stacey, R. (2000). *Strategic management and organisational dynamics.* 3rd edition. London: Pearson Education.

Stacey, R. (2007). *Strategic management and organisational dynamics.* 5th edition. London: Pearson Education.

Watkins, J. M. & Mohr, B. J. (2001). *Appreciative inquiry: Change at the speed of imagination.* San Francisco: Jossey-Bass.

Weisbord, M. & Janoff, S. (1999). *Collaborating for change: Future search.* San Francisco: Barrett-Koehler.

Williams, B. and Imam, I. (ed.). (2007). *Systems concepts in evaluation: An expert anthology.* Inverness, CA: EdgePress.

Zimmerman, B., Lindberg, C., & Plsek, P. (2001) *Edgeware: Insights from complexity sciences for health care leaders.* Irving TX: VHA, Inc.

Appendix A

The Food and Society (FAS) Initiative: An Example Initiative

General Description

Food and Society (FAS) is a large and complex initiative with more than 104 grants for 67 different projects totaling nearly $48 million in the first five years. The initiative is scheduled to operate at a similar level for another five years. FAS has within it many of the strategies, types of activities, and values that are common to other initiatives at WKKF – collaborations, policy, multiple strategies linked to a larger and shared purpose, use of intermediaries, grounded in community, and above all focused on helping people help themselves.

Vision

Launched in 2000, Food and Society is based on a **vision** of a future food system that provides all segments of society a safe and nutritious food supply grown in a manner that protects health and the environment and adds economic and social value to rural and urban communities. Informed by the cluster evaluation of its first five years, the purpose and some of its outcomes have changed; however, the initial vision remains.

About 30 to 40 percent of the funding from foundations for this type of work in the U.S. comes from WKKF.

Intended Outcomes

The following were the outcomes identified at the beginning of the initiative:

a. **Policy.** Public and institutional policies that support sustainable food and agriculture-based enterprises as vehicles for community, social, and economic development.

b. **Scholarship**. Broadened agenda for scholarship in higher education institutions to include engagement with communities and partners to support community-based food systems.

c. **Farms/Acreage**. Increased number of farms/acreage that use environmentally sound agriculture systems.

d. **Markets**. Increased number of economically successful food-related enterprises that are locally owned and controlled, environmentally sound, and health promoting.

e. **Partners**. Increased number of funders and partners supporting community-based food systems approaches.

With the cluster evaluation of the first five years completed, in 2006 FAS established a concrete goal, revised some of its outcomes, dropped others, and added some new ones:

a. **Goal**. Stimulate a shift that will result in at least 10 percent of the U.S. food system functioning in alignment with the FAS vision. (It is currently about 2 percent.)

b. **Additional and Revised Key Outcomes**

Communications. Encourage the growing cultural shift about food and its connection to health, community, and the environment.

Research and Education. Provide the best information available to guide planning and actions in the creation of sustainable, community-based food systems.

Market for Good Food. Support strong and resilient models of community-owned and community-based enterprises.

Youth Involvement. Engage youth as active leaders in their communities and in the national movement for sustainable, community-based food systems.

Policies that Contribute to Good Food Change. Attract new voices to inform the establishment and implementation of public policies that reward public and private actors who support community-based and sustainable food systems.

Community and Stakeholder Involvement. Connect diverse stakeholders across age, gender, race, class, and sector (private, public, academic) lines to engage in projects that accelerate innovation.

Grantees

Grantees for the initiative – all nonprofit organizations – include national, regional, state, and locally based organizations. Individually, they vary in their primary areas of emphasis with the grantees collectively catalyzing change through means including policy, communication, technical assistance, research, and facilitation of organizational change. The average grant size was about $450,000 during the first five years.

Most grantees focus on some rather than all aspects of the vision, e.g., some focus on environmental issues, some on economic issues, some on food safety. Grants are made throughout the life of the initiative (i.e., not all the grants were made at the beginning of the initiative). Here are examples of FAS grantees.

Grantee or Project	Project Purpose
Appalachian Sustainable Development	Build a more self-reliant, just, and ecologically sound food system in central Appalachia as a model for other rural communities.
Chicago Food Systems Project (CFSP)	Address the issue of community food security as it affects low-income communities by bringing healthy foods to the community through a food co-op served by local farms.
Farm and Food Policy Project	Collaborative of 5 grantees – American Farmland Trust, Community Food Security Coalition, Environmental Defense, Northeast-Midwest Institute, and Rural Coalition – to inform the 2007 Farm Bill on issues critical to achieving the FAS vision.
First Nations Development Institute	Assist Native communities in building their agricultural assets while also building capacity within Native communities for systemic change that results in more healthy and sustainable community food systems.
Food Project, Inc.	Expand a program that builds self-confidence, teaches work skills, develops leadership, and increases civic awareness for youth and shares effective practices with other youth-serving organizations.
Iowa State University	Foster the development and growth of value chains in which small and mid-size farmers are rewarded for production practices that have the highest standards of environmental and community stewardship.
Roots of Change Fund (Canopy Institute)	Support the launch of Roots of Change, a funders' collaborative that will expand the resources available for sustainable food systems projects and programs in California.
The Food Alliance	Expand food producer and processor certification capabilities via increased organizational capacities and revised business plan.

Other Features of FAS

FAS is within Food Systems and Rural Development (FSRD), one of the four major programming areas of WKKF. The initiative is led by a program director, supported by other FSRD team members with administrative, communications, policy, and evaluation skills. They are called the "initiative leaders" at WKKF; they are all considered essential to the initiative and are not outside the initiative.

Theory of Change

Underpinning the 2006 FAS goal and outcome revisions is a theory of change, which was not clearly articulated at the outset of FAS. WKKF chose to illustrate the theory of change using a systems dynamic tool – a causal loop diagram (Figure 1).

In its August 2005 FAS strategy paper, WKKF explains the diagram:

- The amount of healthy, fresh, and local foods purchased (HFL PURCHASED)[20] is a key variable to track as we make progress from two to ten percent of the food system functioning in alignment with our vision.
- The amount of HFL PURCHASED will be determined by both individual and institutional consumer demand.

20 Healthy, fresh, local food is sometimes referred to as HFL and at other times as Good Food within the initiative.

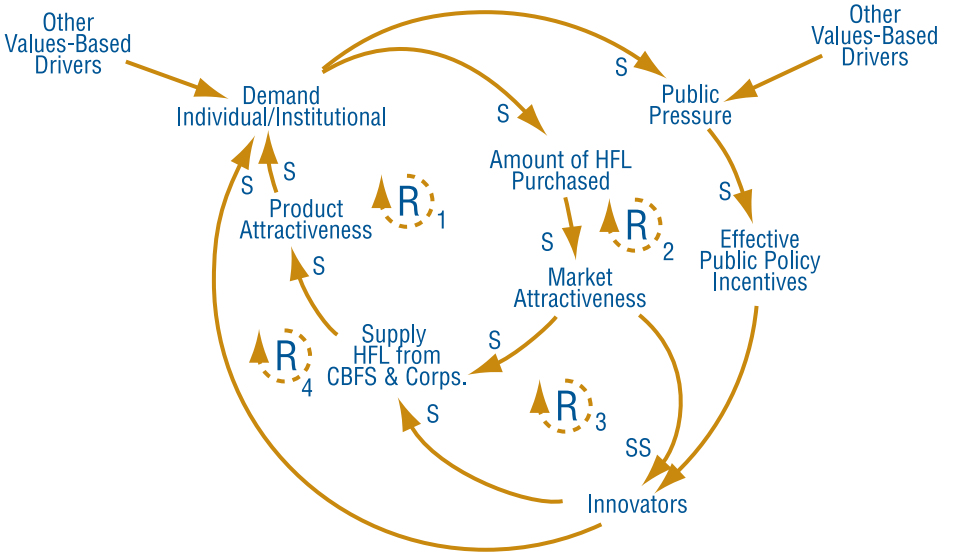


Figure 1. Food systems map – moving from two to ten percent.

- As the amount of HFL PURCHASED increases, we should see an increase in MARKET ATTRACTIVENESS, which incorporates market stability, reliability, and adequate price to attract more farmers, processors, and distributors of "HFL" into the system.
- As MARKET ATTRACTIVENESS increases, we will see an increase in SUPPLY OF HFL from both community-based enterprises and from larger corporations.
- As SUPPLY OF HFL increases, which is another key variable to track, we expect to see an increase in PRODUCT ATTRACTIVENESS, which incorporates product quality, accessibility, and affordability.
- The increase in PRODUCT ATTRACTIVENESS will cause an increase in DEMAND and the entire system will continue to shift toward the Food and Society vision.

This market-based "reinforcing loop" has created the 2 percent of our food system currently in alignment with the Food and Society vision. In order to stimulate a shift that will result in at least 10 percent of the food system in alignment with our vision,

- DEMAND FOR HFL also needs to be translated into increased PUBLIC PRESSURE for more EFFECTIVE PUBLIC POLICY INCENTIVES that will help drive increased INNOVATION in both supply and demand (for example, senior farmers market coupons, community-supported agriculture, and smaller-scale community processing facilities).

The theory of change represented in Figure 1 will increase the percent of the food system functioning in alignment with the Food and Society vision. Many of

the strategies for Food and Society are specifically geared to strengthening the links in Figure 1 (especially between market and policy innovations). However, there is one "blind spot" in this theory of change that also needs awareness and investment in order for the system to truly distribute benefits (economic and nutritional) in a more just and diverse manner.

The blind spot relates to corporate economies of scale. As MARKET ATTRACTIVENESS increases, we will see an increase both in the VIABILITY OF COMMUNITY-BASED FOOD SYSTEMS and in the ADOPTION OF SUSTAINABLE PRACTICES IN THE CORPORATE FOOD SECTOR. These both lead to an increase in SUPPLY OF HFL. However, because of CORPORATE ECONOMIES OF SCALE there will always be a tendency for the VIABILITY OF COMMUNITY-BASED FOOD SYSTEMS to decline as larger corporations enter the supply chain. New innovations in ownership structures are needed for this blind spot to be resolved and for community-based food systems to expand alongside of increased corporate supply. This is why FAS targets resources to the strategy of supporting strong and resilient models of community-owned and community-based enterprise. It is through these community structures that the food system will be able to distribute benefits in a more just and diverse manner.

W.K. KELLOGG
FOUNDATION

One Michigan Avenue East
Battle Creek, MI
49017-4012
USA
269-968-1611
TDD on site
Telex: 4953028
Facsimile: 269-968-0413
Internet: http://www.wkkf.org

EV 4490
Item # 9023
1007 .25M SIS
Printed on recycled paper