DEVELOPMENTAL EVALUATION

Evaluation isn’t something to incorporate only after an intervention, innovation or major change is underway. The very possibility articulated in the idea of making a significant difference in the world ought to incorporate a commitment to not only bringing about meaningful social change, but also thinking deeply about, evaluating, and learning from social innovation as the idea and process develops. However, because evaluation typically carries connotations of narrowly measuring predetermined outcomes achieved through a linear cause-effect intervention, we want to operationalize evaluative thinking in support of social innovation through an approach we call developmental evaluation. Developmental evaluation is designed to be congruent with and nurture developmental, emergent, innovative, and transformative processes.

Helping people learn to think evaluatively can be a more enduring impact from an evaluation than use of specific findings generated in that same evaluation. Findings have a very short ‘half life’ - to use a physical science metaphor. They deteriorate very quickly as the world changes rapidly. Specific findings typically have a small window of relevance. In contrast, learning to think and act evaluatively can have an ongoing impact. The experience of being involved in an evaluation, then, for those actually involved, can have a lasting impact on how they think, on their openness to reality-testing, on how they view the things they do, and on their capacity to engage in innovative processes.

But not all forms of evaluation are helpful. Indeed, many forms of evaluation are the enemy of social innovation. We want to be clear then: We are not suggesting that evaluation is inherently good. Evaluation is no more inherently good than power is inherently corrupt. It’s a matter of how each is used that makes it good or bad, helpful or hurtful, empowering or disempowering. This distinction is especially important at a time and in a world where funders of all kinds are demanding accountability and shouting the virtues of “evidence-based” or “science-based” practice. Social innovators are likely to be ahead of the evidence and in front of the science. Indeed, social science thus far has proved especially inept in offering solutions for the great problems of our time like those noted earlier -- hunger, violence, poverty, hatred. Let’s look, then, at contrasts between traditional and new, more developmental forms of evaluation.

Traditional evaluations… Complexity-based, developmental evaluations…

Render definitive judgments of success or failure

Provide feedback, generate learnings, support direction or affirm changes in direction

Measure success against predetermined goals

Develop new measures and monitoring mechanisms as goals emerge & evolve

Position the evaluator outside to assure independence and objectivity

Position evaluation as an internal, team function integrated into action and ongoing interpretive processes

Design the evaluation based on linear cause-effect logic models

Design the evaluation to capture system dynamics, interdependencies, and emergent interconnections
<table>
<thead>
<tr>
<th>Aim to produce generalizable findings across time &amp; space</th>
<th>Aim to produce context-specific understandings that inform ongoing innovation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accountability focused on and directed to external authorities and funders.</td>
<td>Accountability centered on the innovators’ deep sense of fundamental values and commitments.</td>
</tr>
<tr>
<td>Accountability to control and locate blame for failures</td>
<td>Learning to respond to lack of control and stay in touch with what’s unfolding And thereby respond strategically</td>
</tr>
<tr>
<td>Evaluator determines the design based on the evaluator’s perspective about what is important. The evaluator controls the evaluation.</td>
<td>Evaluator collaborates with those engaged in the change effort to design an evaluation process that matches philosophically and organizationally.</td>
</tr>
<tr>
<td>Evaluation engenders <strong>fear of failure</strong>.</td>
<td>Evaluation supports <strong>hunger for learning</strong>.</td>
</tr>
</tbody>
</table>

In our Social Innovation Think Tank meetings the issue of evaluation arose regularly. We would find ourselves reviewing case examples of major social innovation and inevitably the questions arose: How solid are the data on results and impacts? Can the causal chain between intervention and outcomes be substantiated or even traced? What things didn’t work along the way and how did those involved learn from failure as well as success? We also looked at what the profession of evaluation had to offer in light of complexity theory. That led us to the preceding contrasts.

*Developmental evaluation* refers to long term, partnering relationships between evaluators and those engaged in innovative initiatives and development. Developmental evaluation processes include asking evaluative questions and gathering information to provide feedback and support developmental decision-making and course corrections along the emergent path. The evaluator is part of a team whose members collaborate to conceptualize, design and test new approaches in a long-term, ongoing process of continuous improvement, adaptation and intentional change. The evaluator’s primary function in the team is to elucidate team discussions with evaluative questions, data and logic, and to facilitate data-based assessments of where things are, how are things unfolding, what directions hold promise, what directions ought to be abandoned, what new experiments should be tried – in other words, data-based decision-making in the unfolding and developmental processes of innovation. Adding a complexity perspective to developmental evaluation helps those involved in or leading innovative efforts incorporate rigorous evaluation into their dialogic and decision-making processes as a way of attending to, being mindful about, and monitoring what is emerging. Such social innovators and change agents are committed to grounding their actions in the cold heaven of reality-testing.

Complexity-based, developmental evaluation is decidedly not blame-oriented. Removing blame and judgment from evaluation frees sense and reason to be aimed at the light, the riddled light, for emergent realities are not clear, concrete, and certain. Emergent realities have to be confronted, wrestled with, engaged, played with, until they release interpreted direction and paradoxical or
ambiguous insight – but direction and insight nonetheless. What is this jibberish? The research findings of Sutcliffe and Weber help explain. In a Harvard Business review article entitled “The High Cost of Accurate Knowledge” (2003) they examined the predominant belief in business that managers need accurate and abundant information to carry out their role. They also examined the contrary perspective that, since today's complex information often isn't precise anyway, it's not worth spending too much on data gathering and evaluation. What they concluded from comparing different approaches to using data with variations in performance was that it's not the accuracy and abundance of information that matters most to top executive effectiveness, it's how that information is interpreted. After all, they concluded, the role of senior managers isn't just to make decisions; it's to set direction and motivate others in the face of ambiguities and conflicting demands. Top executives must interpret information and communicate those interpretations -- they must manage meaning as much as they must manage information.

Interpretation in teams works best through an interactive process where various perspectives on the data can be examined openly in the glaring light of evaluation logic: Do the interpretations really follow from the data? What do the data really say? And what can't the data say at all because of inevitable data inadequacies and real world uncertainties.

Interpretation is ongoing and emergent, interpreting both direction and results.. Innovative initiatives have as their purpose the sometimes vague, general notion of systems change through ongoing development. The vagueness of terms like “systems change” and "development" invites the disdain of traditional evaluators who need clear, specific, and measurable objectives to ply their trade. Miller, in The Book of Jargon, defines development as "a vague term used to euphemize large periods of time in which nothing happens" (1981:208). The process is interdependent with the outcome. Change agents in complex systems learn to eschew clear, specific and measurable goals up-front because clarity, specificity and measurability are limiting. They've identified an issue or problem and want to explore some potential solutions or interventions, but they realize that where they end up will be different for different participants in the change process -- and that those participants themselves should play a major role in emergent and iterative goal-setting.

Shorter-term goals may emerge but aren't fixed -- they're milestone targets for assessing progress, subject to change as learning occurs. As a complexity-based, developmental evaluation unfolds, social innovators observe where they are at a moment in time and make adjustments based on dialogue about what's possible and what's desirable, though the criteria for what's "desirable" may be quite situational and always subject to change.

Summative judgment about a stable and fixed program intervention is traditionally the ultimate purpose of evaluation. Summative evaluation makes a judgment of merit or worth based on efficient goal attainment, replicability, clarity of causal specificity, and generalizability. None of these traditional criteria are appropriate or even meaningful for highly volatile environments, systems-change-oriented interventions, and emergent social innovation. Developmentally-oriented leaders in organizations and programs don't expect (or even want) to reach the state of "stabilization" required for summative evaluation. Staff in such efforts don't aim for a steady state of programming because they're constantly tinkering as participants, conditions, learnings and context change. They don't aspire to arrive at a fixed model that can be generalized and disseminated. At most, they may discover and articulate principles of intervention and development, but not a replicable model that says "do X and you'll get Y." Rather, they aspire to continuous progress, ongoing adaptation and rapid responsiveness. No sooner do they articulate and clarify some aspect of the process than that very awareness becomes an intervention and acts to change what they do. They don't value traditional characteristics of summative excellence such as standardization of inputs, consistency of treatment, uniformity of outcomes and
clarity of causal linkages. They assume a world of multiple causes, diversity of outcomes, inconsistency of interventions, interactive effects at every level -- and they find such a world exciting and desirable. They never expect to conduct a summative evaluation because they don't expect the change initiative -- or world -- to hold still long enough for summative review. They expect to be forever developing and changing -- and they want an evaluation approach that supports development and change.

Moreover, they don't conceive of development and change as necessarily improvements. In addition to the connotation that formative evaluation (improvement-oriented evaluation) is ultimately meant to lead to summative evaluation (Scriven, 1991), formative evaluation carries a bias about making something better rather than just making it different. From a complexity-sensitive developmental perspective, you do something different because something has changed -- your understanding, the characteristics of participants, technology, or the world. Those changes are dictated by your latest understandings and perceptions, but the commitment to change doesn't carry a judgment that what was done before was inadequate or less effective. Change is not necessarily progress. Change is adaptation. Assessing the cold reality of change, social innovators can be heard to say:

We did the best we knew how with what we knew and the resources we had. Now we're at a different place in our development -- doing and thinking different things. That's development. That's change. But it's not necessarily improvement. (Jean Gornick, Executive Director, Damiano, Duluth, Minnesota)

The thrust of developmental evaluation as an approach to operationalizing the evaluative thinking mindset involves integrating hope and reality-testing, simultaneously and, perhaps paradoxically, embracing getting-to-maybe optimism and reality-testing skepticism. The next section offers an example of what we mean by integrating hope and reality-testing.

**Hope and Reality-Testing**

In 1977 three Roman Catholic nuns started St. Joseph's House in the inner city of Minneapolis. They were inspired by Dorothy Day's philosophy of “comforting the afflicted and afflicting the comfortable.” The sisters took their passion public and convinced individuals and churches all over the metropolitan area to support them. Over the years thousands of women and kids found compassionate shelter, dozens of volunteers came to the inner city, women and children who were and had been homeless built a community around St. Joe's hospitality, and the sisters became leaders in fighting against violence and injustice. But by the early 1990's their environment had changed. The block surrounding St. Joe's had become the center of a crack cocaine epidemic; drug dealers had claimed the streets; and landlords had abandoned many buildings. St. Joe's guests and families living on the block hid their children inside, police regularly ran through the block with guns drawn, drug dealers and prostitutes (desperate themselves) broke into abandoned buildings. At the north end of the block where two major Minneapolis streets intersected, once-thriving small businesses (gas station, grocery) were abandoned.

This is what Deanna Foster and Mary Keefe faced when they took over the leadership of St. Joe's (now Hope Community, Inc.). They decided to attempt a housing revitalization project and began by trying to talk with local residents, but people were afraid to talk, afraid of the drug dealers and perpetrators of violence. Residents on the block wouldn't even come out to talk. They just said, “We tried for many years and failed... we’re burned out. We’re not going to try again.”
The Hope Community began confronting this reality in light of their vision of a vital, engaged community. They built a playground at their center and renovated a duplex that shared a driveway with the largest drug house. They put fences around the yard and then added fences around the porch because the drug dealers would run through the porch to get around the fences. The drug house was a triplex filled with little children who were terribly neglected and abused. Those little kids were so desperate for something to do that they would climb onto the garage or climb over the fence, anything to try and get to the playground. The playground a crawl tube on the ground that was five or six feet long. They cut a hole in the fence and put the tube through it so that the little kids had their own doorway into our place. The drug dealers would have to embarrass themselves to crawl through the tube -- and some did. But the kids themselves were delighted.

   The kid's private door powerfully expressed the idea that drug dealers are not welcome here, but children are always welcome. One day in desperation we called several of our donors and raised the money to buy that house (it felt like we were paying ransom), and now there's a new duplex it its place.

The purchase of the drug house was not part of some strategic plan. It was an emergent reaction to what they faced on a daily basis.

   St. Joe's leadership became the renamed Hope's leadership, symbolic of their vision. Based on their early success in ridding the community of one major drug house and their long-term commitment to that area, the leaders and community came together to shape a new vision and found support for that vision when a door suddenly opened. They garnered unexpected support from a major philanthropic donor in response to a request they had made.

   We didn’t fully understand at the time, but it really was a unique vote of confidence in Hope. One day, the mail comes, and we open it up, and there’s a hand-written check for $500,000! We put it in the bank and for the next three months I don’t think I slept more than two hours a night. I worried, “How are we going to be good stewards? How will we not waste it?” This serious investment totally called our bluff. We had this big plan, and suddenly someone believed in it and backed up that belief in a big way. We had to refine our own understanding of what our future was going to be, and how we were going to shape it. It’s one thing to have an idea about something, it’s another thing to be responsible for actually nurturing that idea and bringing it forward in a responsible way.

   The door opening brought both terror and delight, sleepless nights and energetic days. They had created a vision for a major community revitalization effort centered around a Children's Village. Hope and vision brought out the skeptics.

   We never said we were going to build the whole thing. Children's Village was a vision. But it shocked people. It really shocked people. Some were pleasantly shocked and then said, “Well, that was fun,” and went on their way. Other people were critical, saying “It’s totally unrealistic and ridiculous for a small organization like Hope to even contemplate. It will never happen.” Everyone picked different parts of it to criticize. Suddenly we were out there in the public eye, and we didn’t know how Children's Village was going to happen. We only knew it would.
They faced the criticism. They faced the critics. But they did so emergently, by finding the flow in the community, facing the daunting reality of what might lie ahead, and began working day to day – acting, monitoring, getting feedback, learning, acting, in a cycle of emergence. In their own words:

We almost had to do it, not backwards, but in alternate order. Normally, when an organization gets half a million dollars they have spent a lot of time in a more linear process thinking through what they are going to do. What is the goal? What is the work plan? What will it cost? Who is the staff? You get the community input, all that stuff, and then have this whopping proposal, right? But it didn’t happen that way at all. It was “Here’s the vision, here’s the money, now, make it happen.”

And the very absence of a traditional linear planning process became a new source of criticism and complaint.

One of the criticisms we get is that we don’t have a linear, goal directed approach. We don’t assume where we are going. We ask: Who’s here? What are people experiencing? What are they believing and hoping? What is their understanding of community? And what is our understanding of all the things we’ve done? We keep trying things, we keep building understanding and building community around ourselves. We are about uncovering, discovering, and creating. It really unfolds itself. It grows organically. It’s just such a natural process.

But it’s more complex than that because, at the same time, there’s a whole set of strategic thinking that’s going on. We also have to ask: Where is the land out there? Where’s the money? What are the opportunities? Where are the potential partners? What are the potential pit falls? How could all this fit together? What would happen if we did this?

While these questions are evaluative in nature, they differed radically from the kinds of “linear, goal directed” questions which would be key to most traditional evaluations. Evaluators speak of “summative” evaluations which are focused on finding out did the program work, were the goals realized, should the program be continued and on setting up the data gathering methods to determine the answers to these questions, early in the process. These questions even differed from the less common but still familiar “formative” evaluations, which are focused on establishing programs strengths and weaknesses and progress, relative to the goals, as it unfolds. Instead, Hope’s leadership focused on an open-ended approach to data gathering, where the questions and concerns were emergent, and where trial and error was carefully mined for learning.

Often we may try things that don’t necessarily succeed on their own, but end up teaching us something and creating other opportunities. We bought a house and ended up selling it a short time later, but we recouped our money, learned about the block the house was on and from that house came one of our best tenant leaders. Another lesson came when we were smaller. We tried having our own construction company, learning quickly about the limits of that strategy and acting accordingly.

A lot of it has to do with intuition, but intuition is not just a thought that comes to you randomly. This intuition grows out of very strategic integrated thinking. We’re constantly operating in this huge matrix of reality. We’re not just focusing on our relationships with people in the
neighborhood and ignoring, for instance, all the real estate developers. People are out there buying and selling real estate, and if you look closely, often ripping people off. But we immersed ourselves in that community because we had to – it was a major part of what was going to impact our neighborhood. We have to deal with the city and the planning department and a multitude of other public agencies. You are constantly immersed in that total picture and informed by it, and then strategically respond to opportunities.

This approach to reality testing took a form different form most evaluations. It defined reality as messy, not orderly, emergent, not controlled and social innovation as an iterative process of experimentation, learning and adaptation. The Hope Community leadership lived out a complexity perspective, seeing and engaging the connections between the micro and macro. They monitored the big picture, the whole picture – national housing, community development, and real estate patterns; interest rates and international finance; government policies, philanthropic funding trends and priorities; research on community revitalization. They had a keen sense of the history of the community. At the same time, they were fully enmeshed in the day-to-day reality of work in the community, including engaging local government inspectors, city planners, social service agencies working in the community, local businesses, and local funders.

**Damiano and Developmental Evaluation**

Damiano is a community-based employment program in Duluth, Minnesota. The Community Kitchens model out of Washington DC provided inspiration and a place to start, but the founders “tweaked” the model to fit Duluth. The goal was to use an established food service program to train and graduate twenty hard-to-employ adults and assist them with employment. That first year the program trained and graduated fourteen people with employment retention statistics of well below sixty percent. The second year they began efforts to move people beyond entry level positions and introduced the concept of career development. Their outcomes changed very little for three years.

Jean Gornick, Damiano’s Executive Director reflects: “In hindsight, Damiano’s early efforts at training people for employment were incredibly naive. So much so, that today I am surprised we experienced even the limited degree of success we did.” She then describes beginning to use developmental evaluation.

It was at that time we became acquainted with useful evaluation practices. A logic model of Opportunities Cooking helped us to recognize that we were making program decisions based on assumptions rather than facts and so we made several changes that greatly improved outcomes for our students.

First and foremost we recognized that our goal was employment and not graduation so we began to use our greatest resource, staff time, in a more productive way. Prior to the change, staff people were spending more than a month a year celebrating graduation (planning, inviting, cooking and decorating). We retained the celebrations, recognizing their importance to people, but scaled them way back. The lead staff person began to spend more time bringing good employers to the table and instituting incentives for work retention. An example of this is the three-month bonus for employment. Mentors and support groups as well as staff support are now provided post graduation.
Our training retention rate was below sixty percent. Using principles of evaluation we improved our intake and orientation procedure. Efforts at screening through a complicated intake procedure were not producing the results we wanted so the procedure was changed. We now use a simplified intake sheet that does a bottom-line screening of math/reading levels, cognitive functioning, and obvious mental health issues. We use a three-week orientation period, with a contract signed the first day of week four. Recognition is given to students who sign the contract and they are asked for feedback on the process and their goals. Focus groups are used with staff and students at the end of each 12-week session.

Equally important is that the new procedures -- and greater success rates -- give both students and staff a greater sense of accomplishment.

The result of these changes is a placement rate in employment of 98% and a work retention rate of 70% after one year in placement.

Interviews with independent leaders and funders in Duluth confirmed that Damiano had come to be regarded as an effective and respected program that is making an incremental but important difference in the community. However, even as the program attained a high level of success, the environment in which it operated was changing. Partly because of its visible success, the program staff became the target of a new labor unionizing effort aimed at non-profits. The recession of 2001-2002 hit Duluth hard. The employment picture changed. Funding available for programs dried up. New government regulations threatened some established operating procedures. The building needed major renovation. Just as success seemed assured, the program found itself faced with new crises in a volatile and uncertain environment. Because the program director was engaged in ongoing developmental evaluation, she was already scanning the environment and identified these trends early. She knew better than to let the program’s past success crystallize into either a recipe (simple problem solution) or formula (complicated problem solution). In an uncertain environment mired in political controversy about the future of employment and social service programs, the program adapted by building new collaborations, seeking new funding partners, and shifting key elements of the way clients entered and worked through the program. They were also able to take an assertive and proactive stance in the face of increased demands for accountability. Rather than lodging accountability in adherence to a recipe or formula, they lodged it in learning and adaptability.

**Accountability**

Complexity-based developmental evaluation shifts the locus and focus of accountability. Traditionally accountability has focused on and been directed to external authorities and funders. But for value-driven social innovators the highest form of accountability is internal. Are we walking the talk? Are we being true to our vision? Are we dealing with reality? Are we connecting the dots between here-and-now reality and our vision? And how would we know? What are we observing that’s different, that’s emerging? These become internalized questions, asked ferociously, continuously, because they want to know.

That doesn’t mean that asking such questions and engaging the answers, as uncertain as they may be, is easy. It takes courage to face the possibility that one is deluding oneself. Here the individual’s sense of internal and personal accountability connects with a group’s sense of collective
responsibility and ultimately connects back to the macro, to engage the question of institutional and societal accountability.


**Bibliography on Evaluation and Systems Thinking**


