MEASURING CHANGE WHILE CHANGING MEASURES:

Learning In, and From, the Evaluation
of *Making Connections*

By Leila Fiester

CONTENTS

INTRODUCTION

By chance and by choice, from the day *Making Connections* began it rode a wave of change that swept through its sponsoring foundation, the field of community change, and the evaluation world. The ambitious, multi-site, decade-long community change effort by the Annie E. Casey Foundation, which started in 1999, aimed for nothing less than improving outcomes for the most vulnerable children by transforming their neighborhoods and helping their parents achieve economic stability, connect with better services and supports, and forge strong social networks. To some degree, in some sites, and for some populations, *Making Connections* achieved many of those goals over the next decade-plus. It generated a wealth of hard-earned knowledge about how to succeed as well as what not to do. And, along the way, it tested the potential and the limitations of tools and strategies available for evaluating community change efforts.

Making Connections' evaluation, which would span eight years and cost almost $60 million, was complex and multidimensional, with many moving parts. This case study focuses on just one slice of the evaluation:[1] measurement choices and challenges. It emphasizes three challenges, among many, that are particularly relevant to evaluations of community change initiatives:[2]

- **How to measure population-level change** for an initiative that seeks community-wide results but often operates at a smaller scale at the program level, with aspirations for change that shift over time;

- **How to frame an evaluation** when it has multiple, evolving, and sometimes competing purposes: measuring outcomes, building local capacity, empowering neighborhood residents, enabling implementers to "manage to results," and enabling the funder to "learn while doing"; and

- **How multiple data needs and uses,** at differing levels of implementation and management and different phases of the work, drive evaluation options and choices.

Making Connections' evaluation struggled with those challenges as the initiative's evaluators, implementers, and managers strove to simultaneously satisfy the need for real-time learning, results-based accountability, and genuine improvements in outcomes for residents of deeply troubled neighborhoods—and to do it all in a constantly fluctuating environment. The story of *Making Connections'* evaluation, therefore, can only be told in the context of Casey's evolution as an institution, the individuals who influenced those changes, and the changing field into which *Making Connections* was born.

ABBREVIATED TIMELINE OF *MAKING CONNECTIONS* AND ITS EVALUATION

Pre-1999	• Foundation relocates from Greenwich to Baltimore • Ralph Smith joins Casey, leads Foundation through strategic planning/restructuring • Start of Casey cross-disciplinary site teams • Design and development of Neighborhood Transformation/Family Development Initiative • Identification and selection of 22 cities for *Making Connections'* exploratory/relationship-building phase
1999	• UPS's IPO doubles Casey endowment • Identification of 10 target neighborhoods • Local Learning Partnerships start to form • Staff begin creating menu of target outcomes • Start of Technical Assistance Resource Center, headed by Frank Farrow
2000	• Use of diarists and local documenters begins • Early activities in sites focus on "seeding ideas," resident engagement, neighborhood small grants projects, and Earned Income Tax Credit campaigns
2001	• Cindy Guy becomes manager of cross-site survey • NORC/Urban Institute team selected for survey • Casey Board pushes for more attention to results
2002	• 10 of original 22 sites funded for full implementation • Core results and common core outcomes identified • Family Economic Success framework developed • Tom Kelly develops "fish" graphic for theory of change • Mary Achatz hired for process evaluation
2003	• Tom Kelly becomes manager of cross-site evaluation • Frank Farrow becomes manager of *Making Connections'* implementation • Work in many sites focuses on initial results strategies, community mobilization, and cultivation of social networks • Cross-site survey Wave I (2002–2004) • Ralph Smith increases attention to school readiness result • Close-the-gap framework adopted in sites

2004	• Casey Board pushes for more data on specific strategies and for fewer target outcomes • Streamlined indicators are developed • Frank Farrow creates site assessment tool for performance and management data • Mary Achatz and implementation staff define community core capacities and develop assessment tools; Achatz and Kelly develop evaluation framework matrix
2005	• Ralph Smith and Frank Farrow abandon aspiration to achieve neighborhood-level change within the decade • Smith pushes for more clarity on result areas and theories of scale and sustainability • Focus schools strategy begins • Participant Family Data Collection begins • Cross-site survey shifts from evaluation to research tool
2006	• Cross-site survey Wave II (2005–2007) • Three sites transition out of full implementation to smaller, targeted investments • Cross-site survey baseline briefs produced and disseminated • Farrow's implementation team creates Web-based "results database" for management purposes; includes local data and data from cross-site survey related to results • Planning begins for local management (in preparation for end of Casey funding)
2007	• Selection and negotiation begins with local management entities (LMEs) to continue the initiative past Casey's decade of investment • Wave II survey reports produced and disseminated
2008	• Cross-site survey Wave III (2008–2010) • Evaluation "data dress rehearsals" begin • National economic crash constrains work in sites
2009	• Eight sites increase attention to grade-level reading work in focus schools • Performance agreements with LMEs concluded • Publication of cross-site survey reports on family mobility and defining neighborhoods • NORC makes cross-site survey data available to other researchers
2010	• Organizational realignment/repositioning within Casey; Doug Nelson retires and Patrick McCarthy becomes new CEO/president of Casey • Final year of Casey funding spread over 2010–2011

A NEW INITIATIVE for NEW TIMES

Making Connections' start date, 1999, was a time of transition on several fronts. During that year, the Annie E. Casey Foundation's endowment doubled. The sudden wealth put new force behind Foundation leaders' vision for what Casey might accomplish, even while a major internal restructuring (planned before the influx of money) reorganized staff, leadership, and priorities within the foundation. The field of comprehensive community initiatives (CCIs) was arguably at its peak of popularity, with several major foundations sponsoring place-based work in multiple sites. "Theories of change" had emerged to add rigor to the movement and support strategic planning, implementation, and evaluation in an integrated manner. New sources and tools for physical and economic development (e.g., tax credits, the HOPE VI public housing revitalization program, community development financial institutions)[3] added new actors and activities to the community-change arsenal. New ideas for evaluating comprehensive initiatives were fueling debate about how to define neighborhood boundaries, how to capture community-level changes, whether it was possible—or even necessary—to establish causality, how evaluation relates to learning, and what role evaluators should play when interacting with the subjects of their study.

A FOUNDATION IN TRANSITION

At 5 a.m. on November 10, 1999, the 51-year-old Annie E. Casey Foundation, created by one of the founders of United Parcel Service (UPS) and named after his mother, was moderately well off. It had an endowment of $1.6 billion, an annual payout of about $92 million,[4] and a staff of about 70. Twelve hours later, Casey was worth $2.8 billion, and in the months ahead its value rose to $3.6 billion. The sudden wealth was generated by UPS's initial public stock offering. Nearly 81 million UPS shares traded that first day, opening at $50—almost twice the value of the 42 million UPS shares Casey already held—and closing near $70.

Some things changed for Casey after November 10, and lots didn't. Foundation leaders vowed to stick with their planned 2000 grant-making budget of about $146 million. They didn't rush to expand the staff or establish new lines of work. They cautioned that they still planned to "lead with ideas," not dollars. But there's no denying that the money opened up opportunities for greater influence, reach, and accomplishment at a very fortuitous time. Just a few months earlier, Casey had quietly begun the local partnership-building and planning phase of a new initiative, *Making Connections,* in 22 sites around the country. *Making Connections'* budget alone was then estimated at $50 million to $60 million annually, for 10 years.

In some ways, however, the money was the least of the changes that Casey faced. In 1994, the Foundation had relocated from Greenwich, CT, to Baltimore, MD, to get closer to the issues it was trying to address, affordable housing for employees, and access to policymakers in Washington, DC. That same year, Ralph Smith, who would become *Making Connections'* chief visionary and architect, joined Casey's staff to lead the organization through a strategic planning process. Smith, a law professor at Harvard and the University of Pennsylvania, had been chief operating officer of Philadelphia's public schools and an advisor on child and family policy to that city's mayor, and he founded the National Center on Fathers and Families at the University of Pennsylvania. He was a good fit for Casey: deeply committed to kids and focused on achieving positive results. But Smith also brought to Casey a more interdisciplinary approach and a sense that the best way to improve outcomes for children was to address the overlapping needs and opportunities of kids, families, and communities.

The organization Smith joined was low-key and informal. With the exception of a five-person evaluation unit and two or three staff dedicated to the KIDS COUNT national data project,[5] Casey's small corps of program officers were organized around professional disciplines, such as child welfare, juvenile justice, health, and education. Most decisions, a long-time staffer notes, were made "by Doug,* Ralph, and Tony** as if they were still sitting around a kitchen table in Greenwich."

A NEW WAY OF WORKING FOR CASEY

As part of the strategic planning process, Smith initiated the practice of holding "Casey consultative sessions" to elicit ideas from an array of key players in the field, test out ideas, and anticipate potential criticisms. He also directed efforts to mine previous Casey initiatives and those of other major foundations for lessons about what to do—and what not to do—in future initiatives. These activities led, in 1997, to development of the Neighborhood Transformation/Family Development initiative and its demonstration project, *Making Connections.* A study of Casey's work offers this description of the initiative and its place in the Foundation's evolution:[6]

> *Casey's initiatives had always been child-focused, but most tried to achieve better child outcomes by reforming powerful service systems. During the early 1990s, awareness was growing within the Foundation that improving child outcomes also required devolving accountability and authority to the local (neighborhood) level. In order to take advantage of this opportunity, community leaders and organizations needed to acquire and develop stronger capacities, which the Foundation sought to facilitate through its Rebuilding Communities Initiative (1994–2001).*

> *Foundation leaders also realized that many of the families whose children Casey hoped to help were so stressed that they couldn't function effectively as parents or caregivers, so 'if we wanted to change*

---

*Doug Nelson, Casey's president and CEO from 1990 to mid-2010.

**Tony Cipollone, then director of research and evaluation.

children's futures, we would first have to change their families' present conditions.'[7] The attention to families soon zeroed in (via the Jobs Initiative, 1996–2004) on families' economic success, because 'whatever else it takes to raise healthy, safe, and successful children, it first of all takes the resources to meet children's basic needs, to purchase a measure of family stability and security, and to build a foundation for future aspirations.'[8]

The importance of place and families came together for Casey with the insight that improving neighborhood conditions could be a strong and necessary step toward strengthening families.[9] Place was especially important when we realized that the kids who experience the worst outcomes in America are concentrated in several hundred extremely tough neighborhoods. Although children's outcomes were improving generally, more than 15 million American children lived in families with incomes below the federal poverty line, and that factor alone put them at much greater risk of poor outcomes. Data showed that those vulnerable children lived in families with multiple risk factors, which in turn lived in neighborhoods of highly concentrated poverty.

The Foundation's previous initiatives, while incrementally helpful, hadn't produced 'compelling evidence of our efforts' efficacy, scale, or sustainability.'[10] In the new generation of Casey work, Foundation leaders wanted to be able to point to children and families who had benefited, rather than taking it on faith that system reforms had, at some time and for some population, done some good.

[As a new launch pad for change], the Neighborhood Transformation/Family Development initiative made place-based community change an explicit focus. NT/FD was based on the premise that children do well when they have healthy, supportive, economically secure families, and families do better when they live in opportunity-rich neighborhoods with connections to services and resources. Unlike other visions for community change that focused either on improving residents' economic prospects or revitalizing the neighborhood, however, NT/FD viewed the community as a place to strengthen and support families—a strategy Casey believed would be less likely to drive residents out of their neighborhoods as conditions improved.

To promote NT/FD's principles and point of view, Casey created a campaign that partnered with national policy and civic organizations to engage their members as champions and co-investors in making the case for "place-based family strengthening." Concurrently, Casey created *Making Connections* as the vehicle to make NT/FD "credible, visible, and actionable" in real communities.

A FIELD IN TRANSITION

Casey's changes came at a critical juncture for community change efforts in general. In 1992, the Casey, Ford, and Rockefeller foundations had sponsored a conference in Cleveland, *Building Strong Communities: Strategies for Urban Change,* at which senior staff and managers candidly discussed the successes and weaknesses of their community change initiatives

The Work of *Making Connections*

....

Making Connections' implementation began in 2000 in 22 cities spread across the United States. There was no single lead organization within sites. Local site teams consisted of Casey staff, national and local consultants, and staff from collaborative partners such as community foundations, local United Way affiliates, city-wide and neighborhood-based nonprofit organizations, city agencies, neighborhood associations, and community leaders. Each site had a site team leader who was a senior member of Casey's staff.

Within each site, the local teams and Foundation-based staff selected specific neighborhoods of 25,000 to 30,000 residents in which to concentrate their work. (The exceptions were San Antonio, which focused on the city's entire West Side—138,000 residents—and Camden, which never selected a target neighborhood). Site team leaders and teams engaged local stakeholders in various ways. Some began with neighborhood small grants programs, some with family or neighborhood study circles, some with community mapping and assessments of community assets and needs, and some by convening neighborhood-wide meetings and summits.

In 2002, Casey selected a subset of 10 sites to continue with full implementation of *Making Connections* while the rest shifted to more targeted investments. At that time, all sites began an Earned Income Tax Credit (EITC) campaign. Over the next few years, sites had the most success with employment: developing a neighborhood pipeline strategy to connect residents to metro-area jobs and, in some cases, creating Centers for Working Families. The employment and EITC strategies augmented asset-building and -protection strategies for families. In later years, some sites also addressed issues of access to financial services, medical debt, predatory lending, employment for ex-offenders returning to the community, and small business loans.

Sites also funded resident and community engagement. Several implemented or adapted leadership development programs or academies. Others involved residents as messengers and connectors in the forms of trusted advocates, *promotoras,* or Walkers-and-Talkers, and as part of the Local Learning Partnership. Denver focused on community organizing institutions and infrastructure, while Louisville developed a resident membership network.

In addition to grant making, Casey supported the work through social investments in all sites, including certificates of deposits in local community financial institutions and program-related investments in the form of debt and equity.

In 2003–2005, Casey began to provide technical assistance on the use of family, friend, and neighbor care to help prepare more young children to enter kindergarten ready to succeed in school. Casey also began to focus on improving reading proficiency by the end of third grade. Eight sites selected specific schools in which to concentrate their efforts around grade-level reading for all students.

All sites pursued opportunities to influence city and state policies affecting neighborhood residents and systems. For example, Denver organized to change the state school funding formula for struggling schools; Des Moines worked with the state attorney general to regulate predatory

mortgages; Oakland worked to change city planning and real estate development. Casey also helped sites access federal funding opportunities, such as Food Stamp Employment and Training, Second Chance Act, Faith & Families funding, Investing in Innovation Fund, and Promise Neighborhoods Initiative.

The average budget per site, per year during *Making Connections'* startup phase was $300,000 to $500,000 for infrastructure and programming (including local team staffing, administrative and operating expenses, travel, and subgrants to consultants and nonprofit organizations that provided programming), $150,000 to $200,000 for data and evaluation activities, and $100,000 for communications and other activities. During the full implementation phase, the annual budget per site was about $1.25 million for programming, $250,000 to $300,000 for data and evaluation (mostly for Local Learning Partnerships), and $150,000 for local technical assistance, communications, and other activities. The Technical Assistance Resource Center, which served all sites, received between $5 million and $7 million per year.

(New Futures, the Neighborhood and Family Initiative, and the Persistent Poverty Project, respectively). A year later the National Community Building Network was created, followed closely by the Aspen Institute's Roundtable on Comprehensive Community Initiatives (now the Aspen Institute Roundtable on Community Change), to help foundations and others exchange information, strategies, and practices. Several research reports and evaluations of that period captured the field's tensions and were broadly disseminated and discussed.[11]

In the midst of this widespread reflection, "many foundation leaders began to reassess their grant-making practices, investments, and strategies. *Foundation Grantmaking for Children and Youth,* a 1998 study conducted by the Harvard Family Research Project for the W.K. Kellogg Foundation, reported that many national

foundations were in a time of major flux. Of the 19 foundations surveyed, 11 had completed a major strategic planning process within the past two years. 'Foundations are rethinking their basic ideas and assumptions about how to create and sustain change in order to improve the status and well-being of youth on a widespread basis,' wrote author Heather Weiss."[12]

By 1999, the Aspen Roundtable was deep into the task of helping to clarify concepts, indicators, and measures for "community building." As Roundtable co-founder Anne Kubisch later wrote:[13]

> *CCIs intended to go beyond the achievements of existing community-based organizations, notably social service agencies and community development corporations, by concentrating resources and combining the best of what had been*

learned from social, economic, physical, and civic development in a particular place. They aimed to implement an intervention where the whole would be greater than the sum of its parts, a vehicle that would catalyze the transformation of distressed neighborhoods….

Although CCIs varied enormously depending on location, sponsor, and community capacity, the 'classic' CCIs were generally designed as distinctive and discrete efforts that analyzed neighborhood problems and assets holistically, created a plan to respond in a comprehensive way, engaged community actors, and developed a structure for implementing the plan. Each sought to achieve multiple results with a combination of inputs centered around some conception of 'community.' Their goals included individual and family change, neighborhood change, and systems change. They operated according to community- and capacity-building principles. A wide variety of programmatic activities was open to them, from human services to economic development to social capital building strategies.

Concurrently, the philanthropic field's use of program evaluation, impact assessment, outcomes evaluation, and performance measurement had grown steadily for many decades. The quest for results, and for increasingly sophisticated ways to measure and analyze them, accelerated in the 1990s with the advent of new approaches, such as venture philanthropy, and innovations in digital and Web technology. Online resources like GuideStar made data-based decision-making easier. New concepts, such as the method for calculating social return on investment created by the Roberts Enterprise

The story of *Making Connections'* evaluation can only be told in the context of Casey's evolution as an institution and the changing field into which *Making Connections* was born.

Development Fund, entered common discourse if not mainstream use.[14] Within a few years, the Center for Effective Philanthropy would be established as a source of comparative data, assessment tools, and research on effective practices for foundation leaders and staff. Over the next decade, while terms like "corporate social responsibility" and "sustainable responsible business" reverberated in the for-profit world, leaders and constituencies within the philanthropic field and nonprofit sector attempted to define their own standards of practice.

Challenges remained when it came to evaluating community change, however. Previous Casey evaluations had grappled with some of the issues. For example, the evaluation of *Plain Talk,* conducted by Public/Private Ventures between 1994 and 2001, incorporated ethnography as a way to capture the influence of local conditions on youth outcomes documented through baseline and follow-up surveys. It focused on community-level change but struggled to pinpoint what happened to individuals who were directly touched by the intervention (finally using a statistical model to compare actual and predicted outcomes for the same youth). And it established correlations but stopped short of claiming casuality.[15]

Casey convened national conferences in 1995, 1996, and 1997 to explore issues and ideas for evaluating community change. The Aspen

Roundtable, meanwhile, pushed the field to think more clearly about theories of change, which had revolutionized how many foundations and their evaluators thought about achieving and measuring results. Aspen's second volume on approaches for evaluating community initiatives, published in 1998, described ways to incorporate theory of change into measurement and analysis strategies and was widely hailed as a step forward both for initiative design and evaluation. Another influential analysis came from *Urban Problems and Community Development,* published in 1999 by the Brookings Institution. Its authors surveyed the history of urban development and concluded that, "because so much progress comes from learning by doing, the goal of evaluation research should not simply be to document success or failure. Instead, evaluators should aim more systematically to distinguish factors that produce failure from those that produce success."[16]

In other quarters, the concept of developmental evaluation was taking root, and it resonated with Casey's desire to let *Making Connections'* evaluation evolve with the initiative and to have the Foundation play a hands-on role in the evaluation. Michael Quinn Patton defines developmental evaluation as:

> *Evaluation processes, including asking evaluative questions and applying evaluation logic, to support program, product, staff and/or organizational development. The evaluator is part of a team whose members collaborate to conceptualize, design, and test new approaches in a long-term, on-going process of continuous improvement, adaptation,*

and intentional change. The evaluator's primary function in the team is to elucidate team discussions with evaluative questions, data and logic, and facilitate data-based decision-making in the developmental process.[17]

All of these factors would shape the design and implementation of *Making Connections* and its evaluation.

UNDERLYING ASSUMPTIONS

Making Connections **was like previous Casey initiatives in its firm commitment to building community capacity** for leadership and action. Several other factors set it apart, however, including:

• A 10-year time frame;

• A commitment to "moving the needle" on family and child outcomes by concentrating investment in a small set of neighborhoods rather than spreading it thinly over a wider area;

• A simultaneous focus on children and adults (which Casey would eventually dub a "two-generation approach to addressing poverty") and communities;

• The inclusion of hard-to-measure goals, such as building public will;

• A decision to redeploy senior Foundation staff, who were largely system- or issue-specific experts rather than community-change specialists, as leaders of the site-based implementation teams. These staff were already managing or leading portions of major initiatives in their chosen fields, and they continued those responsibilities while also assuming responsibility for *Making Connections* sites. Several staff were responsible for multiple

Making Connections sites; one, in fact, oversaw five sites while also leading a separate portfolio of work for the Foundation. Furthermore, the lines of reporting for these staff did not change when they took on *Making Connections* responsibilities. About half reported to Patrick McCarthy, then head of Casey's service and system reform work. Most of the rest reported to a handful of Ralph Smith's staff, and only a few reported directly to Smith. The absence of a change in reporting lines meant that the site team leaders had authority but mixed accountability, with many not directly accountable to the initiative's designer;

• The decision to establish a Technical Assistance Resource Center (TARC) to bring the best available knowledge to sites for implementation. Frank Farrow, director of the Center for the Study of Social Policy (CSSP), was recruited to head TARC. In 2002, Casey established an internal Community Change Initiatives (CCI) unit, with Farrow as the director, and in 2003 Smith shifted the management responsibility to Farrow while retaining authority to make changes in *Making Connections'* design;

• The decision to forego a local management intermediary during the early stages of the initiative, transitioning to one only after completing the fine-tuning and recalibration that seemed inevitable for any major initiative.

Ralph Smith felt that Casey needed more "proximity, access, and the ability to respond with more agility and nimbleness" than working through an intermediary could allow. He also wanted to engage a broader cross-section of actors, including local United Way affiliates, community foundations, and others, than would be possible if *Making Connections* was filtered through a gatekeeper. This left Foundation staff directly involved in sites—"developing and testing ideas in the crucible of practice," as Smith liked to say;

• A commitment to "learning while and by doing," without a prescribed model, which allowed each site team leader and local team to take advantage of local opportunities but also permitted huge variation in priorities and strategies. At the same time, Casey expected the initiative to be results-driven rather than model-driven—in other words, the funder was willing to let sites use locally derived strategies to achieve results; and

• The decision to select 22 sites (cities or metropolitan areas) for a two- to three-year exploratory and relationship-building phase and then determine which places possessed the qualities needed to stay the course for a full decade of investment (10 sites were selected, later culled to 6). The original site selection process involved extensive analysis of demographic and economic data on 96 locations that had participated in earlier Casey interventions and three "civic sites" (Baltimore, Atlanta, and Washington, DC), where the Foundation has ongoing civic commitments. Planners compared profiles of these places, developed by the Center for Assessment and Policy Development, with case studies, evaluations, and other research data on more than

500 high-poverty neighborhoods located in 10 metropolitan areas around the country. They then narrowed the field through consultation with Casey staff, longtime consultants, and national experts.

From the consultative sessions that led to *Making Connections*, Ralph Smith distilled a set of "operating principles"[18] that he expected to guide Casey's interactions with local participants in the initiative:

• **Do no harm.** We will try to understand local politics, cultures, and systems so we don't alter relationships in negative ways.

• **Add value.** We encourage the adoption of proven practices as well as fresh ideas that make families and communities stronger and more effective.

• **Lead with ideas, not money.** Lasting change comes from good ideas. Money is necessary to support ideas, but it shouldn't lead the charge. Funding from a national foundation should never overshadow local investments and resources.

• **Break new ground.** We believe in trying new ideas while building on lessons from past experiences. This applies to our own foundation structure and grant making as well as our practices in the field.

• **Make new mistakes.** We believe in knowing and understanding history. We're willing to take some hits as long as they are *new* mistakes—part of the natural process of learning and innovation.

• **Take time to build local relationships and ownership.** Our agenda resonates with many

people, but not everyone. We will ask local partners to make major, and often difficult, transitions, so we must allow time to introduce our ideas, find common points of interest (and conflict), foster local ownership, and test alliances.

- **Engage multiple sectors.** We do not invest in a sole strategy or lead agency at any site. Our partners include residents, local government, employers, financial institutions, large and small businesses, faith-based groups, community-based organizations, cultural clubs, hospitals, universities, schools, law enforcement officials, and grassroots community groups. This broad base increases ownership, ideas, and energy.

- **Cultivate local leadership.** Local leadership is vital but frequently disconnected. We try to support all community leaders and unite them around a common agenda.

In Smith's vision for *Making Connections*, it was especially important to establish and sustain a widely shared sense that results matter, a practice of shared accountability for results, and an ability to collect, analyze, and use data to measure results. Today, Smith uses a billiards metaphor to describe the importance of using data to set targets and to manage efforts to achieve results. "Do you play pool?" he asks:

> The favorite shot in pool is the break shot, because whatever goes into the pocket is yours. With a baseline of zero, you put three balls in and you've made progress against the baseline. The rest of the game is a bit more complicated. You've got to call the ball, call the shot, and call the pocket in order to get credit. That—calling the ball, the shot, and the pocket—is

what we mean by target setting and accountability for results.

But although Smith and other Foundation leaders were attentive to results from *Making Connections'* inception, not everyone charged with implementing the initiative on the ground shared that early focus. Differences in the messages about results—in the way that people understood what Casey wanted to achieve, how to measure the changes, and how to communicate the changes—would complicate *Making Connections'* implementation and evaluation for years to come.

Several other decisions made at *Making Connections'* start reveal expectations for what the evaluation would and would not accomplish. These decisions were made after extensive discussion of the pros and cons—what Tony Cipollone, Casey's longtime director of evaluation (among other roles), called "the gets and give-ups." In many cases, staff made the decisions knowing full well that there would be difficult consequences.

Decision 1: Casey would not begin the evaluation until after the exploratory phase was completed and Foundation leaders had decided which sites would continue into the full implementation phase. This decision had practical motivations: Evaluating all 22 sites from the very beginning would have dramatically increased the cost of evaluation, with no offsetting benefit in the eyes of *Making Connections'* developers. The Foundation staff serving as **site team leaders** (STLs) also resisted being evaluated during the initial relationship-building phase. After all, most were risking their reputations on a high-stakes but loosely defined undertaking, "neighborhood transformation"

Making Connections' Core Assumptions

....

Families matter. No service, public institution, or system can substitute for a family's crucial role in a child's development.

Place matters. Efforts to strengthen families cannot be separated from the neighborhood contexts in which families live.

Connections matter. Outcomes for children and families can improve if families are connected to supportive and empowering human relationships, services and systems, and economic opportunities.

Resident engagement matters. Residents must be at the center of the work in order for authentic change to occur, remain relevant, and be sustained.

Data matter and results count. In order to make families and communities stronger, we have to know exactly what problems must be solved and what assets are available.

A multifaceted approach is essential. No single policy, entity, sector, or strategy can solve the problems alone.

Broad-based support is essential for stimulating and sustaining social change. Local partners have to mobilize to lead and, ultimately, own the work in their communities.

and "family strengthening," that lay outside their field-specific skill sets. Although *Casey's* leaders might want to learn while doing, and expected evaluation to help them do so, most of the *site team* leaders who reported to them still viewed evaluation as a mechanism for declaring success or failure.

Initiative leaders did commit to collecting baseline data on child, family, and neighborhood conditions as early in the full-implementation phase as possible so that *Making Connections* would get credit for improvements stemming from the initial activities in sites. Meanwhile, to prevent valuable information from being lost during the exploration phase, Smith accelerated

plans to create a Diarist Project. This was an informal documentation effort that aimed to capture the broad storyline and emerging issues within each site. The diarists were modeled after the role freelance writer Joan Walsh played for Angela Glover Blackwell during the formation and early years of the Rockefeller-funded Urban Strategies Council, in which Walsh interviewed Blackwell and others about the change process and what they were learning. *Making Connections'* site team leaders were similarly expected to engage diarists to capture the process in real time. By "real-time learning," *Making Connections'* developers meant learning that occurs while implementation is under way; knowledge

that is both informed by and able to inform ongoing practice.

The Diarist Project struggled to find its feet, going through three different coordinators within the first three years. Each coordinator had a different vision and focus, and none exactly tracked the model Smith had in mind. As the coordinators changed so did the diarists' job requirements, and the "inherited" writers often weren't suited to the changing role. The final coordinator, who managed the project for seven years, was most interested in capturing the perspectives of neighborhood residents involved in *Making Connections* and was reluctant to dissect their statements in any way that might appear evaluative. Thus the Diarist Project became a source of material that was richly descriptive but not analytical or, according to many evaluation and implementation staff, actionable.

Decision 2: Casey would manage the evaluation and play a hands-on role in shaping it, rather than relying on an independent third party. While previous Casey initiatives had all had an in-house evaluation manager, the decision to forego an independent evaluator was a marked departure from previous approaches. (In the evaluations of *Family to Family, Rebuilding Communities, Plain Talk,* and the *Jobs Initiative,* third-party evaluators were routinely identified through an RFP process and competitive bidding.) Foundation leaders felt that measurement was going to be so complex, staff had to be as close to the work as possible. Tony Cipollone, in particular, felt that the analysis of *Making Connections* would have to be closely linked to the implementation side, and that could not be accomplished through an outsider who didn't have access to all of the

The decision to forego an independent evaluator was a marked departure from previous approaches, but Foundation leaders felt that measurement was going to be so complex, staff had to be as close to the work as possible.

internal conversations and decisions. Cipollone also hoped to increase Casey's own capacity for self-evaluation, reflection, and learning. In deciding to control the evaluation from within the Foundation, Cipollone and his staff realized that they were giving up some objectivity and, perhaps, credibility. To ensure careful oversight of their evaluation methods, they drew on the expertise of a group of external research advisors with long experience working with the Foundation, the National Evaluation Advisory Team (NEAT).

The advisory team met monthly during *Making Connections'* early years and continued to meet as *Making Connections'* survey analysis and management team. Its members included: **Lynn Usher,** an expert on child welfare reform and evaluation who was then a faculty member at the University of North Carolina (Usher left the advisory group in 2005 and is now retired); **Claudia Coulton,** co-director of Case Western Reserve University's Center on Urban Poverty and Community Development and an expert in measuring neighborhood-level change; **Tom Kingsley,** senior fellow in the Urban Institute's Metropolitan Housing and Communities Policy Center; and **Robert Goerge,** a research fellow with expertise in administrative records research, human service system reform, and community change at the University of

Chicago's Chapin Hall Center for Children. Casey's evaluation staff also consulted **Prue Brown** of Chapin Hall on many evaluation matters.

Decision 3: The evaluation design would evolve along with the initiative's design. Casey's initiative directors had seen too many evaluations hew to a rigid framework while the initiative grew in a different direction, rendering the findings meaningless. They also wanted to ensure the evaluation could respond to the implementers' need for real-time data, but the specific data needs couldn't be calculated in advance. And they had a fairly high tolerance for programmatic experimentation. "We were willing to fly the evaluation plane without a flight plan," observes Cindy Guy, a longtime member of Casey's evaluation staff. "The sense was, 'measure now and fit it into an evaluation framework later.'"

Thus *Making Connections'* evaluation was set up as a work in progress. It would develop new goals, measures, techniques, and tools as the initiative grew while also staying focused on the overarching evaluation questions. The evaluators were expected to help coach and facilitate learning. They would develop and share many methods and frameworks with the teams imple-menting *Making Connections;* nothing was ever explicitly *off* the table.

Decision 4: The evaluation would be embedded locally, with an emphasis on self-evaluation and assessment, to build local capacity for evaluation, data collection, and strategic data use. Evaluators would also serve as technical assistance providers, and the evaluation would create measurement tools of use to the local implementation teams and to initiative managers. To support the local data collection and evaluation functions, in 1999 Casey began creating a Local Learning Partnership (LLP) in each site. These coalitions of data experts, neighborhood representatives, and other stakeholders, based on a concept developed by Heather Weiss, were intended to serve as: "the lead partner of a multi-layered learning community; the critical partner in…leading local self-evaluation; the 'critical friend' of the site team, providing reflective feedback on progress and goals; the 'guardian of the theory' of the overall site strategy; and a source of continuous, relevant, useable…information about *Making Connections* neighborhoods."[19] Members of Casey's evaluation staff were assigned to specific sites to serve as the LLPs' evaluation liaisons. The LLPs were managed by Casey's evaluation staff until 2007, when management shifted to the implementation side.

Decision 5: The Foundation would try to steer clear of comparing sites directly to each other, although the evaluation would produce some cross-site findings. This decision reflected an awareness that *Making Connections'* implementation would vary dramatically from site to site based on site team leaders' strengths and local constraints and opportunities.

Decision 6: The cross-site survey would be used to measure outcomes for children and families at the neighborhood level, while local evaluations and Local Learning Partnerships would use multiple methods to measure program-level effects. The focus on population-level change reflected current thinking about community change initiatives as well as *Making Connections'* ambitious goals. Several years into the initiative, while the focus on child and family outcomes remained the same, the initiative's theory for attaining population-level change was adapted to what implementers thought was a more realistic approach, given the available resources and the challenges of implementation: to prototype new ways of working on a small scale, demonstrate results, and use those results to influence public and private systems to adopt the approaches on a larger scale. That shift away from achieving population-level change *within the decade* (although it was still expected in the long term) would move the survey from being a source of outcome evaluation data to being a resource primarily for research, planning, and management data.

Two other choices are worth mentioning, although they happened more by default than intentionality. First, the evaluation covered only *Making Connections'* portion of the Neighborhood Transformation/Family Development initiative, not the national campaign to influence and engage national policy and civic organizations, even though the campaign was an important piece of the change strategy.

Second, although Tony Cipollone, as head of Casey's evaluation arm, had overall responsibility for the evaluation, the day-to-day leadership responsibilities fell to a small group of senior evaluation staff. Cindy Guy had responsibility for implementation and analysis of the cross-site survey and other community- and metro-level research based on national datasets; Tom Kelly for articulating a theory of change and managing the overall implementation of evaluation and process evaluation components, as well as embedded program evaluations; and Delia Carmen for helping to develop local data warehouses and, with Audrey Jordan, Tom Kelly, and Cindy Guy, providing local teams with technical assistance on data collection and analysis. None of these individuals seemed to feel that they, or anyone else, were accountable for making sure the pieces of the evaluation all fit together sensibly. And none of them felt they had the authority to intervene when the evaluation fell out of sync with *Making Connections'* implementation.

The trend for community change evaluations, when *Making Connections* began, was to use an initiative's theory of change to determine what to measure through evaluation. *Making Connections* had a hypothesis—children do well when their families do well, and families do well when they live in strong communities, places where they are connected to economic opportunities (jobs, income and savings, assets, affordable housing); high-quality services and supports (responsive agencies, organizations, and institutions that help families reach their goals); and social networks (kin, neighborhood groups, and other supportive informal ties). The initiative did not, however, begin with a formally articulated theory of change. In fact, a theory wasn't explicitly stated until three years into the initiative.

Instead, Casey's first step was to select target outcomes for *Making Connections.* This process began just before the initiative entered the field but continued long after the field work and evaluation were under way. In fact, the process of refining and streamlining the indicators became a emblem of *Making Connections'* leaders' commitment to learning while, and by, doing. Concurrently with the early work on outcomes, Casey developed a cross-site survey of residents in *Making Connections* sites.

ESTABLISHING TARGET OUTCOMES

In 1999–2000, at Ralph Smith's request, Program Associate Janice Nittoli (now at the Rockefeller Foundation) reviewed evaluations of major initiatives and research on every aspect of child, family, and neighborhood change that might be relevant to the new initiative. In typically thorough Casey style—casting a wide net and leaving no stone of knowledge unturned—Nittoli created a 17-page list of nearly 300 target outcomes and indicators. The list included outcomes for children, families, neighborhoods, city government, public service systems, and other civic institutions in the areas of:

- Family-strengthening alliance building, advocacy, and collective action;

- Connections to informal social networks, formal helping systems, and economic opportunity;

- Building neighborhood assets;

- Family functioning; and

- Child and family well-being (child and youth educational achievement, positive social values and roles, health, and economic security).

Smith's intent was to develop a large menu of results, all of which were relevant to the initiative, from which the sites selected for full implementation could choose the outcomes they most wanted to pursue. On the Foundation side, Cindy Guy led the process of honing the list by eliciting feedback from her evaluation colleagues, other Casey staff, and outside experts through consultative sessions—again, a long-established feature of Casey's approach to initiative design and development. Meanwhile, Casey's site team leaders were expected to engage neighborhood residents and other stakeholders in distilling a list of outcomes that made sense locally. In some sites this early conversation about results happened as planned but in other places it did not. In fact, the failure to build in sufficient accountability for site team leaders to involve local leaders in selecting outcomes would leave *Making Connections'* evaluation vulnerable later on to pushback about outcomes from a few sites.

While Guy and her colleagues worked on the menu of outcomes, a team of Casey staff involved in designing, managing, and implementing *Making Connections,* along with the evaluation staff, developed an abbreviated set of **developmental milestones** (see Fig. 1 on page 20) to aid the process of selecting a subset of the initial sites for long-term participation in the initiative.

The developmental milestones were a tool for gauging sites' readiness to continue from the relationship-building stage into full implementation mode. This team also identified three necessary **enabling conditions**—key factors related to the site's overall environment that were not necessarily linked directly to *Making Connections* work. These included "a supportive

policy environment; favorable reinvestment potential; and complementary and compatible traffic."[20] (By "traffic," the team meant activities and initiatives—Casey's own and those of other players—that could either be complementary or distracting. In Philadelphia, for instance, Casey ultimately chose to support Safe and Sound, part of the Robert Wood Johnson Foundation's Urban Health Initiative, rather than continuing to invest in a competing change effort.) The idea was that without these developmental milestones and enabling conditions, sites wouldn't have a good shot at achieving the key objectives of "efficacy, scale, or sustainability."

At the time, the developmental milestones' mention of leaving behind a durable infrastructure for change struck some Casey staff as premature and even superfluous. In later years, however—after *Making Connections'* developers stopped aiming for population-level change within the decade and shifted to building capacity to achieve it in the long term—it seemed prophetic.

Moreover, because the developmental milestones were a tool for sorting the sites—not for shaping the work in the sites that ultimately were selected—the milestones did not include the explicit expectation that local participants would be involved in helping to define local outcomes. That omission contributed to the uneven level of conversation about results within sites—and, consequently, to the feeling in some sites that local partners were not adequately involved in setting *Making Connections'* target outcomes.

Two events shaped the multi-year process of distilling *Making Connections'* target outcomes. The first occurred in spring 2001, when Ralph

Developmental Milestone	Progress Objective for January 1, 2000 through June 30, 2001
I. A range of stakeholders embrace and sign up for the *Making Connections* premises.	By mid-2001, there should be no major sector with whom the team has not found a way to engage in some fashion, ranging from deep engagement with some to less-intense exploration with others. There should be a wide array of groups who have embarked on specific activities, and leadership from a variety of sectors should have discovered common interests and begun to shape alliances within and between sectors around this work.
II. A number of committed, visible champions exist among external stakeholders.	By mid-2001, the site team leader should have in their Rolodexes a "short list" (but diverse) of local champions who are putting their support and resources behind the *Making Connections* premises, having begun to make visible efforts to incorporate these ideas and priorities into their own work and to actively recruit others.
III. Gain support of leaders who live and work in neighborhoods.	By mid-2001, the family strengthening agenda should have become visible to, and visibly inclusive of, parents and residents, with the support of the leaders of various groups, networks, associations, and organizations.
IV. Efforts to help families do their jobs are more visible and better linked.	By mid-2001, there should be several tangible examples that a range of parents, residents, and external stakeholders can point to as meaningful, comprehensive efforts to help families do their job by connecting them to opportunity, supportive networks, and high-quality services.
V. Neighborhoods are better equipped to collect and use data to set priorities, make decisions, advocate change, and measure progress.	Site teams, with significant help from evaluation liaisons, should be focused on developing a complement of activities that include roles for a variety of actors (especially neighborhood residents and leaders) and a variety of kinds of data (quantitative and qualitative; synthesis of existing information as well as newly developed data).

Cross-cutting: Contribute to a durable infrastructure for change.
While not among the five developmental milestones, our work during 2000 should be conscious of the intention to make a value-added and durable contribution to the local landscape, whether ultimately as a "leave-behind" in places where we exit or as foundation for the ongoing work in places where we stay.

Progress Objective: By mid-2001, there should be tangible progress toward the development of a data warehouse and specific activities under way that durably strengthen leadership networks, leadership development efforts, and vehicles for accessing models and examples.

Smith and other Casey leaders went to the Board of Trustees to get approval for some sites to transition from the capacity- and relationship-building phase into full implementation mode. They proposed dividing the 22 sites into three categories. One category was for five sites that clearly would not continue with *Making Connections'* broad agenda but had specific programs or collaborations that warranted ongoing investment. Of the remaining places:[21]

> The [five] sites we identified as being ready to enter Phase 2 could demonstrate that they were acting on the core principles of Making Connections and had the resources, civic support, and capacity to put in place a powerful array of investments, programs, policies, and strategies to strengthen vulnerable families in the target neighborhoods and, eventually, city-wide. A second set of [12] sites, which we called extended Phase 1 or 'presumptive' Phase 2 sites, appeared to have the interest and capacity to move forward but also had unresolved issues or rapidly developing situations that suggested we should wait another year before determining their status. Meanwhile, we wanted them to stay the course in Phase 1.

The Board reacted sharply. How could so many sites still be wavering on the bubble after Casey had invested so much time, effort, and money? How could initiative managers not know enough to just make the decision right away? Smith responded by amping up the emphasis on results, both in his interactions with site team leaders and at a cross-site meeting of local partners in September 2001, a few months before the start of *Making Connections'* second phase.

The next iteration of Casey's target results, which was circulated for discussion in early 2002, featured 36 core outcomes. The list was messy, however, as Tom Kelly acknowledged at a meeting of site team leaders and local coordinators: "Often what counts as an ultimate outcome in one strand of work is an indicator of progress or an intermediate outcome in another strand." Kelly also was frustrated that the selected outcomes were primarily ones that the survey could quantify, which he felt locked the sites into tracking things that could only be measured every three years.

In fall 2002, at another cross-site meeting of local teams, a heated discussion ensued between Casey staff and a handful of local participants whose site team leaders hadn't involved them in selecting from the Foundation's menu of outcomes and, therefore, felt that Casey was usurping their authority to determine their own outcomes. After a day-long debate between this small but vocal group and Casey representatives, a compromise emerged. Smith used a Venn diagram to plot out everyone's target outcomes and both sides agreed to a shared set of outcomes they called the **common ground results.** Both sides would continue to collect data on additional outcomes that mattered to them, but the shared set would serve as the centerpoint. The common ground results soon led to a focus on measuring six core outcomes:

1. Families have increased earnings and income.

2. Families have increased levels of assets.

3. Children are healthy and ready to succeed in school.

4. Families, youth, and neighborhoods increase their civic participation.

5. Families and neighborhoods have strong informal supports and networks.

6. Families have access to quality services and supports that work for them.

The second pivotal event that changed *Making Connections'* outcomes and indicators would not occur until 2004, when the first wave of survey data were reported (see pp. 54–55).

THE CROSS-SITE SURVEY

While *Making Connections'* target outcomes were being established, evaluators turned their attention to how to measure change within and across sites. One of the first measurement processes they developed was a combined neighborhood snapshot and longitudinal survey, to be administered three times over the course of the decade-long initiative. The survey wasn't the only measurement tool in Casey's evaluation plan, which over the years would grow to include many other components. But it was the first and, arguably, the most expensive ($23.3 million between 2001 and 2010), labor-intensive, and ambitious measurement component, and thus was a testing ground for many of Casey's choices and assumptions.

Cindy Guy, then a senior associate in evaluation, now associate director of policy research, directed the survey. Guy had an extensive research background, first as a cultural anthropologist and then a research associate for MDRC in Chicago and New York City. Although MDRC specializes in scientific impact studies and cost-benefit studies, Guy found a niche in the qualitative implementation and process research included in every

comprehensive MDRC study. At MDRC she worked on a range of welfare employment, teen pregnancy prevention, and community-based human-service integration projects.

After joining Casey, Guy directed the evaluation of the Foundation's *Plain Talk* initiative and oversaw the design and launch of the *Rebuilding Communities* and *Jobs Initiative* evaluations. Those experiences reinforced her sense that "evaluation is not just a management or technical assistance tool but has its own legitimacy and purpose: 'to find out what works.'"[22]

Under Guy and Tony Cipollone, Casey had a history of and commitment to measuring conditions before and after its initiatives and tracking progress in between. For most previous initiatives, that had meant mining administrative databases and, occasionally, developing and fielding surveys. *Making Connections* presented several reasons to rely on surveys as a primary evaluation tool: the Foundation hoped to measure changes in civic participation, social ties, and other key outcomes that aren't reflected in administrative records. Moreover, local administrative data weren't defined and collected in the same ways in all *Making Connections* sites, making it hard to pool or compare data.[23] Casey evaluations also made extensive use of national databases that could be analyzed at the neighborhood level.

Casey's reliance on survey research was based on an "almost naïve" expectation that *Making Connections* would produce change on a sufficiently broad scale that a survey of neighborhood residents could provide definitive answers about such a complex initiative's effectiveness, Guy recalls. "We figured we would baseline all this stuff in the neighborhoods and the counties

to which they belong, and then do it again at the end of *Making Connections,* and we'd see a narrowing of the gap [between outcomes]."

Guy issued a Request for Qualifications to nine organizations and a Request for Proposals to four of them. After a formal proposal review process and half-day interviews with two finalist teams, the Casey evaluation team (Guy, Cipollone, Kelly, Delia Carmen, Audrey Jordan, and consultant Tony Hall) and the initiative's National Evaluation Advisory Team chose the National Opinion Research Corporation (NORC) but stipulated that the Urban Institute would also play a significant role. At Casey's request, Margery Turner of Urban's Metropolitan Housing and Communities Policy Center teamed up with NORC Project Manager Catherine Haggerty to co-direct the study. This enabled Casey to build on the relationships with local data intermediaries in several *Making Connections* sites that the Urban Institute had developed through its National Neighborhood Indicators Partnership.

Survey developers faced these design issues:

The breadth of Casey's data interests. Casey wanted to be able to examine the interconnections among many different characteristics, conditions, and outcomes[24] and understand how subpopulations experience change differently, both of which necessitated using a large number of measures. The survey had to simultaneously capture longitudinal data on a sample of families with children (including families that moved out of the neighborhood after the baseline survey) and cross-sectional data "providing a snapshot of neighborhood

The cross-site survey was the first and, arguably, most expensive, labor-intensive, and ambitious measurement component, and thus was a testing ground for many of Casey's choices and assumptions.

conditions at each wave, based on current residents including newcomer families that moved in post-baseline."[25]

Technical issues related to sampling and analysis. The family was a crucial unit of analysis for *Making Connections,* but what definition of "family" would be most relevant, authentic, and consistent across sites? Who would be the most appropriate respondent(s) to report on the family's status? What sample sizes would be needed to assess changes over time for residents in a given neighborhood?[26] Would it be feasible and cost-effective to collect and compare results from face-to-face interviews in neighborhoods with results from a telephone survey for the whole city? Was it reasonable to expect the surveyors could maintain contact with, and collect data from, respondents who might move from the target neighborhoods? Would certain types of information—such as birth and citizenship data—be collected for every member of a sampled household, or just the respondent and his or her randomly selected "focal child"? (For longitudinal tracking purposes, the developers ended up identifying a "focal child" for each surveyed family; tracking was based on the location of the focal child.) How about neighborhood boundaries—should underpopulated areas or higher-income enclaves be excluded from the sample? (NORC argued against exclusion,

· ·

WHEN: The baseline survey (Wave I) was administered between 2002 and 2004, depending on when each site began implementing the initiative. Follow-up surveys occurred in 2005–07 (Wave II) and 2008–10 (Wave III).

WHERE: Baseline surveys were conducted in representative sample of households in the 10 *Making Connections* neighborhoods and counties. Wave II was conducted in baseline households in the 10 neighborhoods plus representative sample of households in new housing units in those neighborhoods, plus families from the baseline sample who moved out of the neighborhood. Wave III conducted in seven sites (three dropped out of full *Making Connections* implementation).

WHO AND HOW: Survey was administered in person to the primary caregiver of a randomly selected "focal child" in each selected household by local interviewers recruited by NORC and *Making Connections* site teams (trained and supervised by NORC). In the baseline survey, county households were surveyed via telephone by Chicago-based NORC employees. Wave I covered 700–800 households in *Making Connections* neighborhoods plus 700–800 households in the county. Waves II and III covered 800–850 households in *Making Connections* neighborhoods and the communities to which baseline families with children moved, by phone with in-person back-up.

WHAT: 45-minute survey covering standard demographic characteristics; household composition and living conditions; child health, education, and well-being; family economic hardship; family employment, income, and assets; informal connections to neighbors, neighborhood, and the larger society; civic responsibility and activism; financial and human services; civic and commercial amenities; and organizational participation, leadership, and volunteerism. In addition, each site inserted up to 15 site-specific questions developed by the local team.

saying that these areas "affect overall neighborhood character and connections." Besides, the survey might be able to monitor "changes in the relationship of these population groups to the rest of the [target] community.")[27]

Size of the instrument. Survey developers drew extensively on previous national surveys, based on a large-scale review of existing instruments conducted by the Chapin Hall Center for Children at the University of Chicago,[28] and the draft survey instrument quickly grew into a massive list of items that would take hours to administer. Representatives of NORC, Urban Institute, the Casey evaluation team, and the National Evaluation Advisory Team met monthly from mid-2000 through early 2001 to refine the initiative's target outcomes and whittle down the measures of change. Given the broad scope of evaluators' data needs, however, there was only so much they could do to reduce the number of questions. As Cindy Guy explains:

If we want to study female-headed, single-parent households we have to create a list of everyone who lives in the house-hold, collect data on the relationship of a selected child to every other member of the household, and collect data on the relationship of the selected child's primary caregiver to every other member of the household. That inquiry alone might require dozens of questions just to get to the point where we can categorize a set of female-headed households, about which we can then ask more questions.[29]

The final version of the survey contained 180 questions and took 45 minutes to complete. Although collectively the questions paint a nuanced picture of life in *Making Connections* neighborhoods, Guy acknowledges that viewed in isolation the multiple items can seem "nit-picky, redundant, or…profoundly tedious… and of no relevance to the serious issues facing residents, community leaders, funders, and policymakers." Nevertheless, she adds, those individual data items are essential ingredients for the kind of analyses needed to address complex issues in the lives of many different program- and policy-relevant subgroups.

Variation across sites and neighborhoods. Survey items had to be specific enough to capture meaningful data but generic enough to accommodate differences between the participating cities and, within cities, between multiple neighborhoods if *Making Connections* targeted more than one. In addition to the usual racial, socioeconomic, and political differences, the size of the target areas ranged from approxi-mately 20,000 to 40,000 residents (with the exception of the neighborhood in San Antonio, TX, which had 138,000 residents).

Language barriers. Because many immigrants live in *Making Connections* neighborhoods, the survey would have to be translated into other languages, primarily Spanish, Vietnamese, Hmong, and Chinese (Cantonese). Translated versions were used in neighborhoods where more than 10 percent of the population spoke the corresponding language rather than English.

The need to collect data on families outside the target neighborhoods as well as within them. The rationale for this, in *Making Con-nections'* case, was not so much for compara-tive purposes as to accommodate an expanded notion of place. *Making Connections* sought to transform neighborhoods into places that sup-port better outcomes for all residents, and the only way to know whether the overall commu-nity is changing is to collect data broadly, Guy wrote in 2009. Keeping the larger community on the data screen would also help *Making Connections'* leaders determine "which strate-gies need to scale up and what policies need to change," she continued:[30]

Having data on the larger community keeps us nimble, able to make changes in response to emerging factors and real-time population shifts. In White Center [Seattle] our local partners developed and honed their strategies by concentrating on Asian immigrants, who compose 12% of the population. But we continued to track other populations, and those data revealed that the Hispanic immigrant population doubled during the early years of Making Connections. Knowing about that popu-lation shift, well before it showed up on the U.S. Census, indicated a logical next step for scaling up strategies in White Center.

While survey developers grappled with these challenges, Guy had her own set of worries. Would local leaders focus on the survey as the whole of the evaluation rather than as one limited tool in a much richer, multi-method enterprise? Would they resist the idea of a cross-site survey altogether? Would preexisting tensions among local stakeholders or organizations adversely affect efforts to involve them in implementing the survey? Would the NORC/ Urban Institute survey team alienate neighborhood residents in the field, derailing the survey process?

In keeping with *Making Connections'* aspirations to engage neighborhood residents in meaningful ways, the NORC/UI survey team and the Foundation's evaluation liaisons talked frequently with Local Learning Partnership coordinators and their point people. To help neighborhood residents develop new skills and capacities, Casey required NORC to make a special effort, in coordination with the sites, to include local neighborhood residents as field interviewers. This generated a host of legal, ethical, and practical issues involving the use of inclusive and culturally competent recruitment and training practices; data quality control; the impact of payments on residents' eligibility for public assistance; the challenge of paying residents who lack bank accounts or social security numbers; the transportation needs of residents who worked as interviewers in communities with poor public transit; the need to provide vehicle insurance for interviewers who could not afford to purchase it on their own; and family issues (from lack of child care to domestic violence) that sometimes interfered with the performance and work output of residents who served as interviewers.

To give sites a sense of control and a stake in the survey, Casey made some decisions that weren't necessarily in the best interest of data collection. For instance, Guy wanted to conduct the baseline (Wave I) survey at roughly the same time for all of the sites that moved into full implementation of *Making Connections* at the same time—that is, in 2002 for the first cohort (Denver, Des Moines, Indianapolis, San Antonio, and Seattle/White Center) and in 2003 for the second cohort (Hartford, Milwaukee, Oakland, Providence, and Louisville). But local partners in White Center protested that it was too dark in winter to go door-to-door safely in the neighborhoods, and they asked to delay their survey until spring of 2003. Similarly, partners in Louisville, the last site to enter full implementation of *Making Connections,* said they were too busy with startup activities to conduct the survey until 2004. Advised by Foundation evaluation staff, Guy agreed to both requests, gambling that the discrepancy in baseline dates would have minimal effects.

In August 2002, the long design meetings, cross-country phone calls, and sleepless nights paid off as NORC took the first survey into four sites. Five more sites were surveyed in 2003 and the last site in early 2004. This first wave of data collection captured information on the composition and demographic characteristics of households in *Making Connections* neighborhoods as well as "a broad range of…result areas, including employment, income, assets, civic participation, social networks, human and financial services, neighborhood characteristics and child well-being."[31] The survey achieved a response rate of about 75 percent, a rate well within the standards of best practice for the field, Guy notes.

The survey data offered insights into key issues that *Making Connections* faced. For example, Casey staff knew going into the initiative that families who live in *Making Connections* neighborhoods have difficulty getting adequate medical care for their children. Yet the survey data showed that most children in *Making Connections* neighborhoods were, in fact, enrolled in public health insurance programs. Guy later wrote:

> *The [survey] data indicated that the root cause of the problem was not solely ineligibility nor the parents' failure to sign up for coverage. We now know that families can be unable to use their health insurance when they need it because of regulatory barriers, such as onerous and inefficient reauthorization practices…. From the family and program perspective, the end result is the same: inadequate care due to insurance problems. But the fuller understanding of root causes has practical implications. While outreach and eligibility reform continue to be important, it is also essential to reform insurance reauthorization practices…[32]*

Similarly, Casey staff knew that residents of low-income neighborhoods use check-cashing services more than residents of affluent areas, and the survey confirmed that assumption. But the data challenged the stereotype that low-income residents are "helpless prey of sharp-dealing financial predators," Guy notes. The data showed that many neighborhood check-cashing customers also have bank accounts and revealed an unmet demand for check-cashing services among residents outside the neighborhoods—evidence that reliance on check cashers "is not based on victim-like characteristics of ignorance or passivity [but] represents a rational choice based on an informed, critical assessment of options."[33]

When it came to learning in real time, however, the cross-site survey wasn't *Making Connections'* best tool. The survey was conducted only every three years, so it was hard to use insights gleaned from the data to correct the initiative's design or implementation. By the time data were available, sites' budgets had been negotiated and subgrants made with local partners to pursue tactics that initiative planners thought were appropriate—commitments that typically don't change until the next grant cycle. A series of investments to strengthen family, friend, and neighbor (FFN) child care providers and child care centers, for example, was well along by the time the survey revealed that the most vulnerable children in *Making Connections* neighborhoods weren't in FFN or center-based care but at home with their mothers.

The survey and the target outcomes laid important groundwork for measuring change in *Making Connections*. But the question of what to measure, and how, couldn't be fully answered without understanding *Making Connections'* theory of change—the topic of chapter IV.

THEORIES of CHANGE and SCALE
(2002-2005)

Making Connections' theory of change was prompted not by designers or implementers of *Making Connections* but by changes on the evaluation side. One day in late 2003, Tony Cipollone, Casey's vice president for civic sites and initiatives, then Casey's vice president for assessment and advocacy, called Tom Kelly (then a senior associate in evaluation) into his office. Cipollone had joined Casey in 1989 to work on the *New Futures* initiative and to build an in-house evaluation unit—one of the first at a major national foundation. He had an extensive background studying, evaluating, and writing about at-risk youth, school improvement, collaborative service delivery systems, and other youth and family issues for Casey and had earlier consulted for Abt Associates, Public/Private Ventures, and Education Matters. Cipollone and his team had developed evaluations for several multi-year, multiple-site Casey initiatives, including *New Futures, Family to Family,* the *Juvenile Detention Alternatives Initiative,* and *Plain Talk.* Cipollone's approach to evaluation encompassed "a drive to understand not just whether something works but *how* it works; an appreciation for the complexity of social issues; high standards for accuracy and clarity; and sensitivity to the importance of collaboration. [He] also was interested in developing new evaluation models and approaches."[34]

Over the previous year, however, Casey's neighborhood redevelopment work in East Baltimore had commanded much of Cipollone's attention. He knew he was stretched too thin to continue overseeing an evaluation as complex and labor-intensive as *Making Connections,'* and he asked Kelly to take over as manager of *Making Connections'* cross-site evaluation. "By the way," Cipollone added as Kelly left his office, "If you take this job, by the time the evaluation is over you might be the most-hated person in this building."

Cipollone's joke found a receptive audience. Kelly had been at the Foundation for three years, and he came not only with a Harvard education but extensive experience in the child welfare, public health, and human services fields. Kelly's pre-Casey evaluation experience mainly involved federally funded demonstration programs—a pretty straightforward, experimental-design task. He had evaluated state welfare reform and Medicaid waiver experiments as well as services for abused children, homeless youth, and hospice patients; developed a series of evaluation training handbooks for the National Center on Child Abuse and Neglect, the Head Start Bureau, the Children's Bureau, and the Family and Youth Services Bureau; and edited the *Program Manager's Guide to Evaluation* for the federal Administration on Children, Youth, and Families. Kelly is a quick study, the kind of person who seeks out challenges. (A framed

photo in Kelly's office captures his decision a few years back to take a trapeze course at circus training camp.) Still, after accepting Cipollone's challenge, Kelly recalls thinking that "this *[Making Connections]* was over and above anything I could have wrapped my head around."

A year earlier, Kelly had grown concerned about whether the evaluation could keep up with *Making Connections'* constantly changing theory of change. No evaluation that he knew of had ever succeeded in adapting to a changing initiative. So he asked site team leaders a simple question: What are you trying to do? If he could name the pieces of work accurately enough, the evaluators could measure them, Kelly reasoned. He boiled down the responses

from site team leaders into a set of common elements of *Making Connections*. When he depicted the relationship between the elements graphically, the shape resembled a fish, and the picture (see Fig. 2 below), presented to a joint meeting of site team leaders and technical assistance providers in 2002, became widely known as the "fish theory."

Technically, the fish was not a complete theory of change, nor was it intended to be anything more than a prompt for local teams trying to articulate their own theories of change. But it had the advantage of making sense on a broad level and giving an air of coherence to all the terms and concepts Casey staff had been using

Fig. 2: The Fish

SHARED VISION

RESIDENT LEADERSHIP

PARTNERSHIP

IMPLEMENT STRATEGIES

CAPACITY TO TRANSFORM SYSTEMS

LEARNING/ DATA

COMMUNI- CATION/ PUBLIC WILL

A COMMUNITY MOBILIZED FOR ACTION AND RESULTS

POWERFUL STRATEGIES OF IMPACT INFLUENCE LEVERAGE

SCALE

ECONOMIC CONNECTIONS

SERVICE CONNECTIONS

SOCIAL CONNECTIONS

RESULTS FOR

CHILDREN FAMILIES NEIGHBORHOODS
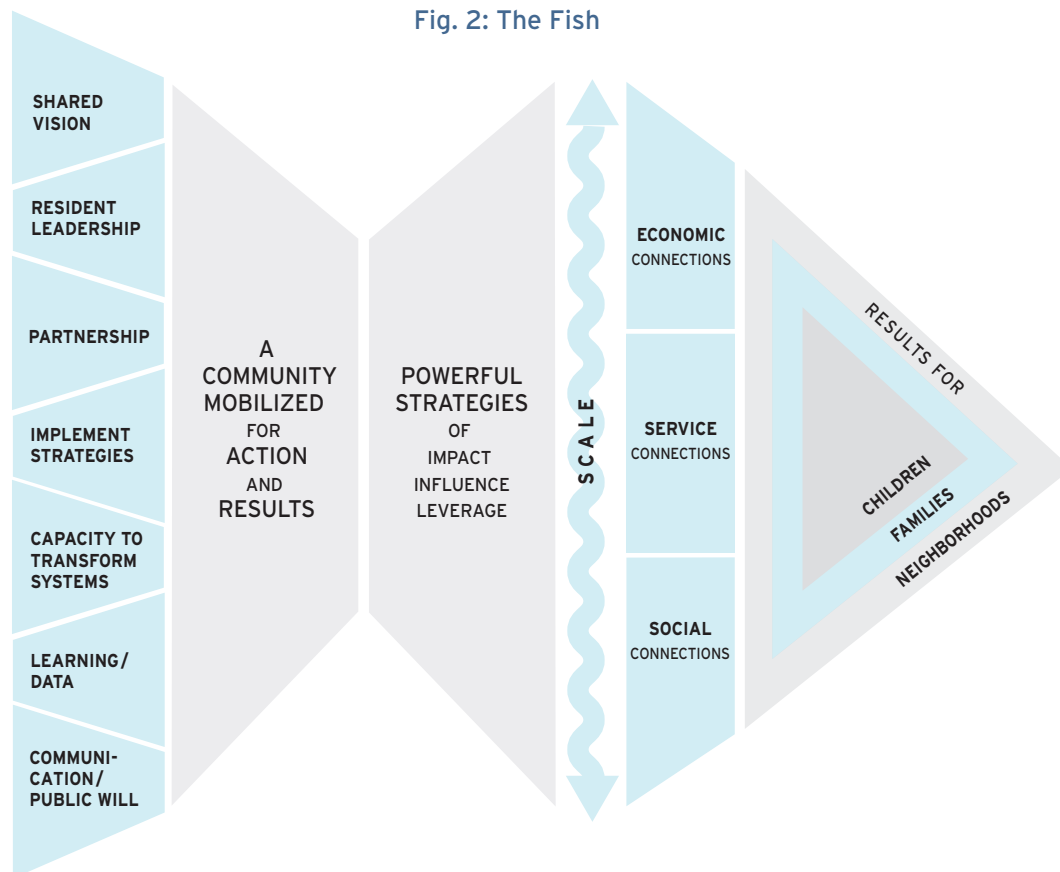
with local partners. *Making Connections'* local coordinators loved the fish.

At the time, Kelly and his colleagues spent a lot of time "debating whether the fish had all of the necessary anatomical parts needed to swim. Did it include all the important aspects of *Making Connections* and what it [was] trying to achieve?"[35] They didn't go much further to define the fish's individual parts, however, until late 2004,* when a workgroup concerned with the initiative's sustainability took up the matter of the theory. "What does an exemplary fin [a core capacity on the graphic] look like in an ideal situation?" Kelly asked in a memo that continued the fish imagery:[36]

> *What kind of fins and gills does the fish need to survive in Denver versus Des Moines? And how do our fish begin to swim on their own? Are core capacities the tailfin that enables the fish to move or the body that keeps everything swimming in synch? Or is 'a community mobilized for action and accountability' the fins propelling the fish forward or the gills and brain enabling it to breathe, think, and react? How will our fish survive, grow, and move forward on their own?...*

> *We need to examine and document the context within each site to help explain the results achieved or not achieved and why….[Moreover], past investment experience tells us that foundation-driven change probably will not last without true resident and customer demand and accountability. How are we ensuring that*

> *our fish swim close to this sunlight to continue to obtain nutrients to grow and succeed?*

And how about the Foundation's role in the theory of change? Kelly went on:

> *At this stage, are we still providing first aid to fish or are they already swimming in the right direction and under their own power? Should we be feeding the fish more or less? Or should we be providing more adult food choices and advanced nutrients now? What role does the Foundation have in helping to diminish the negative forces of riptides and whirlpools stopping or slowing fish from swimming in the right direction? Do we need more fish to add to our school or bigger, stronger fish?*

The looseness of *Making Connections'* theory and the non-standardized nature of site leadership led to great variation in implementation on the ground. Not surprisingly, each site team leader began with the types of work, partners, relationships, and theories of change that he or she knew best: community organizing in Denver, service systems change in Oakland and Des Moines, workforce development in Milwaukee, and so on.[37]

Making Connections' leaders hoped that the site team leaders would learn from each other, cross-pollinating each others' sites until knowledge, ideas, and strategies had spread across the initiative. Some of that did happen. But an additional level of complexity was introduced as site

---

*In 2003, Kelly reissued the graphic with the addition of roles played by the site team in assisting the "fish": Foundation staff were to enhance capacity, create favorable conditions, identify opportunities, and bring knowledge needed for change. Local partners were to foster learning, build relationships, leverage resources, and reduce barriers to change.

team leaders hired local coordinators to serve as their constant "face" on-site, and these consultants' skills were similarly diverse. The local coordinators' role also changed over time, from serving as largely administrative representatives of Casey to playing the primary leadership role. To a great extent, during this time the Technical Assistance Resource Center became the main vehicle for making sure that messages reached site team leaders, local coordinators, and community partners—but this expectation wasn't always clear to TARC, whose staff concentrated on providing technical assistance on implementation strategies. Given these circumstances, it isn't hard to see the potential for wide variation within and across sites.

Across Casey departments, Foundation staff who had expertise in *Making Connections'* core result areas developed frameworks and training to help their colleagues. For example, Bob Giloth, who now heads Casey's Center for Family Economic Success, helped TARC develop cross-site meetings on affordable housing and community development strategies along with workforce, income, and asset building. But the community development piece didn't take root in most sites (with two or three exceptions). Recalls Frank Farrow:

> *It wasn't a failure to see the connection between people and place, it was a matter of how much the sites could handle at this stage and who was being engaged. We found that the development of new strategies around employment and asset development, both of which required major changes in the perspective of local leaders, left very little time for also doing broader community economic development. In some sites this work did go forward, even*

though it was not a focus of the initiative as a whole. Later in the initiative, with employment and asset development strategies better established, it was the appropriate time for sites to link physical redevelopment with job creation.

In this highly varied environment, initiative managers tried to infuse some consistency in theory by building specific expectations into the written reports periodically required of all sites. At first, they asked local partners to describe (a) how they planned to build the capacities deemed important by *Making Connections* and (b) what results they planned to achieve. Then they realized that asking the question that way encouraged respondents to think of results and capacities as two separate things, when really they were (or should be) deeply entwined. So they began asking instead how sites would build the capacities needed to achieve the desired results, and they developed the image of a double helix, with capacities and results intertwined, to represent the concept.

Making Connections' **implementation approach also varied according to the developmental status of each area of work in which Casey expected to produce results** (i.e., workforce development, income and asset building, children healthy and prepared to succeed in school, resident engagement/authentic demand). Casey's prior *Jobs Initiative* had already established a starting point for workforce development in most sites, and it was relatively easy to turn that into a *Making Connections* "neighborhood workforce pipeline" strategy. Farrow's team had to spend more time developing a pathway to family asset development and a framework for preparing children to succeed in

school, however, because this type of work was newer to the Foundation and to the sites.

In mid-2005, Ralph Smith pushed Frank Farrow, the site team leaders, and technical assistance consultants to clarify the result areas. "If we're serious about scaling and sustaining the work in all of these result areas, we ought to have more clearly articulated our idea of what each of these components would look like and what capacities are needed. It feels to me that this work, which is so important, is built on too shallow a foundation. And I think that contributes to the fact that we're saying different things in different places," he said.[38]

Farrow and his team responded in several ways. They held cross-site meetings in each of the result areas to train local participants on the relevant theory. They developed guidebooks on establishing neighborhood workforce pipelines, cultivating resident engagement, and preparing children to succeed in school. And they proposed a theory of scale (see Fig. 3 below), which Tom Kelly helped to develop, that took up where the "fish" graphic left off. This theory of scale was not a competing framework but, rather, an attempt to explain ideas that weren't captured by the fish: how *Making Connections* intended to achieve greater scale over an extended period of time.

Fig. 3: Theory for Getting to Scale in *Making Connections* (2005)

THEORY OF **IMPACT**

Performance Measurement, Evaluation, Documentation

| Build Capacity and Alliances | > | Develop Strategy | > | Implement and Demonstrate | ≥ | Evaluate Analyze | ≥ | Document Learning and Best Practices |

Impact on a Small, but Growing Number of Families

Influence Service Providers to Adopt Approach

Influence Systems to Adopt Best Practices

Leverage Funders to Reallocate More Effectively

Build Public and Political Will

Advocate for Better Public Policies

IMPACT

CLOSE THE GAP

The theory of scale was based on the idea that *Making Connections'* small-scale, prototypic interventions, along with simultaneous cultivation of community capacities, could "change business as usual," generate evidence that the small-scale interventions achieved positive results, and then influence policies and systems (e.g., city government, state agencies, local United Way affiliates, local funders, and advocates) to embed the new way of working in the way they used their own influence and dollars to reach an ever-widening population. Farrow began to call this theory "scale through influence."

But while Fig. 3 expressed this theory for getting to scale, from the evaluators' perspective it didn't give a clear definition of what scale was. Was it about getting bigger numbers—moving the needle farther—across *Making Connections* neighborhoods? About going deeper with certain populations? If the initiative reached every member of a limited population, would that be scale? Another concern, which got to the heart of the population-measurement challenge, was whether place-based community change and working at scale were compatible activities. Since scale happens at a jurisdictional (not neighborhood) level, would an increasing emphasis on scale mean an inherent *de-emphasis* on place, if place was a neighborhood? Come to think of it, is a single neighborhood big enough to qualify as "place"? Is one neighborhood consequential enough to leverage scale at the jurisdictional level? What economic contexts—for families, communities, and metropolitan areas—might affect the pathway to scale? These questions bounced around at evaluation team meetings over the years, but most were never definitively answered.

"A disproportionate amount of my time was spent trying to find out what sites were being told behind closed doors and what site team leaders, local coordinators, and [other point people onsite] were sharing with others, and then backing evaluation and data into that."

– Mary Achatz, evaluator

The work to develop a theory of scale is one of several instances in which Kelly and his evaluation consultants tried to help *Making Connections'* management and implementation staff develop their own capacity for evaluation, a role Kelly views as crucial. "With Frank [Farrow], Ralph [Smith], the site teams, site team leaders, local coordinators, [technical assistance] consultants, and Local Learning Partnerships, every meeting and encounter was a learning opportunity," he recalls. "It slowed our work down—often when we were simultaneously trying to implement a new evaluation tool—but it was where all our relationship-building effort and attention had to go." Kelly suggested questions to include in the periodic reports that Farrow's team required from sites. He developed PowerPoint tutorials on the theory-of-change approach and on defining targets, for site teams to use. He spent time one on one, trying to overcome the misunderstandings and negative perceptions about evaluation that site team leaders held from past experiences. He sent a steady stream of emails and memos describing what he and his evaluation contractors thought the implementers should be doing and thinking

about, based on what the evaluators observed in sites. He tried to be a "critical friend."

While Kelly saw this capacity-building role as important, he was frustrated by how often it seemed needed—and yet how rarely it was initiated by someone other than an evaluator. It irked him that Farrow didn't invite evaluators to all of the meetings involving implementation staff, technical assistance providers, and local representatives, when previous initiatives seemed to have established that practice as the Casey way of operating. Farrow also seemed reluctant to share with his staff or sites the memos that evaluation contractors wrote after their site visits.

"Being the internal staff person [for evaluation] meant constantly being the person to bring these issues up, to keep making the point that [evaluator] Mary [Achatz] is not just being a pain in the ass but that her role as an evaluator is to give feedback about what she observes," Kelly says. "I think Frank eventually got it, but not the rest of [the Community Change Initiatives unit] or the [technical assistance] team. They were consultants who sat elsewhere, and so my access to them and my influence was very minimal."

Farrow's perspective is somewhat different. He says that Kelly played "a tremendously important and helpful role…[as] a helpful colleague [who] was always ready to bring to bear the power of capturing theory and condensing and expressing it in useful ways—many times at just the right moment." The challenge, Farrow continues, was that the job of supporting and guiding sites' evaluation and data activities, including the Local Learning Partnership,

was divided between the evaluators and the implementation side. Consequently:

> *We heard chronic complaints from sites about their being caught unawares by something they were hearing from the LLP and something I hadn't heard about from the evaluation team. In the middle years of the initiative, in some sites we achieved better coordination, mutual understanding, and integration, but in others it remained an issue. That was one reason we tried to figure out how to get better, more frequent, more predictable communication about implementation and evaluation happening at the Foundation level.*

Farrow recalls that Kelly was present at or invited to almost all of the implementation meetings, beginning with cross-site operations and technical assistance meetings and later expanding to the the implementation team's monthly staff meetings (eventually accompanied by evaluators Mary Achatz and Scott Hebert). Farrow also recalls attending the evaluation team's monthly meetings. In retrospect, however, he acknowledges that "we all should have been even more direct and explicit about where evaluation staff felt they weren't being included. These problems weren't raised at the time as issues that needed to be addressed."

PROCESS EVALUATION, CORE CAPACITIES, and CROSS-SITE FRAMEWORK
(2002–Ongoing)

By mid-2002, *Making Connections'* evaluators had grown concerned that valuable information about *Making Connections'* path to results wasn't being captured either by the Local Learning Partnerships, the diarists, or the survey. The loss of that information would make it more difficult to learn from the initiative, both in real time and in retrospect. So, with leadership from contractor Mary Achatz of Westat, Casey began to develop a process evaluation for *Making Connections.*

In the course of figuring out what to document, Kelly and Achatz realized that much of *Making Connections* implementers' effort and resources had gone into developing communities' capacities to achieve results. The implementation team also was working with sites to define what the most essential capacities were and how they would develop over time. So Achatz also began to develop an evaluation process and tools for assessing core capacities and outcomes. And, in addition to the process evaluation and core capacities assessments, Kelly and Achatz created an overarching framework for the cross-site evaluation.

PROCESS EVALUATION

Although evaluation liaisons had taken a stab earlier at having Local Learning Partnerships *document what they were doing* to implement the initiative, the effort was diluted by the LLPs' more immediate task of getting established, the LLP staffs' lack of training in process documentation, and the dynamics of taking on a function that, to local partners, seemed threateningly evaluative. Hiring a third party to manage the process evaluation would take LLPs out of the hot seat and make the activity more of a partnership, Tom Kelly reasoned. He also realized that documenting what was being done would help sites that were struggling to identify useful tactics and, hopefully, illustrate the link between specific strategies and specific outcomes.

In August 2002, Casey contracted with Mary Achatz to provide technical assistance to local process documenters, a role that would grow to include developing and leading the cross-site process evaluation. In 2006, Kelly and Achatz brought Abt Associates' Scott Hebert on board to help with the process evaluation, core capacities assessment, outcomes measurement, and local capacity building around evaluation. Hebert continued in those roles after becoming an independent consultant in 2007.

Achatz, a senior study director at Westat, brought a useful set of skills to the *Making Connections* evaluation. She had experience designing large-scale evaluations using random assignment as well as interviewing and

community action research. She had trained survey researchers and conducted ethnographic research. And she had specific expertise in evaluating community change initiatives, having played a key role in designing and implementing the evaluation of Casey's *Plain Talk* initiative while employed by Public/Private Ventures.

Although Achatz began her research career from an academic perspective, through a project for the World Bank in Africa and other experiences, she had come to appreciate the value of helping non-evaluators ask useful questions and collect data to understand and resolve those questions better. "My whole orientation is toward building capacity and sharing skills and ways of thinking," Achatz explains.

Those qualities, accompanied by Achatz's no-nonsense manner, were like a breath of fresh air to Kelly, who felt as if he was on the cutting edge of tool development with little support to back him up. Although the members of the National Evaluation Advisory Team played a critical role in survey design and analysis and had extensive experience designing and implementing independent, third-party evaluations, they were not able to give Kelly the advice he was looking for on designing and implementing an incrementally developed, internally focused process evaluation. But Achatz just rolled up her sleeves and set about figuring it out.

Achatz's concept for the process evaluation was that each of the 10 sites would have a designated process documenter who would be part of the Local Learning Partnership (which, in turn, would be well-integrated with the local site implementation team so there was a structure for collaborative learning). The process documenter would keep a running record of what people were doing by collecting, collating, and analyzing material from documents, interviews, and observations.

The process evaluation "was never meant to be just a research project," Achatz says. "It was to establish some capacity for inquiry and reflective practice—to help people ask where they were, where they wanted to be, and how they were going to get there." She viewed her job as supporting the local process documenters by helping them identify appropriate questions to ask and answer and by juxtaposing local learning against a broader, cross-site vision. Accordingly, she began by drafting examples of the kinds of questions that local process documenters might seek to answer. She worked with the documenters, one-on-one, to try to create process documentation plans that reflected solid evaluation principles and incorporated both qualitative and quantitative data.

Almost immediately, Achatz ran into trouble. Casey's evaluation unit controlled the budget for Local Learning Partnerships through 2006 but allowed the site teams, local coordinators, and LLP directors to decide whom to hire locally. Achatz therefore had to depend on Local Learning Partnerships or sites teams to designate a process evaluator. Some sites didn't hire anyone; those that did tended to choose someone who had little sense of research design and no experience with evaluation or reflective

practice. Achatz recalls one hire, a secretary who was selected "because she could take notes at meetings," as typical:

> *She would write down the most arcane things and miss the important points. So I spent a lot of time with her on the structure of evaluation and the big research questions and how the events she was documenting fit into that. We started by recording the meetings and transcribing them so I could help her learn how to pick through the information….She did almost no process documentation for several months while she was learning, and after that she could do observations and meeting notes but nothing else. So then I said, 'Let's baseline the context of the community. For example, what supports are available to families in Making Connections neighborhoods?' Her response was to [list entries from] the Yellow Pages. I had to teach her to go out and identify what types of services and supports families needed and where they were available, maybe put them on a map, and then find out if the programs actually served families. Those were eye-opening experiences for her but a lot of work for me. It would have been a lot easier to just do it myself.*

There were other obstacles, too. Achatz felt that members of several local *Making Connections* site teams resisted giving her access to their meetings and documents. She speculates that people were nervous because they were still doing the invisible work that would, presumably, lead to coherent strategies and results but there wasn't yet much to show for their efforts. Achatz didn't have a role that was recognizable to local partners, like the liaison role played by Casey's evaluation staff, and she didn't have the

inherent power of a Foundation staff member. She didn't feel supported by the site team leaders or by Farrow and the consultants who provided implementation and technical assistance. And, until Scott Hebert came on board in 2006, she was the only person available to help 10 sites needing intensive, highly individualized supervision.

Farrow, meanwhile was hearing from local participants that this new piece of the evaluation was burdensome and, while local site coordinators understood the long-term value of process documentation, they initially didn't feel that the cross-site evaluators understood the realities of implementation or appreciated the many competing demands on their time. The situation gradually improved. By 2010, the process evaluation had generated three or four products per site that Achatz deemed useful, especially for local knowledge-building purposes.

CORE CAPACITIES ASSESSMENT

The emphasis that *Making Connections* placed on building local capacity to achieve results, which accelerated as the sites entered full implementation mode, raised two issues for the evaluators. One was the challenge of measuring the degree and extent to which sites developed the appropriate capacities. The second was the need for evaluators to help sites develop the capacity to collect and use data to reflect on the adequacy of their interventions and identify further improvements. Both issues played out through the evaluation's strategy for assessing core capacities.[39]

At the end of 2004, Kelly asked Achatz to assess the capacities that sites had built to date and

judge how that capacity might be linked—in theory or in practice—to improved performance and ultimately results. Simultaneously, Farrow's implementation team was working with local site teams to ascertain which capacities were most needed to achieve *Making Connections'* results. A committee of staff from the Community Change Initiatives unit and site representatives drafted a paper defining these core capacities.

As Achatz approached her measurement task, she knew she needed a framework and standards to make meaningful and fair assessments across sites and within the same site over time. The framework would have to be flexible enough to accommodate the diversity in approaches and contexts across 10 sites. And it would have to be developmental—beginning with where sites began and, over time, showing how and where they ended up at different points. After considering several options, she adapted a tool called In-Sites, which defined capacities in the context of community systems and addressed six developmental stages, beginning with "maintenance of institution-oriented systems" and ending with a "sustainable community-based system."

Kelly had Achatz use the draft framework to assess capacities in the first five sites to enter full implementation mode. Over the 2004 Christmas holiday, Achatz recorded her impressions in a 25-page memo to initiative manager Frank Farrow. While acknowledging that the content of assessments was limited to observations made during site visits, conversations, and review of periodic site reports that were written without the capacities assessment in mind, Achatz reported that the development of capacities

was uneven across and within sites. Even more troubling were these two findings:

• The activities, programs, or projects under way or anticipated are not sufficient to achieve most of the short-term targets they set for themselves; and

• "Closing the gap" is [an] appropriate [technique] for setting long-term targets…As [the sites'] figures show, though, it's entirely possible to develop strategies that will close the gaps but have little or no impact on significant numbers of families and children, including those who are most vulnerable or most isolated.[40]

Although the assessment report was not widely circulated, the concept of having a developmental assessment tool that could be used across sites appealed to *Making Connections'* evaluators and implementers alike. They agreed, however, that the tool should be overhauled to focus on the communities' capacity to develop and implement "powerful strategies."

Staff and consultants from the sites and the Foundation's evaluation and community change units were convened to develop a matrix that defined how the community capacities would evolve over time, based on the earlier paper by the implementation team. A draft version of this matrix, which Farrow's team shared at a cross-site meeting in 2005, generated discussion about the definition of key concepts and raised some concerns about language and complexity, how the tool would be used, and whether it might scare some partners away from the table. To address these concerns, evaluation and implementation staff agreed on a process for using the matrix to assess capacity locally,

and the new version was piloted successfully in Seattle/White Center in 2006.

The revised core capacities assessment tool:

• Presents indicators to guide reflective discussions within sites on the status of local capacity to develop and implement strategies;

• Identifies and defines in broad terms five dimensions of effective and powerful strategies—scope, scale, resident engagement (known in *Making Connections* as "authentic demand for results"), use of data for ongoing improvement and accountability, and sustainability;

• Identifies five developmental stages that communities pass through as they strengthen their collective capacity to develop and implement powerful strategies. The developmental stages are anchored by "maintaining business as usual" on the lower end of the continuum and "effective practices taking hold and transformation of business as usual" on the upper end; and

• Uses a 15-point rating system to assign numerical values to the sites' capacity assessment. Each developmental stage has indicators defining success and is assigned three values. A lower value indicates that the capacity-building work is in an early phase of planning/implementation, a medium value indicates that implementation is under way and the strategy is beginning to show signs of progress, and a higher value indicates that the community is positioned to move to the next level or stage.

Achatz and Hebert visited each site, where they used the tool to help local participants reflect on their strategies and flesh out examples and evidence of built capacities. The assessments were performed separately on different areas of work, beginning with the strongest work in each site (usually around workforce and income development) and including additional assessments if the site had made progress in other core result areas.

Local *Making Connections* leaders determined who participated in the assessments; the groups ranged from a handful of *Making Connections* staff and consultants to 20 or more stakeholders including neighborhood residents, representatives from community-based organizations and public-sector agencies, policymakers, and funders. Most groups included at least one resident of the target neighborhood(s). Two groups only involved residents.

After each session, Achatz or Hebert summarized the findings and assigned scores to indicate the site's status on the developmental continuum. The summary included numerical assessments, evidence used to substantiate the score, lessons learned, and "things to think about" or suggestions for next steps. Local coordinators, Local Learning Partnerships, and site team leaders were asked to review the summaries and provide feedback.

In a late-2008 memo, Achatz described the value of the core capacities assessment as follows:

> *The tool and exercise communicates the overall Making Connections theory of change—the big picture—to stakeholders. It articulates key components of effective*

or powerful community change strategies and broadly defines the capacities needed to strengthen and sustain effective results-oriented approaches to broad-based community change. The tool enables communities to assess the actual and projected potential of their strategies and collaborative capacity to achieve and sustain results for significant numbers of families and children in the Making Connections neighborhoods. [It] provides a time-out for partners to step back and reflect on where they're at, how far they've come, and what they might need to begin thinking about or doing to expand or channel the community's capacities to develop and implement successful strategies and community change campaigns. Once routinely embedded in practice, the assessment process can be used over time to celebrate progress, further strengthen community capacities, broaden the base of committed partners at the state, local and community levels, and diversify funding sources.

For program staff, the core capacities assessment offered a way to clarify concepts like "resident engagement" and address head-on the challenge of linking strategies to results. "The core capacities, although part of *Making Connections'* theory, hadn't been well-explicated by us initially," observed Bill Shepardson, who directed the initiative's Technical Assistance Resource Center. "In 2004 we had a revolt from local teams, saying they were still having trouble getting powerful strategies in place. So we had all these conversations [about core capacities]… and then people understood better that [capacities] were being built in the service of results."

The capacities assessment also forced evaluators to be more responsive to the sites' need for technical assistance on measuring change. Until developing the Core Capacity Assessment Tool, Tom Kelly explains, "We knew what had to happen, but when we turned to our tool box all we had was duct tape."

THE EVALUATION FRAMEWORK

In 2004, Mary Achatz created an 18-page, single-spaced list of potential evaluation questions clustered around the following categories: results and what it takes to achieve them; community capacities; the structure and governance of local work; and community change and the role of philanthropy. Achatz's questions asked about the scope, scale, and reach of results. They also attempted to link *Making Connections'* theory and strategies with outcomes and to reveal the interrelated factors that shaped results.

In 2004, after sharing the full list with other evaluation staff, Frank Farrow, and Ralph Smith, Achatz extracted a shorter list of summary questions that were easier to communicate (see Fig. 4 on page 41). Kelly began to use this set of "high-level" research questions to describe the evaluation's focus at cross-site meetings.

Kelly and Achatz then created a matrix pairing the high-level questions with data sources, including the process evaluation; cross-site survey; sites' annual strategic plans, budgets, and quarterly progress reports; Participant Family Data Collection (see pp. 57–59); national and local public administrative databases; *Making Connections'* annual multi-site

Fig. 4: Five High-Level Research Questions

I. What changes have occurred in the *Making Connections* neighborhoods	relative to overall community conditions/key indicators such as: • Employment rates/labor force attachment? • Household income? • Access to affordable housing and/or homeownership rates? • Access to mainstream financial services? • Crime rates/incarceration rates? • Others?
II. What changes have occurred for (subgroups of) families	in these *Making Connections* neighborhoods relative to: • Earnings, income, and work support and/or employment benefits? • (Bad) debt reduction? • Family assets (particularly homeownership)? • School readiness? • Children's health? • Others?
III. What changes have occurred in community capacity, and the systems of support and opportunity	for families in the *Making Connections* neighborhoods, relative to: • Resident participation, leadership, and networks (including mutual assistance activities)? • Improved connections of *Making Connections* neighborhoods and families to support services? • Improved connections to networks for securing employment and advancement opportunities? • Use of data to promote accountability?
IV. What strategies did the *Making Connections* sites pursue	to try to achieve the above results? And what were the contributions of these interventions toward the achievement of the changes that have been observed?
V. To what extent are the capacities, improved trends, and/or positive outcomes observed in the *Making Connections* neighborhoods, families, and children sustainable?	To what extent in each of the *Making Connections* sites is there clear evidence of: • Public will to continue to advocate for a focus on placed-based family strengthening efforts? • Strong, dynamic alliances established by *Making Connections* neighborhoods? • Sufficient on-going funding to support ongoing activities? • Effective management capacity to direct ongoing efforts?

budget; and local data warehouses. The matrix also contained notes on methods and issues involving data collection and analysis. Together with Achatz's data collection workplan, the matrix constituted *Making Connections'* evaluation framework. But although Kelly circulated the matrix to his colleagues on the evaluation staff and the implementation and management side, he received very little feedback or takeup.

After a few years, everyone but Kelly, Achatz, and Hebert seemed to forget that the framework existed. "The matrix wasn't made very visible to sites," Farrow recalls. "It was doubtful that many of the sites could have put the pieces together in this form. Perhaps a stronger use of this synthesized version would have been to structure the work of Local Learning Partnerships and negotiate their budgets more clearly around these five questions."

MEASUREMENT for MANAGEMENT
(2004-Ongoing)

As soon as the number of sites dropped from 22 (exploratory phase) to 10 (full implementation), in 2002–03, Casey's trustees had begun to ask for more data on the sites. Board members were concerned about the evaluation's cost and wanted reassurance that evaluators were capturing data they could use to improve the initiative's management—not just for some far-off determination of success or failure.

In 2004, almost halfway through *Making Connections'* planned 10-year cycle, Casey's trustees again requested more data. This time they wanted to know about the progress being made and, when progress wasn't evident, what managers were doing about it. The information they found most understandable was progress on the specific strategies taking root in sites, such as workforce development, tax credit refund and asset-building campaigns, and efforts to improve family, friend, and neighbor child care. It's easy to see why: Sites had seen impressive early results from Earned Income Tax Credit campaigns, which had brought real money into *Making Connections* neighborhoods. Residents of these areas were so poor that the baseline for success was low; any increase in the rate of returns constituted a major win. Such measures were immensely satisfying and easy to obtain from IRS administrative data. And the trustees

had just formed a new Performance Measures Committee, which ratcheted up the Board's focus on results.

So while Kelly and Achatz focused primarily on how the evaluation could measure capacities, *Making Connections'* **managers faced the challenge of reporting** on progress toward results and on capacities in a way that made sense to the Board. Over the next several years, Frank Farrow and his team developed several data collection tools and measurement techniques to provide an accountability framework within the initiative, meet internal Foundation reporting needs, and satisfy the Board's appetite for data. These tools were developed sequentially and designed to capture what was being invested in and measured at the time. To avoid a sense of "top-down" imposition, Farrow and his team involved local collaborators in co-designing or reviewing the tools and sought input and leadership from community residents, although the extent of involvement varied across sites. In particular:

- They created a **site assessment instrument** (see Fig. 5 on page 45) that implementation staff used to measure sites' progress in developing strategies in the core result areas and to assess progress in developing community capacities associated with movement toward results. The benchmark for this tool was a

10-page continuum chart of the progress sites were expected to make between their starting point and the goal of "neighborhood transformation" and systems and policy change, which was developed in collaboration with local site coordinators. For several years, Farrow submitted these reports to the Performance Measures Committee five times annually, and he and Ralph Smith discussed them with the committee.

- They trained dozens of local partners in an adapted version of Mark Friedman's **results-based accountability (RBA) model,** which combined RBA with the Foundation's concept of "impact, influence, and leverage." The local partners then trained hundreds of community participants.

- In 2006 they created a Web-based **results database,** with data from each site on the core indicators as well as influence and leverage data, and required site teams to update their data quarterly. Sites could generate their own "impact reports" (targeting a specific result area, indicator, strategy, performance measure, or close-the-gap comparison) and "leverage reports" (showing co-investment levels by site, year, source, and result area).

- They helped sites use **close-the-gap analyses** to set targets and assess progress. "Closing the gap" provided a way to engage local partners in setting targets—determining where their neighborhood stood on a given concern (e.g., number of adults who are employed) relative to the larger city or region and what it would take to close the gap in outcomes. "Closing the gap" became a widely used buzz phrase and management tool. It had the advantage

of being equally comprehensible to non-data-savvy participants and data experts; it resonated equally well with neighborhood residents and Casey trustees; it put issues of equity on the table; and it made the murky work of community change seem more concrete and the targets more achievable.

- They helped several sites establish monthly "**results tables**"—forums where public-agency partners, nonprofit representatives, and neighborhood residents could meet to review data about their strategies' successes and failures and adjust tactics accordingly. Some sites also held annual community forums to critically review progress against their performance targets.

About two years after the site assessment tool was created, the Board requested more site-specific data on *Making Connections'* accomplishments. Sites' strategies had matured by then and were generating more (and more trackable) data. So, building on a prototype developed by participants in White Center, and incorporating feedback from other local partners, Farrow designed **data dashboards** (see Fig. 6 on page 47) for sites to report data on impacts in key result areas, the ways in which they were using evidence of their results to influence public systems and other institutions, and their leverage of public and private dollars. Farrow and his team developed a similar dashboard to report cross-site data to the Board. The two dashboards were deliberately aligned to increase transparency on all sides and to ensure that "people up and down the flow of information were looking at the same information," Farrow says. He and his team continued to work with all sites, using the results database, to

Fig. 5: Sample Site Assessment Matrix (2008)

	SITE 1		SITE 2		SITE 3		SITE 4		SITE 5		SITE 6		SITE 7	
NEIGHBORHOOD POPULATION (2000)	19,557		31,702		39,374		14,798		38,718		133,646		28,373	
Number of Households	7,302		10,693		13,981		7,500		12,215		37,046		10,618	
Number of Children (0-18)	5,653		9,453		10,994		5,862		13,706		42,919		7,526	
Percent of Families with Children in Poverty	39.90%		26.20%		30.40%		56.00%		43.30%		38.00%		16.90%	
	DEC.	MAR.	DEC.	MAR.	DEC.	MAR.	DEC.	MAR.	DEC.	MAR.	DEC.	MAR.	DEC.	MAR.
RESULTS & STRATEGIES														
WORKFORCE														
Effectiveness & Comprehensiveness of Strategies (Scope)	9	9	9	9	9	8	9	9	8	9	8	8	10	10
Systems and Policy Reform (Scale)	10	10	10	10	9	8	8	8	8	9	8	8	11	11
Data and Accountability Systems	8	8	8	8	9	9	7	7	8	8	8	8	8	8
Co-investment in Each Strategy (Leverage)	8	7	9	9	8	8	9	9	8	8	8	8	9	9
ASSETS														
Effectiveness & Comprehensiveness of Strategies (Scope)	5	5	8	8	9	9	8	8	7	8	8	8	8	8
Systems and Policy Reform (Scale)	5	5	9	9	7	7	7	7	7	7	8	8	8	8
Data and Accountability Systems	7	7	7	7	9	9	7	7	7	7	8	8	7	7
Co-investment in Each Strategy (Leverage)	7	6	8	8	8	8	8	9	7	7	8	8	9	9
CHILDREN HEALTHY / PREPARED TO SUCCEED														
Effectiveness and Comprehensiveness of Strategies	8	8	7	7	8	8	7	7	9	9	8	8	8	8
Strategy for Systems and Policy Reform (Scale)	10	10	8	8	7	7	7	7	9	9	6	6	10	10
Data and Accountability Systems	8	8	7	7	9	8	7	7	8	8	8	8	6	6
Co-investment in Each Strategy (Leverage)	10	10	10	10	7	7	8	8	10	10	7	7	11	11
CIVIC PARTICIPATION AND RESIDENT NETWORKS														
Effectiveness and Comprehensiveness of Strategies	11	10	7	7	7	7	10	10	9	9	8	7	11	11
Strategy for Systems and Policy Reform (Scale)	10	10	7	7	8	8	9	9	7	7	8	7	9	9
Data and Accountability Systems	9	9	7	7	8	8	7	8	7	7	8	8	6	6
Co-investment in Each Strategy (Leverage)	10	9	6	6	8	8	9	9	7	7	7	7	7	7
TOTAL RESULTS AND STRATEGIES	**135**	**131**	**127**	**127**	**130**	**127**	**127**	**129**	**126**	**129**	**124**	**122**	**138**	**138**
MANAGEMENT														
Management for Results	9	8	7	7	9	9	9	9	8	9	7	7	11	11
Sustaining Strong Alliances	9	9	6	6	9	9	8	8	9	10	6	6	11	11
Effective Communication to Multiple Constituencies	8	8	7	7	9	9	8	9	8	9	7	6	9	9
Long-Range Financial Planning	7	7	7	7	8	8	7	7	6	6	7	7	7	7
Ensuring Effective Policy and Influence Strategies	9	9	8	8	8	8	8	8	7	8	8	8	8	8
TOTAL MANAGEMENT	**42**	**41**	**35**	**35**	**43**	**43**	**40**	**41**	**38**	**42**	**35**	**34**	**46**	**46**
SUSTAINABILITY														
Authentic Demand from Residents	10	10	7	7	9	9	10	10	9	9	9	8	11	11
Alliance of Partners and Champions	10	10	8	7	10	10	9	9	8	8	8	8	10	10
Strong Leadership to Sustain the Work	10	10	7	7	10	10	8	8	8	10	8	7	11	11
Development of Local Management Entity	9	10	8	7	11	11	8	8	8	9	8	7	11	11
Adequate Resources (Public, Private, Philanthropic)	9	8	9	9	8	8	8	9	7	7	8	8	8	8
TOTAL SUSTAINABILITY	**48**	**48**	**39**	**37**	**48**	**48**	**43**	**44**	**40**	**43**	**41**	**38**	**51**	**51**
TOTAL	**225**	**220**	**201**	**199**	**221**	**218**	**210**	**214**	**204**	**214**	**206**	**200**	**246**	**246**
OVERALL AVERAGE	**8.7**	**8.5**	**7.7**	**7.7**	**8.5**	**8.4**	**8.1**	**8.2**	**7.8**	**8.2**	**7.9**	**7.7**	**9.5**	**9.5**

update the data dashboards at least three times a year.

From an evaluation perspective, Kelly criticized many of the early efforts to collect data for management purposes. He believed in using data for management but disagreed with what he saw as leading with Board reporting—and with *Making Connections'* core indicators—rather than cultivating sites' desire and capacity to manage to data that mattered to them. "The forcing of a common dashboard actually undermined the sites having a stake in the numbers," Kelly believes. He also felt that the push for reportable results prompted a focus on short-term measures of change, which in turn encouraged a focus on programmatic rather than population-level outcomes. Farrow and his staff clearly needed short-term measures to keep trustees engaged and sites accountable. But, as Kelly says, "You cannot just add programmatic outcomes up to get population-level outcomes." The focus on programmatic change also inadvertently led some site teams, in Kelly's opinion, to assume that if they simply placed "X" number of people in jobs they would close the gap in employment. That kind of thinking didn't take into account resident mobility, regional job loss, poor job retention rates, and the fact that some gaps were so big they couldn't be solved programmatically.

In addition, the data dashboard scores didn't always square with what evaluators were seeing in sites, and the pressure to update dashboards quarterly, in time for Board meetings, sometimes interfered with the evaluation team's data requests of sites. The results database applied common metrics to 10 sites that were implementing different things in different ways. The data were self-reported and therefore vulnerable

to double counting and guesswork. Moreover, two or three sites reported all outcomes (job placements, for instance) in the result area, including those that would have occurred without *Making Connections.*

"Just looking at the [dashboard] numbers you could tell they didn't make sense," Kelly says. "[One site]'s numbers were always in the thousands. In [another site], as soon as the city employment division became a partner [of *Making Connections*] they began counting all their workforce development cases and just assumed that half of them lived in the target neighborhoods."

In retrospect, Ralph Smith says, it would have been better to include capacity measures in the data dashboards, along with the indicators of family economic success and school readiness. Since what gets measured gets done, the failure to highlight data on sites' capacities meant that the capacity-building work—while crucial for establishing the platform that would lead to population-level change—never got the lift and investment it should have, Smith says.

Farrow views the dashboards differently: "By the time we developed the dashboards, our concern was how the capacities—while also important in their own right—were being used to achieve results. That is what the dashboard was designed to communicate, and in sites where that connection of capacities to results was made, the dashboard was one way of reinforcing it and allowing people to explain their accomplishments."

Kelly found closing the gap especially hard to take as a measurement technique. He knew the size of the gap could change for many

Fig. 6: Sample *Making Connections* Performance Dashboard–Midyear 2009

NEIGHBORHOOD-WIDE INDICATORS				PERFORMANCE MEASURES								
RESULT: WORKFORCE	**TREND**	**2003-2006**	**2010 TARGET**	**RESULT: WORKFORCE**	**2006 ACTUAL**	**2007 ACTUAL**	**2008 ACTUAL**	**TREND 2006/8**	**2009 TARGET**	**2009 YTD ACTUAL**	**2010 TARGET**	**2011 TARGET**
Households with adult employed	↑	18% increase 65% to 83%	73%	Job placements	522	487	375	↓	400	90*	350**	350
Households with health benefits from job	↑	5% increase 30% to 35%	34%	Job placements with health benefits (%)	NA	60%	66%	↑	60%	52%	60%	60%
				Airport Jobs 12-month retention					84%	60%	60%	70%
RESULT: ASSET BUILDING	**TREND**	**2003-2006**	**2010 TARGET**	**RESULT: ASSET BUILDING**	**2006 ACTUAL**	**2007 ACTUAL**	**2008 ACTUAL**	**TREND 2006/8**	**2009 TARGET**	**2009 YTD ACTUAL**	**2010 TARGET**	**2011 TARGET**
Households with savings	↑	9% increase 68% to 77%	78%	Residents claiming EITC	349	431	499	↑	550	539	600	600
				Total EITC returns***	$572k	$618k	$676k	↑	-	$820k	-	-
RESULT: CHAPSS	**TREND**	**2005/06 2007/08**	**2010 TARGET**	**RESULT: CHAPSS FOCUS SCHOOL – MOUNT VIEW**	**2005-06 ACTUAL**	**2006-07 ACTUAL**	**2007-08 ACTUAL**	**TREND 2006/8**	**2008-09 TARGET**	**2008-09**	**2009-10 TARGET**	**2010-11 TARGET**
Children in quality preschool experience	↑	4% increase 46% to 50%	98%	Children having preschool experience (**all-WC/BP**)	169	304	275	↑	480	457	514	514
Children assessed ready for Kindergarten	↑	2% increase 21% to 23%	30%	Kindergarten children scoring at benchmark	23%	29%	31%	↑	39%	13%****	44%	49%
Children reading at 3rd grade level	↑	2% increase 53% to 55%	67%	Children 3rd grade WASL reading test – proficient	68%	52%	58%	↓	65%	58%	73%	77%
Children in K-3 grades absent 20+days	↑	2% decrease 12% to 10%	9%	Children K-3 grades absent 20+ days	12%	8%	7%	↑	4%	*****	2%	2%
				Children with health insurance					100%	87%	89%	92%
				Children receiving developmental screening					100%	100%	100%	100%
				Children with a medical home					100%	81%	85%	88%
RESULT: COMMUNITY BUILDING	**TREND**	**2003-2006**	**2010 TARGET**	**RESULT: COMMUNITY BUILDING**	**2006 ACTUAL**	**2007 ACTUAL**	**2008 ACTUAL**	**TREND 2006/8**	**2009 TARGET**	**2009 YTD ACTUAL**	**2010 TARGET**	**2011 TARGET**
Residents serving in organizations	↓	3% decrease 17% to 14%	20%	Residents serving in organizations	44	61	170	↑	90	108	100	110

* Midyear figures are lower than expected due to unavailable YTD FSET data; however, early trends reveal that the work is exceeding target.

** Targets ramp downward due to shift to influence, retention and advancement strategies.

*** The site does not set targets for the EITC amounts but rather number of EITC recipients.

**** This year's downturn suggests problems in test administration as the results are vastly different from the prior five years of trend data; we do know that the post-test administered in the spring of 2009 revealed 84% of the same kindergartners testing on grade level.

***** Data not released until Sept/Oct.

FAMILY ECONOMIC SUCCESS

## INFLUENCE ON POLICIES AND SYSTEMS TO ACHIEVE SCALE

Food Stamp Employment and Training Funds. The site's advocacy for FSET created a new funding stream that has brought in $5M and strengthened the neighborhood jobs pipeline (enabling providers to cross-refer and collaborate in serving families). The state recently expanded Food Stamp eligibility to 200% of poverty, increasing the number of residents eligible for FSET services.

Community Colleges. Local community colleges now offer education and training to participants in the neighborhood jobs pipeline to increase retention, advancement, and wage progression.

Data & Evaluation. The site will document the results of its job pipeline projects and FSET to make the case for job pipelines and on-ramps to our system partners and other funders, as a way to promote scale and sustainability.

LEVERAGE AND CO-INVESTMENT

Co-Investments in Millions for FES: Total $29.5M

126% INCREASE FROM 2005
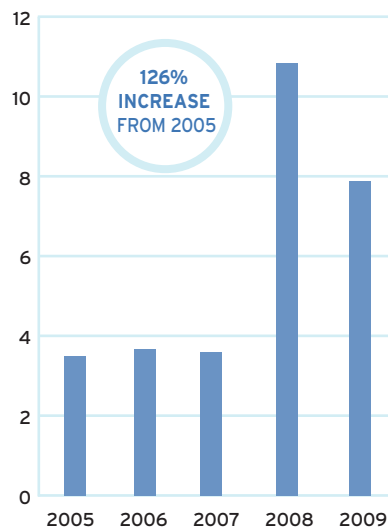
PRIORITY ACTIONS TO ADDRESS PERFORMANCE CHALLENGES

Expand neighborhood jobs pipelines. With the economic downturn, the site will help link residents to additional sectors and job opportunities by expanding its construction jobs project to include green jobs and infrastructure; and its health care vocational ESL project to include additional health care employers. Jobs projects will also help residents retain their jobs as well as achieve advancement and wage progression.

Promote adoption of neighborhood jobs pipelines by other funders. The site will promote adoption of the neighborhood jobs pipeline and on-ramp approach through the region's workforce funders collaborative. This includes participating as an investor in the collaborative and making aligned or matched investments. An evaluation will be done of the site's pipeline project to help make the case for jobs pipelines and on-ramps.

Expand FSET. The site will work with the state department of social services to build capacity around FSET and expand the number of residents served; and explore ways to maximize the use of FSET to support jobs pipelines. The site will document and evaluate FSET results.

# CHILDREN HEALTHY AND PREPARED TO SUCCEED IN SCHOOL

INFLUENCE ON POLICIES AND SYSTEMS TO ACHIEVE SCALE

Making Connections **plays an important role in influencing** the state public-private partnership with substantial funding from a local foundation to create a comprehensive early learning initiative, which will invest millions of dollars to assure universal access to high-quality early care and education and successful transitions to elementary school.

The site's intensive work with a cohort of families is demonstrating what it takes to assure family economic success and improved third grade reading for families in the neighborhood. Partners are using these data to guide policy and practice changes, and the number of families reached in 2009 will be expanded.

LEVERAGE AND CO-INVESTMENT

Co-Investments in Millions for CHAPSS: Total: $42.9M

84-FOLD INCREASE FROM 2005
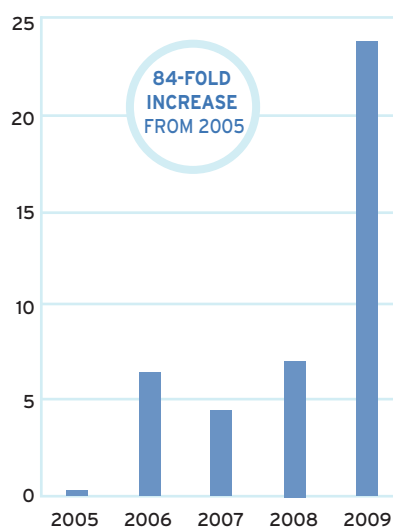
PRIORITY ACTIONS TO ADDRESS PERFORMANCE CHALLENGES

Expand early learning. Influence increase in number of early learning opportunities for children not enrolled in formal, licensed child care. Expand FFN support opportunities for resident caregivers and children birth to five.

Improve third grade reading. Adjust health insurance acquisition efforts and continue to support school district targeted interventions for literacy improvement efforts.

Increase leverage. Continue to present data from the family cohort to secure funding for continuation of two-generation approach for supporting children's health and academic success.

reasons that *Making Connections* didn't control—because a major employer left town, for instance. Moreover, sites tended to identify the gap (say, 1,000 parents who needed jobs) and average it across the five-year period remaining before *Making Connections'* transition out of Casey's management. With one or two exceptions, most didn't take into account that in order to get 200 people into jobs every year they probably would have to start with a far higher number, because retention rates are low and resident mobility is high. And there were issues with duplicate counting from year to year.

But it's not as if the evaluators had easy answers for what to report to the Board. They also struggled mightily with the issue of where to look for impact—at results for families who experience any one of *Making Connections'* strategies or an array of strategies? Results for systems that developed new capacities and more family-supportive ways of doing business? Results at the community level, or the local partnership level? "It was very hard to look at all those things simultaneously, and they weren't the same in each site," Kelly says.

"Our biggest failure was the inability for all technical assistance—by evaluation and by [*Making Connections'* Technical Assistance Resource Center, Farrow's team]—to address not only the technical and programmatic needs of sites but also their management capacity," he adds. "Instead of pounding them with dashboards, we should have been helping them manage to the performance data. We did not model the behavior."

From Frank Farrow's point of view as *Making Connections'* on-the-ground manager, using data and measurement for management

purposes—"results management"—became a crucial part of *Making Connections'* implementation. It takes deliberate attention for multiple change agents to stay on course toward the same target outcomes, and managing to results helps them do that, Farrow pointed out in a 2010 essay.[41] Moreover, Farrow says,[42] when done well and with the appropriate time and resources for training and alliance building, results management:

• Heightens accountability and makes it possible to justify specific methods and approaches to multiple partners. It's always difficult to make the case for community change, but the careful assessment of performance that results management brings to the table reassures other funders and decision-makers.

• Keeps the focus on core outcomes for families and children when other factors pull attention away.

• Make visible, often for the first time, whether the theory of change can actually produce results.

• Surfaces equity issues in a way that forces people to deal with them, because the process involves examining results for specific populations.

• Can help other partners and investors recognize the centrality of community capacities that they might otherwise ignore.

Farrow's perspective on the value of the results reporting within *Making Connections* differs from Kelly's. He points out that the requirements for reporting to the Board about the initiative's management evolved. The site

assessment matrix was an appropriate first tool because "at that point we were trying to send the message to the Board and to sites that the investments had to be in strategies that were simultaneously focused on the result areas and on capacities (which were the foundation for getting to results)," Farrow says:

> The site assessment also ranks sites. At first that bothered me, but I came to understand that our Board was interested in comparative rankings because what they were really worried about was which sites were not able to meet expectations....In the written reports that accompanied these summary sheets, and in our discussions with the Board, the focus would be on what we as managers were doing to help sites that were falling behind.

> When we moved to the data dashboard, we moved to a different level. That format tried to address three questions: is there evidence that there's a difference being made for families? What are the ways in which sites are using this evidence to influence public and private funders and systems? And is anyone else co-investing in this work? By that time in the initiative, it wasn't good enough to just say we were building capacities; we needed sites to say how they were using the capacities to influence their other partners. That's why the second page of the site assessment mechanism is all about influence [and] challenges.

> In both of those efforts, we were trying to keep intact our core supposition that success requires attention to strategies for achieving results and to the capacities that allow sites to implement those strategies. That sounds obvious now, but it wasn't always obvious to everyone…We developed the term 'double helix' to communicate that investing in capacities was inseparable from getting durable results at ever-greater scale.

> To my mind, managing with such a strong emphasis on results (and the capacities to achieve results) was about urgency. Week to week, it could seem to be just about preparing another report for the Board, or for the Foundation's senior management, or for discussion with sites. Without all of us regularly reviewing whether we were making a difference, it was possible to get side-tracked by the complexities of implementation.

> In addition, results management provided evidence that the work by people on the ground was making a difference. The sense of making a difference is ultimately what keeps people engaged in the very tough work of community change, and it was important when it came to making the case to community partners that this work deserved to be sustained even as Casey's investment phased down. Information from the local or cross-site evaluations simply wasn't available in a timely enough way to provide input into critical local decisions about sustaining the work.

During this time, the evaluation and implementation staffs made some effort to communicate better about what they were doing, learning, and requiring of sites. Tom Kelly, Mary Achatz, and Scott Hebert were invited to join the implementation team's monthly meetings, facilitated by Frank Farrow and attended

by *Making Connections'* technical assistance providers. For the evaluators, however, that step was inadequate to produce real coordination, mutual understanding, or integration of the evaluation and performance management strands of work. The evaluators recall having only a few minutes at the end of a long agenda to report on their activities, and often they were bumped to extend the discussion of pressing implementation concerns. Even when stimulating conversation occurred between the evaluators and implementers, there was not enough follow-up after the monthly meeting.

MEASUREMENT STRATEGIES AT ODDS

Making Connections' **management and evaluation teams worked alongside each other for several years but never fully resolved their differences** in data collection and reporting. For Tom Kelly, the difference in views are illustrated by an event that occurred in January 2007 during a planning retreat for technical assistance providers, local site coordinators, and site liaisons. While Frank Farrow was commending Seattle/White Center partners for achieving 1,000 job placements—part of the process of "closing the gap"—Kelly jumped up and strode to a flip chart. Scribbling numbers across the page, he spoke rapidly:

> *Folks, there are 31,000 people who live in White Center. Let's assume half are families with children [i.e., the population Making Connections cares most about]. Now we're looking at 15,000 people. Of those families, let's assume ¼ of them are working and struggling but their kids are doing well in school and we don't have to worry about them. So we're*

"I suspect there are inevitable tensions between the goal of getting a real-time indication of what is occurring, especially in terms of a complicated concept like the interaction of capacities and results, and the longer-term goals of evaluation."

– Initiative Manager Frank Farrow

> *really just talking about 12,000 people, maybe 6,000 families. You told me there are 1,000 job placements. Half are single adults. So you've placed 500 family members out of 6,000 families. Nobody here has shown me a strategy yet that's going to reach 6,000 needy families. If you want this evaluation to show change, I need bigger numbers. Otherwise we've just got a nice narrative story.*

To Frank Farrow, this illustrated "the great strength and the great challenge" of the closing-the-gap analysis. Its strength lay in helping sites set high aspirations and specific targets for what they wanted to achieve. Everyone could grasp the fairness and equity arguments in closing the gap: Children in low-income neighborhoods should have the opportunity (and experience the reality) of doing as well as children elsewhere in the city or metropolitan area. But Farrow also understands that, in measurement terms, closing the gap creates complications because the minute you start to measure the gap, it changes. Families' mobility in and out of the target neighborhoods was especially problematic, he notes, but "contrary to what the evaluators say, we did try to address this.

We developed and distributed to sites, early on, a paper on how to adjust targets and calculate what it would take to close the gap related to employment. This was also part of the training we provided on the results database."

Some sites, notably Seattle/White Center, took such calculations into account when setting close-the-gap targets. Nonetheless, Farrow says, the approach is best seen as "the framework for an aspiration rather than for data reporting," at least until a more complex data tracking process exists that enables sites to track all of the factors that affect outcomes.

For Foundation and initiative leaders like Ralph Smith, the move toward using data for management purposes was a positive step that warranted everyone's support. After all, UPS's former executives sit on Casey's Board, and UPS culture requires knowing where every package is, how each package will get from Point A to Point B to Point C, and what everyone needs to do to move the packages along—all of which requires managing with data. Smith describes his revelation about using *Making Connections'* data for management this way:

> *For the first few years I thought we just had to get better about finding the data and making sure they were clean and told the right story. We were trying to get to a KIDS COUNT-type level of clarity, accuracy, and consistency with the Making*

Connections' data, and I remember being increasingly proud of that.

I didn't fully appreciate that our Board wanted to know how we were using the data to manage. We would be prepared to tell them the good news about how much progress we made and why. But that was ho-hum to them; that's the easy part of managing. They focused on the red arrows—the misses. They wanted to know that we knew why and, as important, had used the data to develop corrective action. In many respects, we wanted to celebrate progress. They want us to focus on that bottom third that either didn't move, didn't move enough, or moved in the wrong direction. What they wanted was to push us to use data to drive a set of questions, resources, and attention toward pulling the bottom third up.

If only the evaluators, implementers, and managers had been as clear as the trustees about the multiple purposes and uses of evaluation data from the beginning, Smith says, "We could have made a more powerful case up and down the food chain for why data were important."

Fig. 7: Summary of Evaluation and Related Costs for *Making Connections*

2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	TOTAL
CROSS-SITE EVALUATION (PROCESS EVALUATION, CORE CAPACITIES ASSESSMENT)										
	$400,000	$800,000	$590,000	$580,000	$727,000	$750,000	$1,008,000	$888,000	$950,000	**$6,693,000**
SURVEY (DESIGN, DATA COLLECTION, DATA ANALYSIS, TECHNICAL ASSISTANCE TO SITES, & CREATION/DISSEMINATION OF SITE-SPECIFIC & CROSS-SITE PRODUCTS)										
$300,000	$2,450,000	$3,100,000	$2,100,000	$3,000,000	$2,800,000	$2,300,000	$3,200,000	$2,700,000	$1,350,000	**$23,300,000**
CROSS-SITE INDICATORS DATABASE										
$0	$0	$120,000	$309,000	$85,000	$95,000	$148,000	$197,000	$198,000	$45,000	**$1,197,000**
SUPPORT TO LOCAL LEARNING PARTNERSHIPS (DATA WAREHOUSES, DATA FOR LOCAL PLANNING, RESIDENT TRAINING IN DATA USE, LOCAL EVALUATIONS)										
$1,500,000	$2,500,000	$3,500,000	$3,500,000	$3,000,000	$2,750,000	$2,750,000	$3,400,000	$2,125,000	$1,250,000	**$26,275,000**
TECHNICAL ASSISTANCE TO LOCAL LEARNING PARTNERSHIPS										
	$250,000	$500,000	$500,000	$350,000	$300,000	In 2007, control of this budget transferred from the evaluation unit to Casey's Community Change Initiatives unit.				**$1,900,000**
$1,800,000	$5,600,000	$8,020,000	$6,999,000	$7,015,000	$6,672,000	$5,948,000	$7,805,000	$5,911,000	$3,595,000	**$59,365,000**

VII

SHIFTS IN THE UNIT of ANALYSIS
(2004-2005)

Making Connections' **design went through several evolutions that changed the focus of work on the ground and, consequently, what and how the evaluators measured change.** Around 2002, for example, the strand of *Making Connections* that focused on cultivating families' economic success shifted from a focus on quick wins, such as increasing Earned Income Tax Credit returns in low-income neighborhoods, to a longer-term focus on strategies for developing job pipelines between the neighborhoods and regional employers, increasing family incomes and access to benefits, and helping families acquire assets. Similarly, a year or two later Smith articulated *Making Connections'* recognition that it takes a two-generation approach to break the cycle of intergenerational poverty. By this, he did not mean it would take two generations to achieve measurable results for children. He meant that the initiative had to focus simultaneously on at least two generations—parents and children—within the families living in the targeted communities. Parents had to be able to succeed in the economy, and children had to be healthy and prepared to succeed in school. Those goals (and later, third-grade reading proficiency as a pivot point for future success) thus became priority outcomes for *Making Connections,* and implementers were encouraged to view the dual outcomes not as two parallel tracks but as intertwined.

Perhaps the most significant changes for evaluators, however, had to do with whether the unit of analysis would be child and family outcomes for the entire neighborhood population or whether the evaluation would attempt to capture changes in child and family outcomes specifically for participants in *Making Connections*-related programs and services. In 2004–05, while *Making Connections'* evaluators and implementers were pursuing their separate agendas—measurement for evaluation purposes and for results management—these larger issues came to the fore. They caused a further refinement of *Making Connections'* target outcomes; were exacerbated by changes in the initiative's design; prompted the development of two new implementation tactics that required measurement tools, Participant Family Data Collection (PFDC) and focus schools; and repositioned the cross-site survey as a tool more for research than evaluation.

REFINEMENTS IN TARGET OUTCOMES

Cindy Guy presented the first data from the Wave I cross-site survey to the Board in early 2004. The trustees, Guy recalls, were underwhelmed. "You just showed us that the residents of *Making Connections* neighborhoods are poor, in about 350 different ways," they said. "We already *knew* that." Doug Nelson responded that it was important to document the

conditions of children and families at the outset of an initiative to create a baseline for measuring progress over time. Board members were receptive to this explanation and to Guy's point that descriptive data on children, families, and neighborhoods were necessary to help explain changes in bottom-line results. Nevertheless, the Foundation's senior staff left with a clear message that Board members did not want to be deeply involved in the early and intermediate steps of the analytical process. They wanted a manageable set of indicators that would enable them to monitor change.

So Smith, Farrow, Guy, and their colleagues quickly mobilized to produce a set of **streamlined indicators** (see Fig. 8 below) that met the Board's desire to know about specific outcomes without confusing the issue with so much tangential data. The process for developing

streamlined indicators was more pragmatic than scientific. Smith, working with the trustee chairing the Performance Measures Committee proposed indicators having to do with family economic success and children's readiness for school. Guy, Farrow, and Bob Giloth, head of Casey's unit on Family Economic Success, then tried to find where those preferences converged with available data and with what it would be useful to know from a best-practices standpoint.

CHANGES IN INITIATIVE DESIGN

Around 2005, following different paths, Ralph Smith and Frank Farrow came to similar conclusions that _Making Connections_ should stop aspiring to population-level change within a single decade. Smith was influenced by what he calls "demographic surprises":

Fig. 8: _Making Connections_ Core Result Areas and Streamlined Indicators

RESULT AREA (AND POPULATION)	STREAMLINED INDICATOR
Families have increased earnings and income. _(Families in Making Connections neighborhoods)_	% of households with children who report earned income and one or more adults employed
	% of households with employer-provided family health benefits
Families have increased levels of assets. _(Families in Making Connections neighborhoods)_	% of households with children who have accumulated savings
Children are healthy and prepared to succeed in school (CHAPSS). _(Children in Making Connections neighborhoods)_	% of children in quality preschool programs*
	% of children assessed as ready for school
	% of children's attendance in the early grades
	% of children reading at or above proficiency level in third or fourth grade

SOURCE: "Evaluation Design Components," draft paper by the Association for the Study and Development of Community (now called Community Science), December 2008.

*"Quality" is not defined.

- Findings on family mobility, derived from the survey, indicated that a large proportion of residents in *Making Connections* neighborhoods weren't staying long enough to be deeply touched by the initiative;

- The primary strategy in one *Making Connections* site (workforce development) had come to a halt because there weren't enough adults in the selected neighborhood who had the level of literacy required to enter the job training and placement pipeline; and

- Several sites had a large subpopulation of residents who were ex-offenders returning from incarceration—people with complicated skill-development issues and employability constraints—which made it difficult for *Making Connections'* economic self-sufficiency strategies to achieve widespread success.

Smith also was concerned that progress in developing the core capacities essential to *Making Connections'* long-term sustainability was scattered and inconsistent. Moreover, he realized that even Casey's investments in East Baltimore and Atlanta, which were dramatically larger than the investments in *Making Connections* sites, did not appear likely to produce population-level changes.

Farrow, meanwhile, had realized that *Making Connections* was not going to achieve population-level change within a decade given the pace and complexity of implementation and the resources available. It wasn't just a matter of doing the job better, he explains:

> *In every result area we had to invent new ways of going about the work. To help get the hardest-to-employ residents into jobs, for example, we had to help sites invent*

neighborhood workforce pipelines. We had to prototype the new approach on a small scale, show that it worked, and then use that evidence to influence everybody from the Workforce Investment Board to the United Way to adopt the approach, because it was their efforts that would perhaps lead to population-level change. In San Antonio, White Center, and Providence, that theory of 'scale through influence' is still reflected in their work and has potential for expanding results.

Farrow also worried that there was no way to tell whether enough children and families in the selected neighborhoods were getting enough "touches," in the correct sequence, to change outcomes at the population level. For *Making Connections'* school readiness work, in particular, he was concerned that the boundaries of the selected neighborhoods didn't align with school attendance boundaries, and so efforts to improve neighborhood schools might not reach enough residents to affect the population. And the sites' weak, non-existent, or antagonistic relationships with school districts undermined the potential to expand *Making Connections'* promising early childhood interventions, where they existed, into a sequenced set of school-age actions that would enable all children to read by the end of third grade.

Smith and Farrow's response was to shift the initiative's focus to producing enough results, and embedding powerful enough strategies, that communities would be able to move their own populations toward results over a period of time that extended beyond *Making Connections'* decade. They called this "building a platform for results" and began to see it as *Making Connections'* most important leave-behind—the ticket to "sustainability."

Smith redistributed to Casey staff and Farrow's consultants a paper called "Public Capital: The Dynamic System," prepared by the Harwood Institute for the Kettering Foundation in 1996, which had influenced his own early thinking about *Making Connections.* He urged everyone to pay more attention to building the tools needed to show whether sites were making progress along the path outlined by Harwood. He encouraged Kelly and Farrow to think about theories of scale and sustainability as related to, but potentially separate from, theories of change. And he commissioned a study by former Indianapolis Mayor Stephen Goldsmith and retired Boeing executive Bob Watt to explore whether elected officials and other prominent stakeholders in *Making Connections* sites perceived that an enduring platform for achieving results had taken root.

Smith and Farrow also began to envision a "more calibrated and disaggregated" notion of *Making Connections'* level of change, characterized by:

• Measurement among a cohort of 100 to 200 neighborhood residents who were most touched by the initiative, rather than the entire population, and tracking their outcomes over time regardless of where they moved (Participant Family Data Collection);

• A focus schools approach, which concentrated on improving education outcomes in one school per *Making Connections* neighborhood and tracking outcomes for a subset of children who attend the school; and

• A commitment to "stay the course" on core results as a way to manage the initiative and to build the results culture needed for the long-term "platform."

PARTICIPANT FAMILY DATA COLLECTION (PFDC) AND FOCUS SCHOOLS

The PFDC strategy was proposed by Tony Cipollone, and it was well-liked by local coordinators who wanted family-level data in real time, but it still made Tom Kelly uncomfortable. *Making Connections* was supposed to be about community change and systems reform, not about providing services directly to families—yet the push to measure the most-touched families seemed, to Kelly, to reverse that assumption. Was the initiative's theory changing, he wondered? "It felt like we'd been selling the wrong thing," Kelly recalls. "We were supposed to be about the mortar, not the bricks. And providing services is about the bricks."

Moreover, because *Making Connections* hadn't been providing anything as concrete as social services, local partners had trouble identifying "touched" (or "participant") families for the evaluators to track. When asked to administer the PFDC survey without confidentiality protections for respondents, NORC staff declined on ethical grounds. Cindy Guy recognized that the PFDC was more of a management information systems project than the kind of research survey that falls within NORC's expertise, so it fell on Kelly and his evaluation consultants to handle this data strategy.

It didn't help that PFDC's implementation was flawed. In Des Moines, the local partner who identified and interviewed families for the PFDC sample found that the families now expected to receive services—something *Making Connections* wasn't prepared to provide. Denver compiled a 900-person database but only a handful of people in it had attended more than one *Making Connections*-sponsored event. Louisville had 3,000 residents who had signed

up to be part of a "network" but couldn't find 20 people who would agree to be interviewed.

From Frank Farrow's point of view, PFDC was implemented prematurely, when sites were just beginning to get traction with their strategies. He believes that PFDC would have been a better data collection strategy later in the initiative.

In the end, Kelly came to view PFDC not as a measurement strategy but as an intervention. "It forced Casey and the sites to get clear about what the definition of success is," Kelly explains. "For me, PFDC unearthed the question of what scale and reach we are trying to achieve. Is it only about getting to scale for the most vulnerable families? For participant families? For all families?"

The second technique for capturing child- and family-level change, "focus schools," attached extra data collection activities to the strand of *Making Connections* known as CHAPSS (Children Healthy and Prepared to Succeed in School). In 2003, Ralph Smith had grown concerned that the effort to make children healthy and prepared to succeed in school had fallen off the radar for people working in the sites. At a cross-site meeting that summer, he recalled the common core outcomes, noting that school readiness "is not an additional outcome, it has always been one of our targets." He went on to challenge the group: "It's not acceptable to say, 'We'll think about it next year.' The children who will enter school next year were not born when we started *Making Connections.*" Site teams responded by focusing on a range of early childhood supports, but by 2005 schools were still not part of the effort. *Making Connections'* technical assistance providers tended to have

expertise in early childhood development and child care, which are primarily programs. Virtually none were knowledgeable about schools or had working relationships either with individual schools or school districts.

Farrow's team and sites responded in several ways. They sharpened the definition of what it would take to achieve the child health and school readiness goal by developing the CHAPSS framework, depicted both graphically and in a paper. Then they held cross-site meetings to train local participants in the framework and share evidence-based information on implementation practices, delivered by knowledgeable technical assistance providers.

When Farrow realized that sites' CHAPSS efforts were reaching many children but not having the cumulative effect of preparing them for school, he proposed concentrating the efforts on one school catchment area per *Making Connections* neighborhood and tracking outcomes for a subset of children who were slated to attend a school in the catchment area. Farrow's intent was to combine a series and sequence of interactions for the same families and children, which he thought was essential to ensure the key goal of achieving grade-level reading proficiency.

Ralph Smith, Tom Kelly, and Doug Nelson all expressed concerns that focusing on a single school or catchment area per site would shrink the unit of change and therefore the overall significance of *Making Connections.* But Farrow argued that this approach was consistent with the overall theory of reaching scale—to prove it was possible to make a difference for a group of children by concentrating multiple interventions on the same kids and families. Otherwise,

interventions were not intense enough to make a difference for any one child or family. This approach also would allow *Making Connections* to build better relationships with schools and school district, Farrow believed.

The focus school strategy had implications for the evaluation. Focus schools were expected to link and integrate strategies for literacy development, child and family health services, individual student learning plans, after-school instruction, parent engagement, and other key ingredients to improve third-grade reading outcomes. This involved collecting student data and establishing data-sharing agreements between the school or district and *Making Connections.*

From an evaluation perspective, the problem was more with nonsystematic implementation of the focus schools strategy than its relevance for measurement purposes. Site teams were limited in the schools they could select, because most catchment areas encompassed only one or two. Given the poverty of *Making Connections* neighborhoods, the schools were generally weak to begin with. One closed after the first year of data collection, requiring data collectors to begin again from scratch. Other sites failed to notify the principal, school district, or superintendent that a school had been chosen, making it difficult to extract the necessary data. Because of district busing policies, some schools were attended by large numbers of children who didn't live in the target neighborhood(s), making it hard to attribute any changes to *Making Connections'* presence.

And, to be fair, by the time PFDC and the focus schools data strategy were deployed Casey's local partners were fairly inured to the

arrival of yet another measurement approach. "We threw so many measurement and assessment [structures] at sites," Kelly admits. "The common reaction was, 'Oh great. Another Casey matrix, another Casey framework.'"

The focus schools strategy became stronger after Casey contracted with Foundations, Inc., to provide intensive technical assistance in the classroom around core education issues. To further boost data collection, in November 2009 Casey contracted with Metis Associates to help eight *Making Connections* sites track results for a cohort of identified "striving readers" in the second and third grades, identify a common set of core indicators that all sites could provide, create a system for sites to submit common data, analyze the data periodically, and report findings on student progress across sites that reflect the focus schools' target outcomes.[43]

REPOSITIONING THE CROSS-SITE SURVEY

While these changes in focus and measurement were under way, the survey was subtly growing apart from the larger *Making Connections* process and outcomes evaluation—mainly because the evaluation's analytical agenda did not include questions that could be addressed with community-level data. As the initiative managers increasingly moved away from the aspiration to "move the needle" on child and family outcomes at the neighborhood level, and the evaluation team followed suit by focusing on programmatic "touches" and core capacities assessment, neighborhood-level survey data ceased to be the relevant measure of *Making Connections'* success or failure.

Accordingly, as survey team members prepared for Wave II they revised the survey instrument to generate more data of immediate relevance to local planning and service delivery. A major innovation, made at sites' urging, was to expand the battery of questions about children from one "focal child" to include all children living in the household. This change produced a larger body of data on children to use in planning and monitoring the CHAPSS agenda.

The survey team and research advisors also developed a policy research agenda that put to use the survey's rich, unique dataset on low-income children, families, and communities. As Guy explains, "It would be too much of a waste to let the survey data just sit there."

Viewed this way, the survey data produced a revealing portrait of residents of low-income neighborhoods; changes in their conditions, attitudes, and perceptions; and generalizable trend data on social, economic, and demographic conditions in their communities and metropolitan areas. For example, a lot of families in *Making Connections'* low-income neighborhoods moved frequently. The initiative's designers anticipated this, but the survey analysts' findings were still startling: The mobility rate among families in target neighborhoods was more than 50 percent. Half moved within three years, and half moved more than a mile away. Survey team analysts Margery Turner, Brett Theodos (also of the Urban Institute), and Claudia Coulton (Case Western Reserve University) set out to learn *why* the families moved, as Guy explains:[44]

> *Urban Institute researchers identified 27 distinct data items in the survey that corresponded to the push-and-pull factors that the field of mobility research views*

as significant contributors to mobility. These included data on household composition and demographics, economic characteristics, tenure in current housing arrangement, receipt of public assistance, neighborhood perceptions, and social connections. The analysts examined these data for each "mover" household before and after the move. Then they looked for clusters of "movers" that shared defining characteristics.

Using this approach, the analysts created a typology of different categories of movers, distinguished by the negative or positive motivation for the move (e.g., were respondents moving to opportunity or reacting to trouble?) and the negative or positive outcomes (e.g., did they end up better off or worse?). Then the analysts mapped out the distribution of these different types of movers within and across neighborhoods and within and across racial groups. The typology doesn't capture all of the complexity and variety inherent in resident mobility, but it does provide some insight into the question, 'Why did they move?'—and it does so in a way that helps us understand how mobility decisions are affected by place and by race.

To complement the survey data in these analyses (and, earlier, to inform the choices *Making Connections'* planners made about site and neighborhood selection), the Urban Institute compiled and analyzed data from the U.S. Census and American Community Surveys, ZIP Business Patterns, Internal Revenue Service, Home Mortgage Disclosure Act, Bureau of Labor Statistics, Federal Deposit Insurance Corporation (FDIC), National Association of Realtors, U.S. Housing and Urban

Development's median income files and housing assistance files, and automobile registries, which they analyzed at the neighborhood and county levels. The Urban Institute also inventoried the data available from participants in the National Neighborhood Indicators Partnership in *Making Connections* sites.

Urban Institute produced a series of analyses over the course of the initiative, using both survey and national data (and sometimes both). These reports addressed metropolitan conditions and trends; social connections; immigrant integration in low-income urban neighborhoods; income and employment; assets, debts, and financial services; primary school trends in *Making Connections* neighborhoods; and shifting trends in labor markets in *Making Connections* metropolitan areas. The mobility analysis, however, was especially compelling. After all, *Making Connections* aimed to help families by strengthening their neighborhoods, and a very large number of families were leaving the neighborhoods—even if many resettled nearby. In that respect, the survey data "added to initiative managers' impetus to advocate for broader policy change, because clearly we're not going to make a difference for families by confining our efforts to small neighborhood areas," Guy observes.

In retrospect, the introduction of PFDC and the focus schools strategy signaled two important developments in the evolution of *Making Connections* and its evaluation. First, the survey—Casey's big, expensive bet for capturing change—was not going to show *Making Connections'* effects at the population level because the interventions weren't operating at enough scale to show population-level change—the kind of change detectable in a randomly

selected sample. The survey continued to be a tool for data-based planning and advocacy, but Guy, Kelly, and their colleagues stopped talking about it as the source of outcome data for the *Making Connections'* evaluation.

Second, the unit of change being sought by *Making Connections'* designers was shifting from community-level change within the decade to a longer-term strategy of smaller-scale efforts that could generate the evidence needed to influence longer-term, broader change. And the lack of coordination and communication between *Making Connections'* implementers and evaluators regarding that shift created a disconnect that would have significant consequences. Even though one of the main reasons for managing the evaluation in-house had been to keep the evaluation and implementation activities closely integrated, Casey's own management structure got in the way. Tom Kelly reported to Tony Cipollone, who reported not to initiative architect Ralph Smith but to then-President Doug Nelson, who wasn't as closely involved as Smith in *Making Connections'* design aspects. Farrow reported to Smith. Moreover, up until 2003 Casey convened annual mid-year reviews, attended by all staff, that detailed the work in every *Making Connections* site. Those sessions made it easy for people to stay informed about what their colleagues were thinking and doing, but when the Foundation switched to a more streamlined review process, staff were less likely to overhear their colleagues' discussion of issues and strategies.

And so, as Smith now acknowledges, the split widened:

> *'Population-level change' is evaluation-speak. That is not how those of us with design and implementation*

responsibilities talked about the work among ourselves or with the sites. Consequently, we [Frank and I] never said in so many words [to the evaluators] that the intensified focus on platform building and sustainability and leap-frogging to scale and all the rest meant that we were no longer committed to population-level change within a decade. And since we continued to use 'closing the gap' as a basis for target setting and reporting, I can see why this could be the point at which Making Connections' implementation and evaluation diverged. What was being evaluated was no longer what the initiative was about.

VIII

DATA DRESS REHEARSALS
(2007-2009)

Although Casey's evaluation staff and consultants were well-experienced in the art and science of evaluation, most of the people implementing *Making Connections*—from the Foundation's own site team leaders to local coordinators—had little or no experience working on evaluations or being evaluated (including most of the 15 original site team leaders). So it wasn't surprising that the sites' local evaluation plans were scattered all over the place in terms of what they would capture and how.

It was Tom Kelly who came up with the idea for data dress rehearsals, which he initially called fire drills, to focus people on the evaluation story. "If we pulled the fire alarm right now, what would we be able to say about each site?" he wondered. "What's the story to be told? How can we make the case for a *'Making Connections* way of work?'"

The site visits that Mary Achatz and Scott Hebert conducted in 2006 to use the core capacities assessment tool had shown how useful local discussions about what site teams were trying to accomplish, and whether they were succeeding, could be. Why not conduct a new set of site discussions to refocus local teams on the initiative's theory of change? The conversations could also ensure that, if people couldn't say anything about what had happened, data

collection and evaluation activities could be tweaked over the next three years to make sure they *could* say something at the end.

The prototype fire drill, facilitated by Kelly, Achatz, and Hebert, occurred over two days in Indianapolis in December 2007. Participants listed the "hot" items in their site (strategies or activities that produced measurable progress or had considerable potential to influence large-scale change), "cold" strategies or activities (those that didn't work out as planned), and "lukewarm" strategies (those that had potential but faced barriers). Then they examined all of the items to see what they had in common. Most teams were able to name factors that looked a lot like core capacities, and some also identified factors that had to do with *Making Connections'* implementation approach.

The fire drill was a hit, both among local participants and the evaluation team. Following advice from a participant who found the name "fire drill" stress-inducing, Kelly renamed the event a "data dress rehearsal," and Achatz and Hebert, joined by implementation staff, took it on the road to every *Making Connections* site at least once in 2008. (Participants in Providence found the experience so valuable they held separate data dress rehearsals to assess their work on family economic success, children healthy and prepared to succeed in school, and resident leadership.)

For sites, the dress rehearsal process involved four tasks:[45]

• Deciding whom to involve in the session (a combination of site team and data partnership members, representatives of the local management entity, neighborhood residents, other key collaborators, and management and evaluation liaisons);

• Creating a timeline of major milestones in the site's history (e.g., major investments and community forums or events as well as social, political, and economic forces that facilitated, hindered, or otherwise shaped *Making Connections'* strategies) and in the site's interactions with Casey;

• Identifying a preliminary list of strategies and activities considered to be hot, cold, or lukewarm; and

• Spending two full days in discussion—the first to review the local initiative's history and discuss the most significant accomplishments and learnings to date and the second to explore implications for the site's data collection, analysis, and communications strategies.

After each dress rehearsal, Achatz and Hebert provided extensive written feedback. Site teams, meanwhile, were expected to produce three documents: a revised, electronic version of the local history timeline; a summary narrative of the hot, lukewarm, and cold strategies/activities with hypotheses for why some worked and some didn't; and an evaluation/data plan that linked *Making Connections'* principles, practices, and investments to improved results for neighborhoods, families, and children—that is, "an action plan for gathering and presenting the evidence necessary for 'telling the story.'"[46]

The data dress rehearsal represented the evaluation team's conviction that the process for developing and using a measurement tool matters as much as the tool itself. The data dress rehearsal had value as a way to focus people's thoughts, mobilize them to look back and plan ahead, communicate with each other and with the Foundation, and clarify both outcomes and definitions of success. It also addressed the issue of making comparisons between sites, something that had plagued *Making Connections'* evaluation from the beginning, by taking the effects of local context into account.

"I felt there was a home run [with the dress rehearsal process] that no one noticed," Kelly says:

> *If we had done it without involving the sites in the process—if Mary and I had written the matrix and gone out and scored the sites against it—I think we would have missed something. I think it's more important now to have a locally relevant ruler [to measure change] than to worry about how we compare sites, as long as there's some accountability check. This was also important because we were bringing people not just into Evaluation 101 but into the most complicated thing that we, professional evaluators, had ever evaluated.*

FINAL YEARS
(2008-2010)

Between 2008 and 2009, *Making Connections* sites underwent a long-planned transition from Casey's direct, hands-on management to local management. This process began in 2005, with initial discussions about what Casey meant by "sustainability." It was guided by two papers developed by the Community Change Initiatives unit, giving guidance on the definitions to sites and talking about what needed to be sustained. The transition was a difficult process that involved reaching consensus on Local Management Entities (LMEs) among a wide array of partners, renegotiating local roles and partnerships, and, often, reintroducing the initiative and its goals to a whole new set of players. By the end of 2009, however, all LMEs were in place and beginning to serve as "accountability managers" for *Making Connections* results—approving and tracking the investment plan, making corrections when performance targets weren't met, and reporting to both the community and Casey, which would phase down its funding through 2011.

In the midst of this transition, the economic crash of 2008 led to significant changes in *Making Connections'* focus and expectations for results. In an environment where even middle-class workers were losing their jobs, it did not make sense to hold *Making Connections* sites to the high targets for family economic success that they had set related to employment and savings. Initiative leaders told the local teams that, while holding true to the same core results of family economic success and children healthy and ready to succeed in school, they could use their Casey dollars to respond forcefully to the economic crisis. Many sites did so by increasing activity related to foreclosure prevention or remediation. This resulted in greater variation in strategies across sites, but sites got the clear message that being responsive to local needs was consistent with the two-generation results framework.

The greater variation in site strategy, combined with stronger local management, moved *Making Connections* toward being deeply rooted in each individual site rather than being a single, multi-site initiative. In a sense Casey had come full circle, returning to *Making Connections'* origin as the demonstration project for an overarching initiative, Neighborhood Transformation/Family Development rather than an "initiative" in its own right.

At the time of this case study's publication, *Making Connections* is still an ongoing endeavor. It is not too early, however, for some of the key players in this story to comment on the "gives and gets" inherent in the initiative and its evaluation.

How *Making Connections* Played Out in Key Areas

....

From documents developed by *Making Connections'* management (not the evaluation), 2008–09.

EARNINGS, INCOME, AND ASSETS

Workforce pipelines: Between 2005 and 2008, in 10 sites, *Making Connections* placed 10,897 residents in jobs.

Earned Income Tax Credit (EITC) campaigns: Between 2003 and 2008, federal tax returns were prepared for 1,169,981 residents across *Making Connections* sites, generating $545 million in EITC claims and $208 million in Child Tax Credits.

Work supports: A project to help help workers access Food Stamp Employment and Training (FSET) benefits generated $10.7 million in resources, across eight sites, during its first year of operation.

Family savings: By late 2008, more than 10,663 households across seven full-implementation sites had established savings, checking, and/or Individual Development Accounts. In one site, agreements with four of the state's five subprime lenders helped borrowers reclaim more than $6 million in assets.

Asset accumulation: Across sites, 341 families became homeowners for the first time. In one site, a multi-funder partnership supported development of 278 affordable housing units and 18,000 square feet of commercial space for small-business owners and community-based organizations.

CHILDREN HEALTHY AND PREPARED TO SUCCEED IN SCHOOL

Preschool experiences: Between 2005–06 and 2007–08, the number of children receiving pre-school services grew from 1,558 to 1,999 across 10 sites.

Grade-level reading: School-based interventions over two years increased the percentage of third-graders reading at grade level in three sites (e.g., from 11% to 33%, 24% to 36%, and 45% to 54%).

Child care: Five sites organized networks of family/friend/neighbor care providers to improve both the quality of child care and the capacity of these small businesses.

Health and school readiness: In one site, links with health screening programs for young children boosted the countywide rate of enrollment in the state Children's Health Insurance Program from 58% to 70%. In another site, *Making Connections* contributed to a public awareness campaign that helped win passage of a sales tax to support preschool slots and quality enhancements.

Source: "*Making Connections* by the Numbers: Impact and Leverage." Prepared for the Board of Trustees, August 12, 2009.

PLATFORM FOR RESULTS

In 2008, *Making Connections'* leadership (not the evaluation) commissioned a report to the Board of Trustees that examined whether, and to what extent, the ideas and strategies that Casey promoted have changed how local leaders think about, invest in, and "do" community change. The authors, a former mayor and a retired corporate executive, interviewed 14 local funders, mayors, and other city officials in eight locales around the country to learn whether local leaders (a) understand and have internalized the key concepts, (b) have changed norms and practices as a result, and (c) have targeted or achieved different results because of Casey's influence, and (d) how deep and far the changes reach. The authors concluded that:

> [W]hile Casey has not magically and completely transformed poor neighborhoods or their residents' lives, the Foundation's approach to community change makes good sense to local leaders and is influencing them in many positive ways. Casey has helped to establish a crucial set of tools, capacities, networks, relationships, and norms that provide a strong platform for local problem-solving and leadership. In fact, that platform of public capital may be the most durable and valuable result that is directly attributable to Casey's approach. It is a robust and essential force for combating poverty's effects on children and families, and one that prepares local leaders and their partners to meet challenges extending far beyond the scope of any one initiative.... Local leaders identified several topics on which they believe Casey's approach has had a deep, meaningful impact: resident engagement, empowerment, and leadership; community fabric, social ties, and relationships; and family economic success.

SUSTAINABILITY

In 2009, *Making Connections'* management side stated that the transition to Local Management Entities had positioned several sites for potential long-term sustainability. In one site, the LME—a partnership between a settlement house and city government—"is resulting in an ambitious, city-wide agenda to ... sustain *Making Connections'* results through [the mayor's] Pathways to Opportunities initiative." In other site, where the LME is a community development corporation, the entity is "leading a broad coalition of partners...committed to the initiative's two-generation approach." The LME in a third site, a community foundation, is "spearheading development of a network of Centers for Working Families that will connect families to job placement services, asset-building programs, child care, and other work supports." A fourth LME, which includes the city's Department of Community Initiatives, school district, Catholic Charities, Family Services Association, and resident leaders, is "committed to a wide expansion of *Making Connections'* core strategies."

Source: Management Report to the Board of Trustees, October 2009.

Tom Kelly, while acknowledging some missteps, is proud of the measurement tools and techniques that the evaluation created or refined, including the methodology for survey sampling and family tracking, identifying neighborhoods, measuring community engagement ("authentic demand"), and measuring change in core capacities; the data dress rehearsal; and the Local Learning Partnership concept. He also finds the implementation team's site assessment matrix for results management noteworthy.

Frank Farrow identifies two major evaluation-related legacies for *Making Connections*. The first involves insights into the "inseparable interaction" between community capacities and achieving results—that is, between investing in programmatic interventions and supporting the ways in which community and philanthropic leaders, public systems, and neighborhood residents work together to achieve these results. "The fact that a new way of working is continuing in sites—and is proving to be the first response mobilized when sites are challenged with new opportunities, like Promise Neighborhoods—indicates that *Making Connections* made a difference," he says.

Farrow also believes that the Local Management Entities comprise an important part of *Making Connections'* ultimate story. If these entities succeed in the leadership role they have assumed and stay the course in an environment of economic recession and diminished Casey support, that will be a major accomplishment, he notes.

Ralph Smith points to several broad lessons about community change that Casey leaders took from *Making Connections.* The first is prompted by the challenge of attempting population-level change. Smith now argues for an even less monolithic and more disaggregated approach to understanding "population":

> *Even with respect to low-income families in low-income neighborhoods there is the need for more calibrated targeting of interventions and supports. For example, our best example of a workforce funnel was abandoned because not enough residents would meet the minimum requirements for literacy and language. Neighborhood pipelines to major employees proved more challenging and less sufficient when the potential workforce had histories of incarceration and felony convictions.*

Casey's Foundation-wide "repositioning" effort, conducted between 2007 and 2009, embraced this realization explicitly by reframing its approach to "disadvantaged children" to focus on children in five cohorts of families:

• Low-income, low-wealth families experiencing and/or at risk of persistent and intergenerational poverty;

• Low-income, low-wealth families living in neighborhoods and communities of concentrated poverty;

• Low-income, low-wealth families facing additional barriers to opportunity due to family structure, low literacy, limited English proficiency, criminal records, incarceration, dislocation due to community development, etc.;

• Low-income, low-wealth families whose functioning is compromised by chronic health and mental health issues, including addiction and parental depression, and food and housing insecurity and homelessness; and

• Low-income, low-wealth families disrupted by the child's removal by the child welfare or juvenile justice systems or the caretaker's removal by the adult corrections or immigration systems.

While Smith argues for this "more textured understanding" of family populations, he also cautions against getting too attached to any cohort since "the borders between these cohorts are highly permeable. A family in one cohort today could have moved into another by tomorrow."

Smith's second insight is that place-based community change needs an economic engine. "Without an economic engine to produce the jobs and entrepreneurial activities we stand no chance of achieving the core goal of aligning people, place, and opportunities. That's a make-or-break factor," he says.

Third, Smith argues for a more complicated concept of place, which he calls "place-plus":

> The impulse to change conditions and circumstances in places, to cultivate a sense of collective efficacy that is grounded in geography, is still sound. But we have learned that a focus on place by itself is not sufficient to produce and sustain significant changes in outcomes for neighborhood children and families. You also need something that's literally and figuratively more concrete: bricks-and-mortar redevelopment, a school reform effort or charter school creation, housing revitalization—something affirmative that can be an anchor and catalyst for mobilization.

Fourth, high mobility among target populations means that even a place-plus approach has to

develop a set of interventions and strategies that connect people to the groups, institutions, and levers of change that persist even if the resident leaves the neighborhood, such as anchor institutions, social networks, jobs, churches, and the like.

Finally, Smith now believes that population-level change requires a commitment of "patient capital" closer to the scale of Casey's investment in its hometown of Baltimore—about 10 times what the Foundation invested in any one *Making Connections* site. The investment strategy should include a combination of grants, program-related investments, and public dollars that play a "catalytic" role, similar to the approach used by the multi-funder Living Cities initiative, Smith says. It should back up the Foundation's investments with contributions by major local investors as a way to "grow the pool" of resources. And it will take time.

These takeaway thoughts are, surely, just the starting point for a stream of lessons from *Making Connections.* In the end, it may be most instructive to consider *Making Connections'* evaluation in terms of the initiative's original guiding principles. In grappling with the challenge of measuring change while changing measures, *Making Connections'* evaluators broke new ground. At times, they clashed over data collection and management issues with the initiative's management and implementation staffs, but the evaluation did no apparent harm. And although they made some new mistakes along the way, the evaluators also added considerable value to the growing knowledge base and arsenal of tools for evaluating community change.

Endnotes

1 This case study is not about the initiative itself, at least not beyond its intersection with the evaluation. Other documents available from The Annie E. Casey Foundation and the Center for the Study of Social Policy describe the initiative in depth.

2 Several important aspects of *Making Connections'* evaluation could not fit within the scope of this case study, which focuses on what changes to measure and how to measure them. Equally important aspects of the initiative and its evaluation—including the challenges of building capacity for evaluation and learning, using data to revise implementation strategies, staffing and managing both evaluation and implementation needs, and integrating evaluation with implementation—are touched on here but warrant much deeper discussion elsewhere. Similarly, because this case focuses on measurement issues, it does not provide a complete description or analysis of the infrastructures for data collection and analysis that *Making Connections* invested in, locally and across sites. These include Local Learning Partnerships, local data warehouses, and a cross-site indicators database.

3 See Kubisch, A.C., Auspos, P., Brown, P., and Dewar, T. (2010). *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts.* New York: Aspen Institute Roundtable on Community Change, p. 23.

4 http://philanthropy.com/premium/stats/foundation/main.php?program=All&year=&state=&keyword=&year=1999

5 KIDS COUNT, Casey's signature project, supports a network of grantees in the 50 states, District of Columbia, Puerto Rico, and U.S. Virgin Islands who collect and analyze data on a common set of indicators of children's educational, social, economic, and physical well-being. Casey publishes an annual national report based on these data and also funds state-level projects that offer a county-level assessment of the condition of children.

6 Fiester, L. (Forthcoming). *Place, Plus: How the Annie E. Casey Foundation and United Way of Metropolitan Atlanta Influenced Each Other's Perspective on Community Change.* Baltimore: Annie E. Casey Foundation, pp. 18–20.

7 Smith, R. and Fiester, L. (2007). "*Making Connections* at Midpoint: Early Choices, Lessons, and Results." AECF: internal document.

8 Doug Nelson, speaking at the Jobs Initiative Showcase Conference, September 6, 2001.

9 Smith, R. and Fiester, L., op. cit.

10 *Ibid.*

11 See for example: *Voices from the Field, Vol. I* (Aspen Institute's Roundtable on Comprehensive Community Initiatives); *Community Building: Coming of Age* (Development Training Institute and the Urban Institute); *Stories of Renewal: Community Building and the Future of Urban American* (Rockefeller Foundation); *The Path of Most Resistance and The Eye of the Storm: Ten Years on the Front Lines of New Futures* (Annie E. Casey Foundation); and analyses produced by the Chapin Hall Center for Children and Families at the University of Chicago and by the Urban Strategies Council.

12 Fay, W. (2004). "External Reconnaissance." An issue of *Casey Dispatch: Opportunities, Challenges, and Lessons from Making Connections.* Baltimore: Annie E. Casey Foundation, p. 3.

13 Kubisch, et al. op. cit., p. 9.

14 Adapted from Bernholz, L. (2008). "Geologic Timeline of Philanthropy Industry (USA).

15 See Fiester, L. (2011). "Plain Talk About Teen Sex: Evaluating a Community-Based Effort to Change Adolescent Behavior Through Adult Communication."

16 Ferguson, R.F., and Dickens, W.T. (1999). *Urban Problems and Community Development.* Washington, DC: Brookings Institution Press, p. 607.

17 Patton, M.Q. (2009). "Developmental Evaluation." PowerPoint presentation delivered to CES, slide 41.

18 Fay, W. (2004). "The Design Development Process." An issue of *Casey Dispatch: Opportunities, Challenges, and Lessons from Making Connections.* Working document. Baltimore: Annie E. Casey Foundation, p. 2.

19 "Expectations and Requirements for Local Learning Partnerships." (June 6, 2003). Memo distributed to LLP members. Baltimore: The Annie E. Casey Foundation, p. 1.

20 Blank, S. and Fiester, L. (2004). "The Site Sorting Process." From the series *Casey Dispatch: Opportunities, Challenges, and Lessons from Making Connections,* p. 6.

21 *Ibid.,* p. 5.

22 Fiester, L. (2011). op cit., p. 20.

23 "*Making Connections* Core Outcome Survey: Background Information." (January 31, 2002. Memo from AECF Evaluation, Data Analysis, and Policy Advocacy unit to representatives of local learning partnerships, p. 2.

24 "Invitation to Apply for *Making Connections* Cross-site Survey Public Research Data File," (Undated), p. 2.

25 *Ibid.*

26 These questions adapted from "Request for Proposals: Surveys in *Making Connections* Sites," pp. 3-4.

27 Guy, C. (May 13, 2002). Memo to evaluation team on May 9 survey management meeting, p. 2.

28 Comments by Cindy Guy on a draft summary of evaluation design components by the Association for the Study and Development of Community, 2008.

29 Guy, C. (2009). "Documenting the obvious: The role of survey research in community change initiatives." Essay for the *Making Connections* Writers' Roundtable, p. 3.

30 *Ibid.*

31 "Invitation to Apply for *Making Connections* Cross-site Survey Public Research Data File," (Undated), p. 1.

32 Guy, C. (2009). op. cit.

33 *Ibid.*

34 Fiester, L. (2011). p. 20.

35 Kelly, T. (2004). "Finding (and Redefining) Nemo: Or, What Seafood Can Tell Us about Sustainability," AECF: Internal document, p. 2.

36 *Ibid.*

37 Fiester, L. (2009). "Staff as Inadvertent Intermediaries." Essay developed for the *Making Connections* Writers' Roundtable, p. 2.

38 July 19, 2005. Unpublished minutes of a *Making Connections* internal management meeting.

39 The text on core capacities assessment that appears on pp. 25–6 borrows heavily from a chronology contained in a memo by Mary Achatz, "Assessing Community Capacity to Develop and Implement Powerful Strategies," dated September 2008.

40 Achatz, M., and Ruiz, M. (January 26, 2005). "Assessment of Local Capacity for Community Change in Cohort 1 *Making Connections* Sites." Internal memorandum to Tom Kelly, p. 23.

41 Farrow, F. (2010). Response essay contained in V*oices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts.* New York: Aspen Institute Roundtable on Community Change.

42 *Ibid.*

43 "Focus Schools Data Tracking." (October 23, 2009). Memorandum from Tom Kelly to Metis Associates.

44 Appendix C: Assessing Capacity to Implement Powerful Workforce Strategies: Questions to Guide Local Discussions. (Undated). Casey cross-site evaluation team product, p. 4.

45 Described in "The Data & Evaluation Dress Rehearsal Exercise: Purpose, Process and Products." (April 2008). Guidance to sites from the *Making Connections* evaluation team.

46 *Ibid.,* p. 7.

The Annie E. Casey Foundation