WHEN THE RESULTS DISAPPOINT, THE PROCESS MATTERS LITTLE

William Meezan, D.S.W., Ohio State University and Bowen McBeath, Ph.D., Portland State University

This narrative describes a research project done in the community that did things “right” by academic standards and at times felt “blessed by the research gods.” It brought together members of the academy and representatives of public and nonprofit agencies, was fully participatory in nature, had adequate funding that was obtained with relative ease, used the strongest of designs and analytic techniques, and was well disseminated. Yet, in the end, politics and behaviors that fostered organizational maintenance trumped findings that led to new knowledge and organizational learning. Despite the academic rewards that the author has reaped and will continue to reap from this experience, it left him feeling more cynical about the system of services that has been put into place for children in foster care and their families.

I arrived at the University of Michigan in fall of 1999 to assume my new position as the Marion Elizabeth Blue Professor of Children and Families. Having left Los Angeles with most of my commitments completed, I was anxious to find my next big project. I knew it would not be easy to establish myself in a new school, with a new set of colleagues. Nor would it be easy for me to break into and learn a new child welfare system. But having to do so with what felt like so many people “watching” — after all, I was in the first endowed chair at what was rated as the top school of social work in the country — was more stressful than I had ever imagined. The self-imposed pressure was intense, to say the least.

This pressure pushed me to be bolder and more entrepreneurial than I normally would have been. I interviewed colleagues at the school, asked people to make introductions for me, and called state officials “cold” to try to get a beginning understanding of the child welfare environment in Michigan and what was “on people’s minds.” I was pleasantly surprised at how open people were to talking with me. Perhaps it was my position; perhaps it was the fact that I had connected with a well-known child advocate in the state who was willing to open doors for me. Whatever the reason, getting in to see people didn’t prove to be as difficult as I expected.

Interestingly, at the state level and to some extent locally, I discovered that one experimental program under way in Wayne County (Detroit) seemed to be getting all of the buzz. Called the Wayne County Foster Care Permanency Pilot Initiative (the pilot), it was a new way of reimbursing some of the private foster care agencies for providing services to children and their families under their care. The pilot had two primary objectives: to reduce the number of days that children spend in foster care and to increase the percentage of foster children that reach a permanent placement. A secondary objective of the pilot was to allow nonprofit agencies to provide services to foster children and families in a flexible manner, thereby (hopefully) better meeting their needs. Who could argue with either of these objectives? I certainly agreed with them.

In order to provide incentives for agencies to meet these objectives, the state child welfare agency — the Michigan Family Independence Agency (FIA) — altered the provisions for agency reimbursement in its foster care contract. Rather than being reimbursed for administrative costs on a standard per-child, per-diem basis, the
agencies that volunteered to participate in the pilot provided services on a reduced per-diem basis. These pilot agencies’ revenues, however, were supplemented by a substantial initial payment as well as additional bonus payments linked to the achievement of performance milestones. Pilot agencies were allowed to use their initial, bonus, and per-diem payments as they saw fit.

This approach intrigued me; in an era when managed care was infiltrating the child welfare sector and wreaking havoc with a number of state systems, this seemed to be a sensible approach. I wondered if it would work. I wondered how agencies adapted to the new model. And I wondered why there was a significant buy-in into the program on the part of the state and some private agencies.

The pilot was planned in 1995-96, when the Republican party controlled state government in Michigan. My poking around led me to a couple of tentative conclusions about why Michigan had decided to test this particular approach. On first blush it seemed that the decision to experiment with the pilot came about due to a confluence of factors. The first of these had to do with Michigan’s approach to child welfare: Michigan law had predated the federal Adoption and Safe Family Act in terms of having stricter time limits to move children through the foster care system, and the pilot was seen as a way to facilitate this movement.

The second factor emanated from some of the private agencies themselves—a handful of these agencies were eager to have flexible funds, particularly when a child first entered care, to meet the immediate needs of families and thus facilitate shorter stays in care. These agencies also saw this reimbursement system as a way to refocus their mission and be rewarded for taking timely actions on the part of children. Some, I vaguely suspected, may also have seen this reimbursement system as a way to make their foster care service more profitable.

The third factor was a reflection of Republican thinking at the federal level about what child welfare delivery systems should look like. “Performance-based” and “accountability-driven” service delivery were becoming important buzzwords in Washington. In addition, the pilot was seen as a step toward controlling the cost of foster care, as it included two fiscal features characteristic of managed care models popular with fiscal conservatives—cost containment and risk shifting (Embry, Buddenhagen, & Bolles, 2000; Wernet, 1999; Wulczyn, 2000).

The pilot was launched in 1997 with 4 of the 19 private, nonprofit foster care agencies in Wayne County choosing to adopt this new way of reimbursement for services (two of these agencies were instrumental in designing it). In March 2000, when two additional agencies adopted this incentive-based reimbursement system and some modifications had been made to it, I met with the officials from FIA responsible for it. Present at this meeting was Jim Beougher, Director of Children’s Services for the state, who was a major force behind the initiative.

It was clear to me that those around the table, and particularly Jim, wanted the pilot evaluated. They seemed sure that it “worked”—that it moved children through the foster care system more efficiently without harming them. They guaranteed cooperation, but no money at that time, and asked that I pursue the possibility with the pilot agencies if I was interested in undertaking an evaluation of the initiative. I was excited about the possibility. Here was an important project that could enhance knowledge about an innovation in child welfare. Here was the possibility of forging important relationships with the state and with agencies that had never existed at the School of Social Work. Here was a way to again test my research skills with a “big” project. Here was a possibility of my “hitting
the ground running” and proving myself to my new colleagues.

It seemed like a natural for me. Child welfare had been the intellectual home for my entire career, and I had grown to be a community-based, participatory evaluator, who, while building on the work of others (Patton, 1997), had even developed and applied a framework for organizational learning through evaluation (Cherin & Meezan, 1998). My previous work had taught me that under the right conditions, university researchers and agency personnel could work closely to design and implement high-quality evaluations that had real, practical meaning for organizational and policy change (McCroskey & Meezan, 1997; Meezan & O'Keefe, 1998a, b).

The next step was clear—I needed access to the pilot agencies to ascertain their interest in such an endeavor. I had no idea if they would be interested. I doubted whether such a project could be pulled off. Would they be protective of their practices as many agencies often are? Would they open themselves up to scrutiny, particularly when they knew that FIA supported the pilot initiative? Nonetheless, in my new emboldened state, I decided to put on my “confident” face and move forward.

I decided that the easiest way to begin to try to gain entry to the private agencies was through their association—the Michigan Federation for Children and Families. Its executive director, Bill Long, was extraordinarily welcoming and offered to make introductions. He also let it be known that he would be willing to support an evaluation in any way he could. While he could not offer money, he offered to make introductions to the executives of his member agencies, to help get me into the right meetings, to throw the weight of the Federation behind the project, and, should we get this project off the ground, to designate a staff member to be involved in its design and execution.

In this first meeting with Bill, something else occurred that would shape my time in Michigan. He mentioned that he was in discussion around research ideas with a very bright doctoral student in social work and political science at the University of Michigan who was interested in nonprofit management. His name was Bowen McBeath—“had I met him?” Admitting that I had not, I told Bill that I would contact Bowen to see if there was a “fit” between our interests and in our styles.

Bowen and I met shortly after this meeting, and we hit it off immediately. Here was an intellectually curious young man who had previously worked for researchers I respected. He was looking for a long-term opportunity to learn to do meaningful research in the “real world.” He was personable and seemingly hard working. I decided immediately (I often form judgments of people quickly) that I couldn’t have asked for a better “second”; he seemed (and proved to be) trustworthy, organized, and committed to his own learning. Perhaps equally importantly, his skills complemented mine—he knew the literature on nonprofits and organizational theory while I knew foster care and child welfare research.

There could not have been a better match; finding Bowen just felt “right.” I doubt if he knew how scared I was of this large and tenuous undertaking; I know he was unaware of the doubts I had that the project could be “pulled off.” And I’m sure he didn’t know how much his presence gave me strength, confidence, and fortitude to go forward. Connecting with Bowen in this serendipitous way calmed me and told me that things might just work out—that the research gods might be shining on me and this work.

When we decided to begin working together, I couldn’t promise Bowen anything more than the opportunity to go “exploring” with me to see if we could put this project
When the Results Disappoint, The Process Matters Little

together — politically, economically, and logistically. But there was a commitment attached. If he took the risk and joined me on this exploration, and if, against the odds, we pulled the project off, I would guarantee him everything a doctoral student could want beyond the classroom: experience in negotiating a community-based research project; exposure to grant writing; the opportunity to directly supervise a complex research project on a day-to-day basis; more-than-fair compensation; data for his dissertation; opportunities to present at conferences; and joint authorship on all publications. I was thrilled when he said “yes,” and for the next few months we laid the foundation for what has become, at this point, a five-year journey. He wasn’t compensated for his time initially, but he was learning from me and from the people we met. It was evident in our long discussions to and from meetings, and from our grandiose planning, that we were excited and were going to have fun!

Bowen and I began to meet with the State FIA officials charged with the administration of the pilot and with key administrators and foster care program managers from the pilot agencies — they had a regularly scheduled meeting to discuss implementation issues. In May 2000, by invitation, we attended our second meeting with this group, and the possibility of doing an evaluation was the major agenda item of the day. At this meeting I explained the evaluation approach I was proposing. I assured them that they would have the major “say” in the evaluation — that after I taught them some basic research and evaluation principals they would basically design and execute the project under my direction. I used every cliché in the book that I believed: that the evaluation would be “bottom up” rather than “top down”; that the evaluation would be done “with” them rather than “on them”; that the evaluation was not about “I gotcha,” nor was it a “compliance audit.” Clearly they were skeptical, since some had been burned in the past by researchers and outside evaluators. But they did listen, and even made comments like “Are you sure you’re a university researcher? So many don’t work this way,” and “You seem to still care about the kids.”

We agreed at that meeting to go exploring together but also thought that the group at the table (probably 40 in number) was too large to plan this effort. We thought it best that those in charge of supervising the foster care program and some other key players in the agencies should form an evaluation committee. And so, through a series of meetings with this evaluation committee between May and December 2000, we came to a common understanding of the role of evaluation in improving program performance, the pilot-related research questions that we might examine, the data that might be necessary for the evaluation, and how these data might be collected.

In order to form a control group, Bill Long helped us recruit three Wayne County foster care agencies that were operating under the normal fee-for-service contracts (the non-pilot agencies). He suggested potential agencies based on their similarity to the pilot agencies in terms of size and other criteria, including his belief that they would participate. With Bill having greased the wheel, I met with appropriate officials at these agencies and they agreed to lend their support to the research; some really didn’t believe that the pilot was good for children or their families and wanted to “prove it” (which is always a great motivator for participation). These executives appointed their representatives to the evaluation committee who were, for the most part, the foster care service directors/managers from their agencies.

From the beginning, the evaluation committee consisted of me, Bowen, representatives from Michigan FIA as well as the Michigan Federation for Children and
Families, and one to two representatives from each of the six pilot and three non-pilot agencies. In addition, there was a time when Wayne County (as opposed to the State) FIA was represented at the table, but they eventually dropped out of the process. Their on-again, off-again involvement with this project could be the subject of its own narrative, but suffice it to say that I was furious, as I believe others were, when they walked away from the table just before data collection began. They had influenced the process and the evaluation questions and methodology, yet when push came to shove they didn’t want to be part of the study even though they provided 15% of the foster care services in the county. There was a series of written and in-person exchanges between me and the Director of the Wayne County FIA, but in the end she would not budge. So we went ahead with a simple comparison between Wayne County nonprofit agencies that were and were not participating in the pilot. And I was left to conclude, once again (I had experienced such reluctance in the past), that too many public sector child welfare agencies don’t want to open themselves up to the same kind of scrutiny that they expect of those from whom they purchase services.

Other than the Wayne County FIA folks disappearing from the table, there was little turnover among committee members; all but a handful of the agency representatives were involved from start to finish. During its first three years, the committee met bimonthly first to learn about the evaluation process and then to design and carry it out. Between 2003 and 2004, the group met monthly to review new analyses, comment on their meanings, and discuss methods of disseminating study results to state and local policymakers and agencies serving children and families. Everything worked as it was supposed to, the research gods were smiling on us, and we were delighted.

### The Participatory Nature of the Evaluation

From its inception, the evaluation committee operated in a participatory manner (Cherin & Meezan, 1998; Patton, 1997). We followed a modification of an evaluation cycle that I had helped to develop earlier in my career (see Figure 1) and accordingly archived each of our meetings.

![Figure 1 - The Action Evaluation Model](image)


All major aspects of the evaluation—central research questions, research design, sampling, and data collection techniques and instruments—were discussed with, decided by, and approved by the cooperating agencies. The following notes from an early meeting demonstrate the tone set for all dialogue around these issues:

*Bill handed out [a] task list and asked all present to read it and see if there were any omissions or incorrect order of task items, as this list would ...become the evaluation committee’s working agenda. Verlie asked if there were any more*
specificity to the notion of “choosing a sample.” Bill responded that there were two questions we need to answer here: a) Who is eligible for the sample...and b) when do we start collecting the sample? ... Verlie also asked about any preliminary ideas concerning the timeline for each task. Bill replied that each task will take as long as we need to take...Rita asked if Bill and Bowen had anything in mind regarding the need to create data instruments. Bill responded that we will bring drafts of questions and ideas for the group to consider, but we’ll rely on the group to be expert judges in order to ensure that the questions have content validity. In other words, the overall group will make sure that the questions [are] able to capture the reality...(11/28/00).

Why did the agencies participate in what would turn out to be an overwhelmingly complex set of studies? I believed at the time that they participated, despite the considerable costs in attention, effort, and time, because they were genuinely interested in learning whether the pilot initiative was effective in reaching its stated objectives. I truly thought (and I believe that they did too) that they had hoped to use the study results to improve the performance of their foster care departments and that they wanted to learn how to carry out rigorous evaluations so that they could be carried out in other parts of their agencies.

As a result of this voluntary and participatory framework, Bowen and I and our research staff enjoyed a close working relationship with the representatives from the participating agencies. This relationship allowed for multiple rounds of data collection, data cleaning, and data analysis, which are often difficult to successfully complete in large-scale, field-based, research projects. This close relationship also improved the accuracy of the databases, the relevance of the analyses undertaken, and the validity of the interpretation of the findings and their congruence with agency experiences.

Funding the Evaluation
It was at the point when we had to find funding for the project that I first used the word “blessed” out loud. I had found Bowen, the partners were on board, the planning was going well, but without money nothing further could happen. Then things really started to fall into place. FIA was shocked that we had gotten as far as we had and were actually “pulling it off.” Jim Beougher said that he would give us a small, non-competitive grant for the first year, and a second grant the next year if we continued to succeed. With the first $25,000 in hand, we went looking elsewhere. Each time we looked for money, we somehow found it. Nobody would fund the whole project in any one year—no single funder really believed that we could pull off a project with a complex design, multiple data points, nine agencies and the State of Michigan, in Wayne County, etc. But the more we succeeded the easier it was to piece together money from different sources.

Even our proposals to funding sources were reviewed word by word during the evaluation meetings; yes, it really was participatory:

> Bill will incorporate the group's revisions to the output evaluation proposal prior to sending it on to Lynn at FIA; Bill will send everyone a draft of the... proposal for an outcome evaluation planning grant (11/01/00).

And so one challenge grant funded another, and in the end we funded ourselves...
When the Results Disappoint, The Process Matters Little

for four years through more sources than I could have ever imagined: FIA, the Aspen Institute’s Michigan Nonprofit Research Program, various offices at the University of Michigan, the Kellogg Foundation, and the Skillman Foundation all contributed at some point. Some grants were bigger than others, and some years were leaner than others, but there was always enough money for the basics. In flush years we could pay the agencies for some of the time it took them to collect data, but when we couldn’t pay they collected the data anyway – and they understood.

The Central Research Questions of the Evaluation

We decided that the evaluation of the pilot initiative should include a process evaluation and an impact evaluation. The process evaluation contained a qualitative and a quantitative component; the impact evaluation was strictly quantitative.

The qualitative component of the process evaluation examined nonprofit service providers’ organizational adaptations to the pilot initiative. Two research questions guided this phase of the study: 1.) What organizational changes did pilot agencies make when they shifted to the pilot contracting environment? and 2.) What organizational accommodations did non-pilot agencies expect to make as they anticipated shifting to the pilot contracting environment? The study examined these questions with data obtained through telephone interviews with administrators, foster care supervisors, and foster care line staff across the six pilot and the three non-pilot agencies in the evaluation.

The quantitative component of the process evaluation examined the effect of the pilot initiative on service provision to foster children and their families and became Bowen’s dissertation (McBeath, 2006). This phase of the study was organized around three research questions: 1.) What types of services did foster children and families receive from pilot and non-pilot agencies? 2.) Were there any pilot/non-pilot differences in the amounts of services provided over time? and 3) What factors were associated with differential service patterns between the pilot and non-pilot agencies?

The impact evaluation sought to identify how and at what rate pilot and non-pilot children moved through the foster care system. This study focused on six research questions: 1.) How many foster children reached each performance milestone? 2.) Were there any pilot/non-pilot differences in the proportion of foster children reaching each performance milestone? 3.) What factors were associated with the achievement of each performance milestone? 4.) Were there pilot/non-pilot differences in the final dispositions of foster children and families? 5.) Were there pilot/non-pilot differences in the length of time needed to reach these final dispositions? and 6.) What factors were associated with the achievement of these final dispositions?

The Qualitative Process Study

We designed this stage of the study to assess changes across various levels of agency employees and with employees who performed different functions within the agency. We gathered information on the effects of the pilot foster care initiative on service delivery, interdepartmental and interorganizational relations, staffing patterns and staff training, staff roles and responsibilities, and financial management and technology use. We interviewed 45 administrators, 19 foster care supervisors, and 20 foster care line staff from pilot and non-pilot agencies based on guides developed by Bowen and me and modified over the course of a number of meetings by the evaluation committee. Members of the evaluation committee purposively suggested respondents for their knowledge of foster care and the pilot initiative. Those who were interviewed from
pilot agencies told us about their actual experiences in the areas that were probed; those from non-pilot agencies speculated on what would happen in these areas if their agency had to move to incentive-based contracting (for a full discussion of this phase’s methodology, see Meezan & McBeath, 2003).

In the next few days, agencies should look through the draft questionnaire with an eye to (1) any important topics or questions that are not present and (2) any topics or questions that seem superfluous. Agency administrators, supervisors and line staff should be given a chance to comment on the questionnaires. Agencies should provide Bill with a list of staff (at the administrative, supervisory, and line staff levels) that should be interviewed. These should be experienced, thoughtful people (6/21/01).

The results of the study delighted us, for they systematically conveyed the experiences of the pilot agencies and showed how naïve the non-pilot agencies were about the depth of change that would be necessary should they have to move toward this form of reimbursement (for a full discussion of these results, see McBeath & Meezan, in press; Meezan & McBeath, 2003, 2004). And the findings delighted our agency partners; those at the table thought that we had really captured their realities.

Our findings described the service delivery changes made in intake and assessment procedures, performance tracking, and managing flexible funds made by the agencies. They suggested that performance-based contracting required the service providers we studied to quickly diagnose clients’ needs, focus additional resources on collecting client and service information in databases, and pool funds to address the multiple needs of clients systematically and simultaneously rather than haphazardly and sequentially. And we captured the fact that many agencies felt the need to create new type of positions within their foster care departments.

Our results also showed the depth of change needed in the agency’s financial management and accounting procedures, and the difficulties in budgeting and forecasting for the pilot agencies due to insufficient information, hidden costs, and other unexpected barriers. We found that moving to the pilot model required foster care departments to have more communication with other departments in the agency.

Our findings also uncovered some discomfort with the pilot, and I only wish we had paid more attention to this finding at the time. Agency supervisors in particular noted that some conflict existed between the pilot program’s goals and their conceptions of the appropriate goals and processes of foster care. For some of them, performance-based contracting was associated with the possibility of reduced lengths of stay for children in foster care and less service to children and their families. Some workers and supervisors feared that agencies might return children to biological families at the ultimate expense of child safety and well-being; others feared that parental rights might be terminated before parents had been given an adequate chance at rehabilitation through the provision of appropriate, long-term services.

This process evaluation allowed us to make recommendations that were developed with the help of the evaluation committee. We agreed that contract incentives, as seen in performance deadlines, directly influenced work environments and employee behavior, and that if performance deadlines were set they needed to be done so that employees had sufficient time to conduct thorough
assessments, provide necessary services, access community resources, and place children in living situations that promote safety, permanency, and positive well-being. We agreed that the data suggested that service delivery should be monitored to ensure that client needs rather than cost considerations dictated the services provided to a family, and that if agencies were to be successful under this form of reimbursement, formal relationships with collateral service providers needed to be established and high-demand, out-of-agency services needed to be available. It was also apparent to the evaluation committee that management had to listen to front-line staff to ensure that performance-based contracting did not jeopardize the quality of service provided to the client, and that both vertical and horizontal communication within the agency was crucial for successful implementation.

We were off to a great start! Our findings and recommendations rang true with the evaluation committee, the agency executives to whom they reported, and the professional audiences to whom they were presented. We were capturing what the agencies were living, and we were providing useful information to those who might move to this form of reimbursement system. Who could ask for more?

The Quantitative Process and Outcome Study

We needed to identify a sample of foster children being served by the six pilot agencies and three non-pilot agencies in order to examine service provision to foster children and families and the outcomes of these services. Since the early 1990s, children entering foster care in Wayne County were assigned to nonprofit agencies through the Family Assignment System (FAS). The FAS requires children entering foster care for the first time to be assigned to the next agency in the alphabetical queue of nonprofit service providers. If an agency was at full capacity and did not have a foster care slot available, the next agency in the queue was contacted, and if a space was available the child was placed with that agency.

The 19 nonprofit foster care providers in Wayne County received all of their foster care cases through the FAS. As a result, children were assigned to agencies in a manner unrelated to their characteristics or family circumstances. Thus, assignment to agencies was random and placement with an agency was decided based only on when the child entered care. Who could believe it? Without planning, conditions in Wayne County were such that we had the equivalent of a randomized trial! We were going to be able to isolate the independent effects of the pilot initiative from other individual-level covariates of service provision. Who could ask for more? As researchers we were jubilant. To the agencies, the rigor of the study was a source of pride – nobody could doubt that their work was being evaluated in the most exacting, fair, and impartial manner. The following, from notes after the first data were in, shows the agencies’ general pleasure with what was being accomplished because of the use of this design:

Bill first led the group through the comparison between pilot and non-pilot children. There are minimal differences between the groups. Therefore, Bill mentioned, any differences in outputs are not due to differences in the children...Rita said, “There goes our first excuse” to general laughter. (1/10/02)

Clearly the research gods continued to bless us.

We began sample identification in May 2001. The evaluation committee agreed to gather a sample containing roughly 250 foster
children – 180 foster children from the six pilot agencies and 70 children from the three non-pilot agencies. This sample size reflected the agencies’ estimates of the number of new foster care case admissions over a two-month period. This sample size also provided sufficient cases to ensure statistical power for some multivariate analyses.

Agencies were responsible for identifying all study-eligible children and for completing a sample screening form for each foster child entering the agency during sample selection. As a group, we agreed on the criteria on which children would be admitted to the study. And as became the routine for the development of all data collection instruments used in the study, Bowen drafted the screening form used by agencies to identify study-eligible children, I modified/approved it, and the members of the evaluation committee modified it to better meet their needs and fit into the data collection that was already occurring at their agencies.

**Homework for each agency:**
- Add/subtract to question 11 of the service provider data collection instrument...
- Review the potential aggregate service data collection instrument... (3/21/01)

**Homework:**
- Bill will be e-mailing final versions of all forms to the agencies. Each agency should run one case through the forms to see if they work and let Bill know if they have problems, within a week (4/5/01)

Bill said that he was very pleased that he’d received requests from a number of people asking for revisions to certain [data collection] forms. He went over a few of the requests... (4/20/01)

In all, agencies used six data collection instruments to gather quantitative data for the evaluation. These six data collection forms concerned the characteristics of the foster child at the point of his or her entry into foster care; the characteristics of the primary caregivers of the foster child at the point of his or her entry into foster care; the services provided to the foster child and his or her family by the foster care agency; the services provided to the foster child and his or her family by other agencies; the characteristics of the caseworkers serving the foster child and his or her family; and the placement status of the foster child. The evaluation committee designed five of these six instruments; only the form pertaining to out-of-agency services was not constructed specifically for the evaluation, since FIA requires foster care providers in Wayne County to list the services provided by other agencies to foster children and their families on a form called the “FIA 67.”

Two of these instruments were completed only once per child. These non-recurring instruments included the child characteristics form and the primary caregiver characteristics form. In contrast, the in-agency services, worker characteristics, and child output data collection instruments were completed 30 days after the child’s entry into foster care and then every 90 days thereafter. This was done at the request of the evaluation committee, since foster care caseworkers in Michigan are required by State law to complete regular reports concerning the status of foster care cases at these time intervals. These recurring data collection instruments were collected regularly through 930 possible days in care, or just over 2 and one-half years.

We began sample selection for the six pilot and the three non-pilot agencies on May 1, 2001. Sample selection stopped for pilot
agencies on September 1, 2001, and one month later for the non-pilot agencies due to the need to admit adequate numbers of non-pilot children and ensure statistical power for multivariate analyses. The final sample was composed of 243 foster children, which included 175 children from the six pilot agencies and 68 children from the three non-pilot agencies.

**Politics Creep In**

In the State of Michigan 2002 was an election year, and there was a chance that the Democrats would regain the state house. With this concern, the strong commitment to the fiscal principles that underlay the pilot initiative, and the belief that the pilot was "working," Jim Beougher and his staff at FIA publicly announced that the pilot would go county-wide (all 19 private foster care agencies in Wayne County) in October of 2002 before any possible change of administration could occur. This happened despite the fact that the State was partially funding the evaluation.

I feared that politics was going to trump knowledge and rational decision making, for we were not going to have findings by October. Since I knew from previous encounters that as a researcher I would not be able to influence Jim directly, I worked behind the scenes with the FIA staff on the evaluation committee and their superiors to try to influence his thinking and get him to reverse his decision. I fed them arguments and reasoning to use with him, such as "If the data were not going to be in, and if the pilot needed to expand for political reasons, why not do another phased entry as had happened in 2000?" "Wouldn't it be embarrassing to go county-wide only to discover that the experiment had failed?" My voice resonated with the FIA folks who had been involved with the evaluation, but their voices did not resonate with Jim. The plans to go county-wide were to go forward, and that was that.

Our disappointment was lessened somewhat by the fact that the preliminary data from the qualitative process study would help ensure that new agencies moved to this reimbursement system in an informed and thoughtful way.

"Lynn [the FIA rep to the evaluation committee] said that she is giving non-pilot agencies a walk through the pilot guidelines and that Cheryl and Carmine [members of the evaluation committee] have helped a lot. Lynn said that she read some of the qualitative themes at the meetings, so non-pilot agencies are benefiting from our work here...Lynn's goal is to use more of our group's data in order to have new trainings later in the fall. Our interview data is helpful...because it helps her keep in mind what's important to talk with agency representatives about." (8/14/02)

**The First Quantitative Findings**

As findings of the study became available, we began to present them to the members of the evaluation committee. Initially, we were most interested in the findings after 300 days of the experiment, since the first incentive payment milestone occurred at 290 days, and we wanted to see whether the incentive-based payment system was having any effect on service provision to and placement outcomes for children and families.

The findings were, for the most part, disappointing to those at the pilot agencies. What we found was that while controlling for many other variables, children and families served by pilot agencies received significantly fewer services from the agencies themselves, and that fewer referrals were completed for out-of-agency services on their behalf than non-pilot children and families. These results fit perfectly with the concerns of the workers.
and supervisors we had interviewed in the qualitative study.

For months, people struggled to make sense of these findings, and the struggles intensified over time. Agency representatives often interpreted the overall findings as suggesting that their agency was primarily responsible for the changes that were occurring. Even representatives from non-pilot agencies tried to find reasons for the differences to bolster the work that their friends/colleagues in the pilot agencies were doing:

Rita then said she was concerned about documenting for services — her workers may not be documenting all that they do...Rita said she’d be disturbed if an interpretation was made that pilot kids were going home faster with fewer services (10/3/02).

Rita suggested that one of the possible reasons why pilot services are lower than non-pilot services is because workers may not be recording the services provided. Bob said that he thought the differences in services could be related to turnover. Kathy replied that this could be true across agencies. Kathy also mentioned that it’s possible that pilot agencies have simply found a more efficient means of doing business ... Rita noted that her relative assessments are done outside of the foster care unit and are therefore not recorded. Rita, Cheryl and Kathy said their workers’ case aide contacts probably didn’t get recorded either. Rita also mentioned that these results made her want to revise how her workers fill out forms. Rita said she was fighting the urge to go back to her cases and change them. She said, “Up ‘til this point I’ve taken a laissez faire approach, and it’s killing me now, and I need to hear you tell me not to do it.” Bill replied “Don’t do it. Trust the process.”(11/21/02)

And we debated some potential rival hypotheses for these findings. Were pilot agencies providing only those services that are seen as absolutely necessary to achieve performance payments and/or improve client outcomes? Did pilot agencies choose to save the initial lump-sum payments and performance bonuses rather than use them for additional services? Were unmeasured differences in organizational goals, structure, processes, and technology impacting service provision? Did this form of payment simply lead to service rationing?

[After a discussion of possible creaming in terms of service provision, which the data suggested might be occurring at that point] Bill: “If I made a statement to you that a worker with too many cases spends time first with easier cases, would that surprise you?” Rita: “But that hides the honest effort of workers to spend time with each case. So subconsciously there is this...” Bill: “Remember, we’re not judging you. This is for improvement. If these things hold, it’ll suggest who you’re serving, who you’re missing, and your programs. ...what lessons can we learn from these data?” Kathy: “That’s what Jim did at LAS. Cases most likely to succeed/fail received different services. And that was because he was bothered by what he knew about who we were not serving.” (12/2/02)
We obviously could not resolve these questions, but those around the table agreed that whatever was going on was, at least at the operational level of the foster care program, unintentional. In the end, everyone around the table accepted the fact that performance-based contracting was associated with a decrease in service provision. And much to people’s dismay, and despite the strong emphasis in professional child welfare training on crafting individualized treatment plans, multivariate analysis led us to conclude that few case conditions influenced the amount of services that were provided to children and family in the foster care system.

The outcome data proved just as disappointing. After 300 days in care, there were, in general, no statistically significant differences in the placement status of children from pilot and non-pilot agencies on the major outcome variables: 36% of the full sample was returned to a relative and 21% was returned home to a parent. And while 55% of the children served by pilot agencies and 45% of children served by non-pilot agencies reached this first pilot payment milestone, this difference was not statistically significant. Furthermore, this milestone was achieved in about the same amount of time under these two conditions. Yes, there were some minor differences that did favor the pilot agencies: they were significantly more likely than non-pilot agencies to have their agency supervision terminated (15% as opposed to 3%); and children from pilot agencies were also more likely to have their court supervision terminated than were children from non-pilot agencies (19% versus 4%), but these did not overshadow the main finding that the two payment systems seemed to be achieving similar outcomes in about the same amount of time.

We did learn some useful information about who did and who did not achieve permanence within this 290-day framework, and that knowledge provided a feeling of accomplishment for those around the table. It was important to know that children that had experienced prenatal drug and/or alcohol exposure were less likely to reach the first performance point; that children that had been neglected or abandoned were also less likely to achieve the first pilot milestone; and that children whose primary caregiver scored highly on the worker-completed assessment of needs and strengths were more likely to reach the first performance point.

But despite these insights, there was general disappointment with the reality we had uncovered. And needless to say, the discussions around the major findings about outcomes were just as intense as those that took place about the service findings.

The Conference
Despite the disappointing findings, the process through which the evaluation took place led to a real sense of ownership of the findings on the part of the participants and their agencies. While Bowen and I had been presenting our preliminary findings of the qualitative and quantitative studies at national conferences, it was clearly time to unveil them to a more “local,” but perhaps a more important, audience. We had money in the budget for a statewide dissemination conference, and we went forward with it.

Held at the University of Michigan in October of 2003, the day-long First Research Conference on the Wayne County Foster Care Pilot Initiative was a smashing success. Fully planned with the evaluation committee over a number of months, over 100 people attended. They represented private agencies from around the state (including the newly inducted pilot agencies in Wayne County), the new top leadership of FIA (yes, the Democrats had won the election and now controlled the state bureaucracy), our funders, academic colleagues, and child welfare
When the Results Disappoint. The Process Matters Little

advocates. Everyone involved in the evaluation committee had a hand in the conference: Jim Beougher and representatives of the two private agencies that helped design the pilot discussed the background and history behind the initiative; Bowen presented an overview of managed-care initiatives in child welfare in the U.S. and of the pilot; and I presented the design and findings of the qualitative process study. After lunch, Bowen presented the quantitative findings through 300 days. Four agency representatives presented the lessons they had learned both from participating in the evaluation and from the findings, then four other committee members talked about the future - issues that new agencies had to be aware of as they implemented the pilot, and issues to be explored as new data were analyzed (we now had data well beyond the 300 days reported at the conference, and were looking at the termination of parental rights milestone at 515 days and the incentive payment attached to it).

After each of these presentations the questions flowed and the discussion was lively. People were clearly excited; most stayed for the whole day. And different people were excited about different things: about the work itself; about the partnership that had developed between the State, the University of Michigan, and the agencies; about the commitment of the agencies to learn from what they were doing; about the rigor of the study; and most of all about the practical importance of the findings and their implications for agencies starting, or who might have to start, down the performance-based path.

Four things about that day stand out for me. First, the agency representatives were magnificent. They not only knew their stuff and presented eloquently, but they weren’t thrown by difficult questions. They “owned” the evaluation process and were sure it was rigorous. They were sure that the data, no matter how disappointing, were reliable and valid. They were honest about what they had and had not learned, both about evaluation and about performance-based contracting.

Second, Jim Beougher was defensive and somewhat upset at the results. No longer in his role as the head of Children’s Services for the State, but now a County Director (he had been reassigned after the election), he first tried to attack the design and then tried to find any uncontrolled variable that might explain why pilot agencies had not achieved better outcomes than the non-pilot agencies. He was joined in this exercise by a higher up at one of the agencies that had designed the pilot but had not been at the evaluation table. They finally seemed to take solace in the fact that these were preliminary findings, and that once all the data were in differences would emerge that would show the true effectiveness of the pilot initiative.

Third, the new administrators from State FIA seemed incredulous at what we were hearing. “Are you saying that less service is being provided in pilot agencies?” “Yes,” someone would answer, “but there might be lots of reasons for that, some of which are not necessarily bad.” “You mean pilot and non-pilot children are as likely to end up in a permanent home at 300 days?” “Yes, that is correct, but there are lots more data to come, and we haven’t even looked at recidivism yet,” someone else would answer.

Finally, I remember someone in the FIA central fiscal office, Vic Bursankas, asking about the cost of the two comparison conditions. The question was asked publicly and we explained that we did not have such data. Later, at the end of the conference he cornered me and asked the questions again. I gave him the same answer. I didn’t think much about this interchange at the time.

More Data

Over the next year, data collection was completed, all the data cleaning was done, and all primary analyses were completed. We
When the Results Disappoint, The Process Matters Little

kept meeting with the agencies to present the findings. And while we had more nuanced information, the basic findings regarding service provision really didn’t change between the conference and the 930-day data collection mark. Controlling for all other factors in the study, children and families in the pilot agencies received fewer in-agency therapeutic and non-therapeutic services (but not fewer phone calls), and fewer completed referrals (but not fewer overall referrals) for out-of-agency service over their time in foster care than non-pilot children and families.

At the end of 930 days, there were some interesting and potentially important differences between the two groups in terms of outcomes. There were no differences between the groups in terms of the percent of children who achieved permanence because they were returned home or in kinship care, although significantly more pilot children went to kin and significantly more non-pilot children returned home, which has enormous ramifications for children and their parents but not performance-based contracting as set up in Wayne County. There was also no difference in the proportion of children whose parental rights had been terminated and were either waiting for an adoptive home or placed in one. The only major difference between the groups was that a significantly greater proportion of the pilot children (22%) were more likely to have been adopted than non-pilot children (4%).

Politics Again!
Shortly after the conference I received a formal request from Vic for the state identification numbers of the children participating in the study; he wanted to do a cost analysis. I felt that the release of such information from the research project would violate the approved IRB protocol; it would also clearly go against the spirit of the participatory nature of the evaluation. I took the request to the evaluation committee—the agencies could release these identifiers but I could not.

The group went through Vic B.’s proposal re: cost analysis comparing administrative payment amounts to pilot and non-pilot agencies, focusing on per diem and performance payments...After reviewing Vic’s memo, the group was very concerned that his analyses wouldn’t be able to capture a number of different types of non-IV-E costs and payments, such as county, local, and even funds that private agencies contribute themselves. The group also expressed great concern regarding the accuracy of the state’s quarterly records.

Bill asked for some direction from the group. The group agreed to invite Vic to the next meeting to discuss the matter further. Bill said that he would do this, and would give Vic a sense of the group’s concern. (1/22/04)

At the next meeting of the evaluation group, we thought this matter had been put to bed:

Bill said that Vic B. hadn’t replied to his message, so Bill expects that Vic is no longer interested in pursuing the cost analysis. Carmine asked if FIA could force us to give up our data at some point. Bill said that they could not do this, since he has a contract with FIA that will prevent this from happening. (FIA would have to use a court subpoena in order to obtain such information.)

Later in 2004 the real bomb hit, and it came from out of the blue. A bill introduced
in the state Senate sought to discontinue the pilot. The language in the bill cited a study conducted by (unidentified) researchers at the University of Michigan and cost analysis conducted by FIA that showed the pilot to be more expensive than traditional fee-for-service reimbursement as the reasons for doing this. I suspected, but never confirmed, that the new head of FIA had asked for this statutory language to be introduced. After all, the pilot was a Republican initiative, the preliminary findings that they heard were not “positive,” and there was a battle raging between the State and Wayne County around who should pay for what when it came to foster care services.

A number of things about this turn of events bothered many of us at the time, and some of what went on felt like personal attacks. First, people from the State never produced the cost analysis that was done, even when it was formally requested by the nonprofit agencies in Wayne County. The “findings” from the cost study were quoted in very general terms, but the study itself – its assumptions, methodology, data sources, actual findings, etc. – were not. To this day we have no idea whether Vic just went off and did the best he could or produced what he thought was wanted. What we do know is that evidence that was used had to be based on agency-specific, rather than child-specific, data, as the state never had our subject children’s FIA ID numbers.

Second, nobody from the State ever contacted me to discuss any new findings from the study or, for that matter, anything about the study at all. When I offered to talk with them about it, I was told they would get back to me, but they never did. They clearly had “heard enough” at the conference despite all of the caveats that were presented. Yet there were innuendos made that such discussions had taken place and that I had confirmed that they had made the “right” decision to close the pilot. Most painful of all was that after all of these years, some of the administrators at some of the agencies with whom I had worked questioned my integrity and thought I had been meeting with State officials behind their backs. After facing their questions and assuring them that I had never spoken to people at the State, my credibility was tentatively (but I never felt fully) restored, and the agencies confronted the people from the State with the fact that they were not aware of the most recent findings from the study or that more findings were still going to be produced.

Third, it became clear that the pilot had become embedded in a battle between the powers in Lansing and those in Detroit – what social services (and social service functions) the State was going to pay for and what the County was going to pay for.

Jim updated everyone as to the state of the Senate bill to end the pilot initiative. Wayne County is continuing to argue that it has no more money available to pay for foster care. Some private agencies have retained a lawyer to possibly push for an injunction against the County. Elizabeth Carey of the Federation [Bill Long’s successor] met with a number of private agencies, and Bill and Bowen, earlier in May to discuss how to gather data to counteract some of the arguments in the Senate bill. (5/23/04)

During the summer and fall, political efforts to save the pilot heated up. A number of groups organized toward this end, but the private agencies were fractured with some hiring their own lobbyists and others relying on the Federation to represent their interests. Interestingly, one of the groups that organized to save the pilot was led by a non-pilot agency in the study that had converted reluctantly to
the pilot system only in 2002—it was the agency that wanted to be a control agency because its leadership did not believe the pilot could be good for children and families. Their conversion to “defenders of the pilot” shocked me, and I didn’t understand it—I still don’t. They would never discuss their position with me, and their representative to the evaluation committee, who had been one of the most loyal, hardworking, conscientious, and thoughtful members of the group, began to withdraw from it. I still wonder if she had been ordered to do so.

Perhaps the executives at this agency, which was now operating under the pilot system, had found what I (and some non-political FIA folks who were involved with the project from the start) had begun to suspect and that the data seemed to be confirming—that less service could be provided and similar outcomes could be achieved and that agencies could profit under this form of reimbursement. My suspicion that agencies were, indeed, profiting from the pilot and doing so intentionally—a suspicion that made me feel very uneasy—started at the conference when the presenters were so severely challenged by the high-placed administrator at one of the pilot agencies, and grew even stronger afterwards.

This administrator was at one of the agencies that had helped to design the pilot and had the largest number of children in the study. Once the initial findings from the study came to light, he started coming to our evaluation meetings. His behavior at these meetings was disruptive (some told me that that was just his way everywhere, but I found him personally attacking and hostile). He questioned everything. Eventually, he also helped force from the agency (and our meetings) the representative that they had been sending to the evaluation committee—a man who was committed to the principles of participatory evaluation and to organizational learning. Ironically, this diplomatic, considerate man who cared deeply about children took a job with FIA in Lansing, but he would never talk with me about his former agency or about what he knew about the dismantling of the pilot.

Jim is leaving... in June to assume the directorship of FIA’s Division of Licensing. This is his last meeting with us. Thanks so much for your hard work and good cheer, Jim! We will miss you tremendously! (5/23/04)

I sensed the project had become a political liability when I was asked to meet with the Executive Director of this large agency—someone relatively new to his job. In the spirit of the participatory evaluation process, and hoping that we could keep the pilot alive until we had analyzed all of the data, we met in the early summer of 2004 in his office. What started out as a pleasant exchange about the study quickly turned into a “defense” of the findings. He called his “bulldogs” into the meeting—his data person and his “second” who had been attending our meeting. Together they questioned everything about the study, and when they realized that I could defend what we had done they tried to pressure me to “interpret things differently” (what a euphemism!). When I refused, nicely and politely, the meeting was adjourned. I should have anticipated what that would mean for the evaluation committee.

A Sad Ending
Bowen took his first academic job in the fall of 2004. Fortunately, by the time he left, the data set was complete and cleaned. There had been no fallout for us—yet. The agencies did all they could to help with the final data tasks.

Bill has closed three of four grants in order to consolidate the
financing of the research project. [He] needs all remaining invoices from agencies ASAP. The project officially shuts down 8/15/04. ...Bowen is leaving for Portland, OR on 8/19. Bill thinks that we should have meetings every six to eight weeks in the fall in order to pass out new analyses and begin planning for the late 2004 dissemination conference. We’ll need to decide when to hold the conference, how to organize it, and where it’ll be held. Bill has already secured the money to hold the conference. (5/23/04)

The summer “farewell” meeting of the evaluation committee was a disaster—only four people showed up to say goodbye to a project staff of five, with whom they had worked closely for a number of years, and to Bowen. And this set the tone for the fall when I was the sole person left at the university who had been involved with the evaluation. Only two agencies consistently came to the evaluation committee meetings, and the two agencies leading the charge to save the pilot—the largest of each of the pilot and non-pilot agencies—stopped attending meetings at which new findings were being disseminated. Their representatives to the evaluation committee stopped returning phone calls and e-mails.

Clearly, with the general findings known, people lost interest and drifted away. And I had become persona non grata, not because any of the project’s goals or procedures had changed, but because nobody wanted to know any more. The response to e-mails trying to rally the troops to learn more was something on the order of “just send us data memos.” And when I raised the need for planning meetings for the final dissemination conference, many didn’t respond, and those that did said basically “No, we don’t need (or want) another conference.” So I wrote and sent an e-mail to the committee, thanking everyone for their efforts through the years, reminding them of the many successes we had experienced, and formally ending the pilot evaluation.

Coda

I left the University of Michigan at the end of the 2004-2005 academic year. I had spent most of my time there putting together and carrying out the pilot evaluation. It was, in many ways, the crowning jewel in my 30-year research career; as a dean I will probably never do another piece of empirical work this challenging or complicated. It was the “perfect” study, done in the “right” way, with a wonderful group of staff and agency people, “blessed” by the research gods who allowed us to find funding when nobody thought we could “pull it off.” It was, I believe, a stellar project carried out with integrity that addressed a topic of considerable importance to child welfare systems wrestling with spiraling costs.

But this experience will always leave me only half-gratified. And it leaves me more cynical than when I started it. Yes, there has been much knowledge gained about the impact of performance-based contracting on the functioning of a child welfare system and the children and families it serves. Yes, there will be many academic papers published in top-tier journals with our names on them. Yes, Bowen has written a wonderful dissertation based on some of the data collected (McBeath, 2006). And in the end, I do believe it was probably right to end the pilot experiment—the findings about fewer services were disturbing; and the findings about outcomes have become even more disturbing to me as we have more carefully and completely analyzed these data. I worry more now about what happens to children and families under incentive-based initiatives than I did before I began this project.
Yet in the end I can take no credit for policy change or the new direction that the State of Michigan has taken in reimbursing nonprofit agencies for providing foster care. I believe those were blindly political decisions made without credible data to support them. Nor can I say that I have helped agencies use our evaluation data to make smart decisions that improved their functioning while also improving how children and families are served. In the end, the goal of organizational maintenance clearly outweighed the potential benefits of any organizational learning that was done, and these agencies (as best as I can tell) continue to function in the same ways they have always worked.

But perhaps, most cynically, I have come to believe that most, but not all, child welfare agencies, and many of the people who work for them, care more about themselves than the children and families they serve. I am only grateful that counterbalancing this heightened cynicism are the few program supervisors and workers I met along this journey who still believe that children and families can be healed even if the agencies they work for are not healers. And our data show that these few people, working in difficult environments, can provide appropriate amounts of services, even when fiscal pressures, are present to the benefit of those they serve.

And as I think of these colleagues and friends, I am reminded of the principles that I have learned over my career as a community-based researcher and that were reinforced by this project: integrity is essential, so guard it at all costs; create a strong team, not a collection of individuals; question your findings exhaustively, since others will certainly do so; enjoy the process, since it may end unexpectedly; and move from disappointment to new challenges. For me, in the end, crafting and conducting this study has come to mean so much more than the “results.”

References


William Meezan, D.S.W., is a Dean and Professor at Ohio State University College of Social Work. Bowen McBeath, Ph.D., is an Assistant Professor at Portland State University Graduate School of Social Work. Comments regarding this article can be sent to: Meezan.l@oau.edu, or to: mcbeath@pdx.edu.

(Footnotes)

1 This narrative is written in Dr. Meezan’s voice and primarily from his perspective. Dr. McBeath is, however, a full coauthor of this work, as he chronicled the process we went through together and was a full partner in the research. Parts of the experience reported here are taken directly from his field notes and his dissertation.

2 Defined as the reunification of a foster child with his or her biological parent, the placement of a child with a relative (“kinship care”), the establishment of legal guardianship for the child, independent living, or adoption.

3 While the foster care pilot did not contain some common managed care characteristics, it was considered by both FIA as well as private providers to be a managed care system. FIA chose to neither capitate foster care payments to private contractors, integrate foster care systems across agencies, nor create a single case management agency for the Wayne County pilot. However, official FIA documents mention the foster care initiative as a managed care system. Moreover, private providers have given public presentations in which they have referred to the Wayne County pilot as a managed care system. Agency personnel often equated the foster care pilot initiative with managed care. Therefore, for both FIA as well as involved private service providers, the foster care pilot clearly was an experiment in managed care.

4 All italicized text comes from the archive notes from the project recorded by Bowen McBeath. Dates at the end reflect the date the evaluation group meeting took place.

5 I suspect I was not named because of the public fuss I had made when Wayne County FIA pulled out of the initial evaluation, and had I been named in any way in this bill I would have spoken up that the data from the evaluation were not all analyzed.