

Speech by Douglas W. Nelson
President, The Annie E. Casey Foundation
at Duke University's Terry Sanford Institute of Public Policy
to the
Foundation Impact Research Group Seminar
March 9, 2005

I am honored by this opportunity. Like the entire philanthropic field, I have immense admiration for Dr. Joel Fleishmann – both as a leader and a thinker.

This invitation grew out of some interest Joel has had in the Casey Foundation and, more particularly, out of a conversation the two of us had on the general subject of “foundation effectiveness” a few months ago. We talked then a little about the use of data, the importance of social change hypotheses, the role of measurement and how all of these things have shaped the Casey Foundation’s identity and the way we try to do business.

At Joel’s request I’ve put together some notes on those subjects and look forward to your guidance and comments to shape my thinking.

I must confess – at the outset of this talk – to some trepidation. I’m a little bit worried that some of my conclusions about innovation, measurement and evaluation will appear painfully obvious to many of you who have thought longer and deeper than I have about social change. I’m also a little afraid that my deeply earnest and personal sense of the importance of a results-orientation to quality philanthropy will render me tediously preachy on the subject. So consider yourself warned.

Let me begin by saying that the Annie E. Casey Foundation’s overarching mission is to find ways to improve the conditions and adult outcomes for America’s least advantaged, poorest and most likely to fail kids. More recently we’ve begun to reformulate that mission by saying we’re working “to narrow the gap in opportunities and outcomes between America’s more advantaged kids and its less advantaged.”

In pursuing that mission, the Casey Foundation prides itself on a thorough-going commitment to data collection, constructing baselines, developing quantifiable and measurable objectives (for our initiatives, experiments and grants), and tracking the outcomes of our investments and interventions against clear baselines and goals. We don’t do all these things well all the time, but we almost always try to do them.

The origins for Casey’s emphasis on data, measurement, and deliberate social change “experimentation” are many. And lots of it is a simple reflection of where the foundation field as a whole is going. But some of the impetus for our approach is a reflection of who I am – the long-time CEO – and, more importantly, of who our founder – Jim Casey – was.

Let me begin with the story of the lesser light – myself. Almost thirty-years ago I was a Ph.D. candidate at the University of Wisconsin – then studying and actually teaching the history

of American philosophy. My absorbing interest was late 19th century American pragmatism. I was particularly taken with two pioneers of pragmatism – Chauncey Wright and Charles Peirce. Wright, in particular, was a wonderfully clear advocate of the view that all propositions, including propositions about society, could be framed as hypotheses, and that the validity of such hypotheses turned on whether they could predict measurable social effects. Charles Peirce, in turn, was adept at making the case that repeated, careful hypothesis formation and experimentation would lead in time to ever more valid and predictive hypotheses or “theories of change.” You simply had to be imaginative, be willing to measure, willing to admit when you were wrong, keep trying, and, sooner or later, you would discover “what worked.” Moreover, Peirce believed that recurring evidence of predicted consequences (i.e., “of what worked”) was usually powerful enough to convince other well-intended thinkers of the validity of sound theories and propositions, regardless of those thinkers’ preconceptions or values or subjective point of view.

Now all of this would have remained an innocent academic pursuit, if it weren’t for a fluke. One of my fellow history instructors worked part-time for the Governor of Wisconsin. This guy thought I gave amusing lectures, and asked me if I would be interested in writing speeches for the Governor. I told him I didn’t know a thing about who the Governor was, what he believed, or what state government did – and said “no thanks.” The Governor called me up that night at home . . . asked me if I would draft a speech for him on the subject of aging and old folks – and I, being 24, said “of course I would.”

I knew absolutely nothing. I sat at the kitchen table pouring over the Wisconsin budget materials on the elderly that the Governor’s office had sent. I could come up with only three thoughts. First, the state of Wisconsin appeared to be spending almost all of its so-called “aging” money on nursing homes – 90+%. Second, my grandmother and all her friends hated nursing homes and it was the last kind of help they wanted. Third, given those two facts, I could only conclude that maybe Wisconsin ought to experiment with alternative forms of helping the frail elderly that might turn out to be more popular, more useful and less costly.

So I turned those three observations into a 13-minute speech in which Governor Lucey announced his decision to experiment and carefully evaluate in-home care for nursing-home eligible old folks. To my amazement (and the Governor’s), the speech was an enormous success . . . a year later I was the Commissioner on Aging . . . two years later the state placed a moratorium on nursing home beds and began to build flexible home care help in every Wisconsin county. Over the next few years, Medicaid expenditures for nursing homes leveled off; less costly home care grew dramatically; consumer satisfaction for the new services soared; and Wisconsin today has the best community care system in the country. Furthermore, it hasn’t licensed a new Medicaid nursing home bed since 1981.

I got promoted to Assistant Secretary for Human Services and a few years later left the state (before people discovered I didn’t know anything) to join a small social policy think tank in Washington. I came off my experience in state government filled with the conviction that the deliberate reform of public policy, state systems, and service spending could lead to dramatically improved outcomes for clients and to improved cost-efficiency for taxpayers. I still believe in

that, although I have a much greater appreciation about how hard it is and how important it is to be lucky as well as “scientific.”

Anyway, I was at the Center for the Study of Social Policy in 1988, when a bunch of UPS executives made an appointment to get our advice on how to spend the recently expanded assets of the Annie E. Casey Foundation (a Foundation I had never heard of) that had been founded by James Casey – the man who launched and built the United Parcel Service – UPS.

I still really had only one idea. I began our consulting session by telling the UPSers the safest thing they could do was build children’s museums and playgrounds, provide scholarships, support youth development programs and continue providing long term foster care. I said that such an approach would surely do some good, would be appreciated and would earn their foundation lots of praise. I did, however, add that this approach wouldn’t actually make much of a dent in what was happening to disadvantaged kids – as a whole – in this country. I pointed out that all of Casey’s expanded endowment (about \$600 million then) was sufficient to run the New York City public schools for about three weeks. I went on to argue that if they really wanted to change the future for lots of at-risk kids, they’d have to find ways to persuade government and traditional non-profits to spend their time and money differently . . . to spend it much more effectively and efficiently. To me, that meant that the Foundation needed to document existing inefficiencies and shortcomings in current performance of key public human service agencies; they had to design and experiment with new and different approaches to at-risk kids and families; and they had to be willing to walk away from those approaches that didn’t improve outcomes while building on and promoting every approach that did show positive results. One guy said it sounded more like an R&D enterprise than it did a charitable foundation. I said he might be right, admitted I didn’t know much and we parted. I thought they were a little disappointed and bored by the advice we had given. A year later, the Chairman of UPS called me and asked me if I wanted to run the Annie E. Casey Foundation – the only condition was I had to go to work for UPS for three months – do a little bit of every job. The notion was that this experience would help me better understand the values and intentions of the Foundation’s founder.

It was during those three months that I learned about UPS, about Jim Casey, and about why all my simple minded enthusiasm for measuring, evaluating and building on what works led UPS executives to take a chance on me.

Jim Casey, who ran UPS from 1907 to about 1970 was an early and thoroughgoing advocate of what was called, in the 1920s, “scientific management.” He believed efficiency produced profit. And he believed that efficiency was achieved by measuring everything – by keeping track of the cost (in time and money) of every step in the process of achieving a result – in this case, the result of successfully delivered packages that met customer expectations. Further, Jim Casey believed that whenever you found a process that improved efficiency, you made it standard practice and you supervised employees to achieve fidelity to that practice.

To this day, UPS is a culture of system analysis, process analysis, cost measurement, and outcome measurement. Drivers always carry their keys on a specific finger; they step in and out of the truck in a particular way, they store their dollies in the same corner – all because these and

scores of other practices have been documented as most time and effort efficient. Further, every UPS driver is measured everyday on the number of stops/hour; percent of mis-deliveries; percent of damaged packages; accidents per 100,000 miles and so on. This data is used to identify problems, to encourage continuous improvement and to reward performance.

People who grew up in this measurement culture have and continue to make up the majority of our board. Not surprisingly, they have not just supported Casey trademarks – documentation of system inefficiencies, the creation of accurate baselines in judging the effects of initiatives, careful identification of new practices or processes designed to produce better results, and continuous measurement of the effects/results of service experiments and system reform initiatives – our Board has not only supported this, but for 15 years they have reminded us we don't do it well enough, completely enough, clearly enough.

Of course, we remind our stewards that changing conditions and outcomes for disadvantaged children and families is more complicated than moving boxes; they agree and tell us that's our problem . . . that's why we get paid all that money without having to make a profit.

In many ways, the history of Casey over these last 15 years has been the chronicle of our struggle to meet a core challenge – to usefully bring data collection, hypothesis formation, and results evaluation to bear on our mission of improving public and private efforts to help disadvantaged kids.

Some of our efforts have been very constructive and helpful, some have been frustrating but instructive, and still others have great promise, but it's too early to declare victory.

For example, one of the first and most continuous Casey investments has been in a national data collection effort that we call "KIDS COUNT." Since 1990, we have collected and published data on the condition of kids and families (things like infant mortality rates, adolescent child bearing, poverty rates, high school completion rates, child death rates, percentage of single parent households) for every state, and now for every county, and every large city. Furthermore, KIDS COUNT tracks changes in those rates from year to year for each of the state and local jurisdictions, and by doing so, allows a huge number of illuminating forms of comparative and trend analysis.

We've spent a lot of money on KIDS COUNT. We spent it because we thought it would raise public awareness about the needs of kids and families and the relative failure of our existing efforts to meet or address those needs. And media tracking suggests it has raised awareness. We distributed 180,000 copies of the KIDS COUNT data last year – making it the nation's most widely read publication on the condition of children. We also thought it would increase and strengthen public and political accountability for improving conditions and incentivizing greater, more innovative actions by states – and we think it has. We did a recent survey that showed 80% of state legislators are aware of KIDS COUNT and they or their staff frequently use it in decision making about child and family policies. Given the reading habits of state legislators, that's amazing.

We also think KIDS COUNT data has focused and sharpened charitable investment on behalf of kids. Last year it was estimated that thousands of applications for foundation funding from kid-oriented non-profits cited KIDS COUNT data as part of their problem statement or goal setting. All in all, we think KIDS COUNT has measurably elevated the data orientation and results accountability in the investment the country makes on behalf of poor children and families.

Other of Casey's data based, hypothesis driven investments and initiatives have had more mixed results. For example, our first effort to demonstrate a large scale reform initiative was an experiment we called "New Futures." Our theory was that the high degree of failure experienced by at-risk youth between the ages of 14 and 19 could be dramatically reduced if communities were given the data documenting the extent of failure; if they were given incentives and opportunities to increase collaboration between public schools and social services; if they were given access to innovative program models and experts to help them create new interventions, and if they had \$9 million in Casey money over five years to do all this.

We set out to prove this hypothesis in five cities. Part of it worked as predicted. The creation of baseline data on the heretofore unknown rates of academic underachievement, dropouts, adolescent pregnancy and the disproportionate share of these and other rotten outcomes that fell on African American youth did, in fact, prompt the hoped for mobilization, sense of urgency, and effort in all five cities. But it also became clear in time that the theory that a little more collaboration, more service integration, and the introduction of some promising new programs was woefully insufficient to meet the wildly optimistic goals that each city set for school improvement, pregnancy reduction, and delinquency prevention.

Casey learned that we needed a much more robust and comprehensive set of opportunity and service-changing interventions, if we were going to make big changes in the outcome trajectory of high risk, older teens.

By 1994, the tracking data told us we weren't accomplishing enough and we had to analyze our mistakes so others could learn from them and so that we wouldn't remake them. We wrote a self-evaluation of New Futures reflection called "The Path of Most Resistance" and I gave several talks on lessons we had learned.

In the course of doing this, I learned something about Casey's "peculiarity" as a big foundation. People were actually stunned that we were publicizing our disappointment, telling people our initiative hadn't worked as well as we hoped, and marshalling the evidence that proved we fell short of our goals in every city. I was proud of our honesty and was frankly confused by my philanthropic colleagues' sense that we were crazy to be so straightforward in our evaluation. But times and the field have changed over the past decade. I think it's fair to say that the field today embraces a much more positive view of foundations' responsibility to be held accountable for candidly reporting results. That's a good thing.

Now, I don't want to leave the wrong impression. Lots of Casey's subsequent data-based, hypothesis-driven, system improvement initiatives have succeeded in showing the way

toward what works; they have yielded the predicted improvements in client and system outcomes, and many of our strongest initiatives have been widely replicated.

I won't bore you with these happy successes, but let me mention a few.

In the early 1990s we launched a foster care reform initiative called Family-to-Family. We began with an analysis of the chaos and lack of results that had long plagued America's child welfare system: more kids entering care; kids stuck in care; kids hurt; bad adult outcomes; rapidly growing costs. Our theory, put over-simply, was that good system data, the recruitment of foster homes in high-need neighborhoods, the introduction of team decision-making, improved training and supervision of frontline workers, and improved communication between foster and birth families could together reduce the number of kids coming into care, shorten the length of stay, reduce the percent of kids in group care, and diminish the level of abuse and instability experienced by children while in care. I'm happy to say our pilot sites met most of these goals, and Family-to-Family is now being effectively implemented in over 15 states and over 80 counties.

Similarly, Casey, in the 1990s, launched an experiment called the Jobs Initiative. Like our other initiatives, we began with a problem analysis – specifically, the very poor employment outcomes experienced by residents of low-income neighborhoods who were supposed to benefit from expensive state and federal workforce development programs. In light of that problem analysis, we got a lot of smart folks together to design a better approach. The hypothesis they eventually developed went something like this: if you get key employers involved to design worker training, if you focus placement efforts on high employment growth sectors, if you equipped prospective workers with certain key soft skills, and if you provide some family supports to newly place workers, you will get much better employment, retention, and earning results for the same investment.

The Jobs Initiative has now been operating in six sites for almost nine years. To date we have helped over 23,000 workers get, keep, and improve their employment. The careful and expensive outcome tracking in almost all of our sites shows, when compared to historical results, a higher percent of applicants get employed; a high percent of those employed get retained; the earnings increases of those retained exceed the targets set by the projects. The strategies and practices of the Jobs Initiative are now being replicated or incorporated into workforce development systems and welfare to work systems across the country.

I would like also to tell you about Casey's Juvenile Detention Alternative pilots – which were designed to prove that it is possible to reduce the waste, harm and mis-targeting that has characterized the terrible and terribly expensive youth detention systems in most U.S. jurisdictions. In our three most "mature" sites—Cook County (IL), Multnomah County (OR) and Santa Cruz County (CA)—reliance on secure detention was reduced by at least one-third (in one site, by two-thirds) through the implementation of objective decision-making tools, increased program options and faster case processing. Skeptics feared that these kinds of dramatic decreases in the use of detention would unleash a juvenile crime wave. Quite the opposite has happened. In each of these sites, key indicators of juvenile crime have dropped to their lowest levels in decades because officials are now better able to identify those youth likely

to continue to offend and, also, because their systems are being held accountable for public safety outcomes. Finally, these jurisdictions saved literally millions of taxpayer dollars by avoiding construction and operation of new detention beds.

I could go on with other hopeful examples, but I want to save some time for a few summary thoughts – and I'll be glad to send data on any of Casey's measured initiatives to any of you who want more detail.

What I want to close with are some conclusions I've drawn after 15 years about the advantages and challenges of data-driven and results-focused philanthropy. On the positive side, there are many advantages to incorporating baseline data, clear hypotheses, and results tracking into the design, implementation and assessment of philanthropic investments:

- First, it makes foundations think. It makes us clearer-headed about what we are trying to do and whether we are actually doing it.
- Second, this commitment to public and objective goals makes foundations and grantees morally accountable. Results measurement makes it possible to determine whether one's investment and actions actually succeed or fail . . . do good or do harm . . . are worth the cost or not.
- Third, it creates the basis for intelligent and effective program management, refinement, and improvement. Results tracking does this by giving funders, implementers and policy-makers some continuous objective feedback on the impact and effects of their implementation efforts.
- Fourth, it provides an objective, non-partisan, non-ideological basis for social change advocacy and for intelligent replication. Data about investments that produce desired measurable results is convincing to Republicans, Democrats, liberals, conservatives, to the sentimental and the hard headed, to business folks and social workers. It allows us to make policy decisions on the basis of what actually works rather than what we may want to believe ought to work.
- Finally, another big advantage of data-based and results-focused grant-making is it tends to lift up critical questions of social justice and fairness. Increasingly, the Casey Foundation is using gap analysis as the broad framework for our goal setting and results tracking. Put another way, we are finding it useful to track the widening or closure in the outcome gap between black and white students, between inner city and metropolitan areas, between cities and states, between poor states and rich states.

This gap theory of goal setting is a complicated subject, but one that has lots of promise. In our *Making Connections* initiative, for example, we are setting goals to increase employment in each of our sites. But instead of saying, for example, that a 5% increase in a neighborhood's employment rate is our goal, we are saying we want to cut the gap between inner city and suburban employment by a half. Such a target not only floats with bigger economic variables, it introduces a long term goal of social equity into the shorter term result of increased jobs for a particular target group of disadvantaged individuals.

Those are some of the arguments for taking the measurement and results path that the Casey Foundation has struggled to follow. But this wouldn't be a very complete overview if I didn't acknowledge the limitations, risks and downsides of outcome-accountability.

- To begin with, you increase the probability of public failure. Public failure is no fun and the prospect of failure – rather than unqualified gratitude for good intentions – may (for some people) reduce the motivation for and satisfaction in philanthropic giving.
- Second, foundations that impose specific results on their grantees that their grantees can't realistically achieve or that are inconsistent with their mission and core capacities can end up actually doing harm to grantee competence, confidence, and sense of purpose.
- Third, an unsophisticated commitment to achieving results (or an excessive fear of falling short of benchmarks) may actually discourage risk-taking, innovation and experimentation. It may lead, instead, to an over-reliance on the tried and true. Results data is essential for meaningful social experiments, but if poor results are heavily sanctioned, grantees may avoid experimentation, creativity and innovation.
- Fourth, the commitment to data collection, data tracking and evaluation is expensive. These expenditures may expose philanthropy to the public charge of spending too much time and money studying and analyzing, and too little just helping.
- Finally, “the results philosophy” probably doesn't make compelling sense for small foundations whose resources and staff may not be sufficient to support a learning, experimentation or field influencing role.

For big foundations, however . . . for foundations whose missions imply an aspiration to make a big difference, to change the world some . . . for these institutions, quantifying problems, defining clear hypotheses, setting measurable goals and tracking results seem indispensable to a real commitment to their mission. At least we think so.

Thanks for listening.