

Philanthropy, Evaluation, Accountability, and Social Change

John Bare, Ph.D., The Arthur M. Blank Family Foundation

Key Points

- Many foundations have substituted process accountability for accountability for contributing to social change.
- While process accountability is important, it sets a floor, not an aspirational ceiling.
- There are tools—such as risk analysis, systems approaches, and game theory—that can help philanthropy engage in work on complex social problems that cannot be deconstructed into a series of small, linear projects.
- Seeking to extend basic human rights to more individuals around the world, seeking to reduce racism in a given city, or seeking to change public-health norms in small town—all of these aspirations require first a willingness to take on challenges that defy short-term, causal, quantifiable results attributable to a best practice.

It is time to acknowledge that philanthropy's accountability movement is something other than what is advertised.

All foundations know the accountability checklist. Avoid self-dealing. File your 990 on time. Pay the excise tax. Be transparent about your grant processes. Avoid excessive compensation. Avoid luxury perks for board members. Commit to board development. Do not bully nonprofits. Keep administrative costs low. Have HR establish annual performance reviews for individual staff members. Benchmark internal processes against those of other foundations. Seek cause-and-effect

attribution regarding quantitative outcome targets for grants.

What is missing is an explicit recognition that this set of behaviors is about setting a floor for minimum standards. Treated as such, it is a useful thing. It is useful in the same way that the American Bar Association's Code of Professional Responsibility defines the minimum standards lawyers must meet on ethics but does not describe what it takes to be a great, world-changing attorney. Every profession or field needs agreed-upon minimum standards. Fall short of these minimums, expect to get dingd.

Yet for several reasons, philanthropy has treated the accountability movement as so much more.

Attentiveness to accountability is sometimes offered up as a sufficient indicator of impact. These foundations point to their internal accountability efforts as evidence of a commitment to effectiveness, despite the fact that funders can hit the mark on every accountability measure and still contribute nothing of external value to society. This bait-and-switch is convenient. It allows foundations to define success through internal metrics, not whether their social investments generate any transformative value for society.

For foundations, it is always easier to show success on bright-line measures of accountability—and take credit for success—than to pursue an

ambitious social change agenda, where definitions (much less measures) of success are murkier. When all else fails, it is easy enough to lower the standard used to define success, as states seem to have done under No Child Left Behind (de Vise, 2007; Sturrock, 2007; Wallace, 2007). Accountability, in the end, is about defining the minimum level of acceptable performance, not the highest level of accomplishment we want to pursue.

Setting the Norm

As a field of practice, philanthropy has begun to use the accountability movement to establish a norm for what sound philanthropic investments should look like. This is where the serious harm can occur. The accountability movement requires that the norm be closer to the floor than the ceiling. Foundations become locked into the notion that philanthropic investments must carry short-term, unambiguous revelations of causal attribution. Simple, linear tools are sufficient to manage and evaluate these efforts.

This is a narrow view of what is possible. Within these limits, foundations are less likely to invest human or financial capital in complex change agendas, where outcomes are far removed from current interventions. The distance between today's investments and tomorrow's outcomes, of course, deprives current investors of the opportunity to take credit for any good that may occur. The ability to take credit is a key component of the accountability movement.

Adhering to the accountability movement requires foundations to avoid social change agendas where outcomes cannot be easily quantified and where problems are too large for any one funder to produce. Accountability means being able to carve out the unique contribution of any individual funder, an impossibility in long-term, complex change efforts.

When foundations apply accountability tools to their grant-making approach, the tail ends up wagging the dog. Internal process draws more attention than external impact. In the same way, research shows that factory workers react to daily production targets by calibrating their effort, so

as not to work too hard or too fast, in an effort to get the most pay out of the least work. The effect is to tamp down production (Surowiecki, 2005).

For foundations, when they attempt to deconstruct complex social change agendas to create bite-size, measureable grant projects—those with quantifiable measures and easy attribution—the foundations lose contact with the larger purpose of their work. As Peter Block (2008) writes in challenging us to think differently about community transformation, we cannot assume that “an aggregation of individual changes” will add up to large-scale change. “The mindset that we can program and problem-solve our way into a vision does not take into account the complexity and relational nature of community,” Block writes.

When foundations apply accountability tools to their grant-making approach, the tail ends up wagging the dog. Internal process draws more attention than external impact.

Further, foundations investing through an accountability model find themselves using fidelity to process and plan as the determination of success. It is the same way in business: The accountability movement has distracted firms from the work of producing a great product or service. Instead, meeting the internal process metrics becomes the strategy itself, and it easily becomes corrupted (Hymowitz, 2005a; Hymowitz 2005b). This is a great irony, and failure, of the accountability movement. Despite all the rhetoric about impact, the real tests are about compliance to process. And linking rewards to progress on short-term, tactical targets, especially financial goals, “causes people to game the system and in doing so destroy value,” Harvard business professor Michael Jensen (2003) writes in a paper, “Paying People to Lie”:

The accountability movement's obsession with fidelity to process can be traced to the rise of standardized processes that powered mass production, which was organized by the "measuring and coordinating of all of the *activities* of the manufacturing plant so that the speed and efficiency of these activities could match the speed and efficiency of the machines themselves" (Head, 2003, emphasis added).

Philanthropy's overreaction to and overuse of tools anchored in the accountability movement have deprived foundations and nonprofits from approaches that can better help them produce complex social change outcomes.

What George Eastman did for mass producing photographic film, what Henry Ford and Frederick Winslow Taylor did for mass producing automobiles, and what James Buchanan Duke did for mass producing cigarettes all required precise control over and accountability to standardized *activities*. When this accountability model is transferred to fields of medicine or customer service, the test is not about better outcomes. The test becomes fidelity to process, for example, physicians following "decision-making algorithms" designed by process engineers working for insurance companies. The effect is to "subordinate the skills" and judgment and creativity of a professional to "an industrial, assembly line discipline" (Head, 2003).

In philanthropy, this aspect of the accountability movement shows itself with human-service organizations being forced to serve both the requirements of the funder and the needs of their client families. When teaching an evaluation class for nonprofit executives, I had one woman explain that her organization had a foundation

grant to deliver health services to low-income families. The funder required all grantees to adhere strictly to the preferred implementation model. To keep their funding, grantees had to deliver detailed reports to the funder describing how they were following the model. For this executive, however, her staff had discovered that they could get better results by varying the model based on the needs of client families. Varying the model was not permitted. To keep the funder happy while also pursuing authentic social change for client families, the organization decided to keep two sets of books. The group used one set of reports to satisfy the evaluation requirements of the funder. The second set of books accurately tracked the work and the outcomes.

The Challenge at Hand

The central thesis is this: Philanthropy's overreaction to and overuse of tools anchored in the accountability movement have deprived foundations and nonprofits from approaches that can better help them produce complex social change outcomes.

One result of this pathology is that philanthropy's leaders have come to define success, and effectiveness, as being in compliance with minimum standards and not in violation of an expressed operational prohibition. This condition is at times a source of frustration and at times a convenient dodge for those seeking to aim for the floor and claiming to have reached the ceiling.

The good news: At hand is a more robust collection of tools that enable foundation and nonprofit executives to pursue, manage, and evaluate complex social change strategies as they unfold in real time. Following a discussion of the rise of the accountability movement, this article offers eight such tools. The set of tools offered here is intended to be generative, not exhaustive, and intended to be used in tandem with accountability tools, not in place of them. The appropriate solution is to fit the right tools to the right jobs, addressing issues of minimum compliance and aspirational social change.

In the discussions below, the examples are assembled from a variety of fields of work. This is intentional, so as to illustrate how the sweeping accountability movement became so dominant that its management and evaluation practices were easily the tools of choice for foundation and nonprofit executives. Likewise, as useful alternatives become evident in various fields, it is appropriate for philanthropy to take note. Amassing lessons from multiple fields of practice is a reminder of the power of borrowing and adapting tools already deployed for other uses.

A Larger Cultural Movement

Office Space, the 1999 Mike Judge film, shines light on the accountability movement that has dominated American life, certainly far beyond philanthropy.

In a key scene, Ron Livingston's character, Peter Gibbons, meets with two consultants—evaluation consultants, in effect—who are on site to interview staff ahead of layoffs. In describing the management at his firm, Gibbons explains that he has eight different bosses, with each waiting to jump on him if he makes a mistake:

Eight, Bob. So that means that when I make a mistake, I have eight different people coming by to tell me about it. That's my only real motivation is not to be hassled, that and the fear of losing my job. But you know, Bob, that will only make someone work just hard enough not to get fired.

The scene reveals two critical shortcomings of the accountability movement. First, by articulating the *minimum performance standard*, it motivates individuals only to clear the lowest acceptable bar. Sports psychologists (Stevenson, 2006), for example, know that the threat of punishment for poor results, a key aspect of the accountability movement, "generally isn't the most effective way to motivate someone." Second, the culture equates success with avoiding (or at least hiding) mistakes. Both variables inhibit the pursuit of ambitious, long-term efforts to transform society.

The *Office Space* scene is a dramatic illustration of the "gotcha" element of our culture, where oppo-

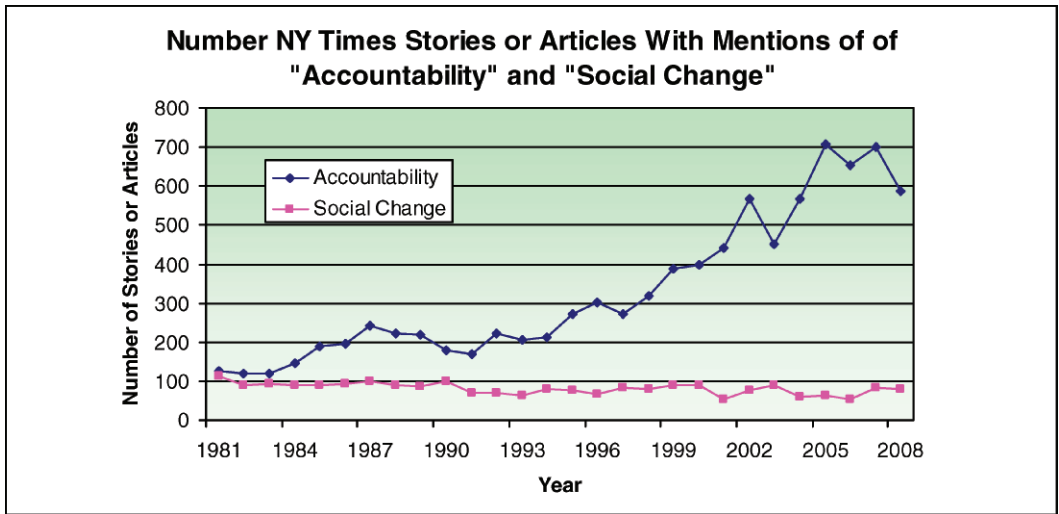
nents stay ready to pounce on any misstep or flaw. In our politics, we are left with a constant state of tit-for-tat charges that creates a continual cycle of investigation and accusation. Success, again, is equated with no one spotting any mistakes that can lead to a "gotcha" moment. Given that humans are flawed, mistakes are inevitable. Thus success can only be achieved by doing nothing. So in practice, individuals most likely to succeed in the accountability movement tend to be managers who maintain their position while doing as little as possible.

In our politics, we are left with a constant state of tit-for-tat charges that creates a continual cycle of investigation and accusation. Success, again, is equated with no one spotting any mistakes that can lead to a "gotcha" moment. Given that humans are flawed, mistakes are inevitable. Thus success can only be achieved by doing nothing.

Then-Vice President Al Gore appeared on David Letterman's late-night show on Sept. 8, 1993, to poke fun at this "gotcha" culture, which he believed was inhibiting the federal government's ability to improve the social condition. To point out the folly of federal procurement regulations, Gore put on protective eyeglasses and used a hammer to smash a glass ash tray. In a scene that could have come from *Office Space*, Gore explained that government rules said ash trays "should break into a small number of irregular shaped pieces not greater in number than 35" (Gruber, 2003). Gore's point: With so much time and money spent monitoring the nature of ash tray breakage, it is a wonder anything substantive gets done.

Gore's effort, formally known as the National Partnership for Reinventing Government, had

Figure 1 Annual "Accountability" and "Social Change" Mentions in the NYT, 1981–2008



an ambitious mission. President Clinton (1993) explained that the effort would “make the entire Federal Government both less expensive and more efficient, and to change the culture of our national bureaucracy away from complacency and entitlement toward initiative and empowerment. We intend to redesign, to reinvent, to reinvigorate the entire National Government.”

In the end, Gore failed to move political and media classes away from the accountability culture and onto more substantive issues facing the country. As shown in Figure 1,¹ the accountability movement continued in ascent after 1993. This content analysis provides a quick way to observe the degree to which the national conversation has been focused more on accountability or more on social change agendas.

Back in 1981, the first year of the analysis, the terms “accountability” and “social change” received

¹ Graph was created using the “Advanced Search” function at the *New York Times* Web site, <http://query.nytimes.com/search/alternate/query?query=&st=fromcse>. For each year, 1981–2008, I carried out separate searches for the term “accountability” and a search for the term “social change.” The *Times* Web site reports a total number of stories or articles that include at least a single mention of the term in question. The count reported is a tally of the total number of stories or articles that include the term in question. The term in question may appear once or more in each entry. The Web site search was current as of June 19, 2009.

about the same number of mentions in the *New York Times*. Since then, “accountability” has dominated the national conversation. Through two Iraqi wars, an impeachment trial of a president, a failed presidential vote-counting system, and the rise of digital media, there has been a rise in the number of proclamations of accountability and charges of failure to maintain accountability. The more nuanced idea of “social change” has not yet become a substantial part of the national conversation.

It Turns Out Everyday Life Is Complex

There is a place for the accountability movement, of course. It is well-suited to the clarity associated with minimum standards, such as a ban on foundation self-dealing. And every field, philanthropy included, needs agreed-upon minimum standards. These are the non-negotiables, the lines we shall not cross. Yet as much as we need these bottom-line standards, these are only a necessary component of a strong philanthropy sector. We also need to challenge ourselves to pursue greater and higher aspirations than are defined by accountability regulators.

As with any instruments, accountability tools can be overused. Figure 1 shows this almost certainly is the case here. As the accountability movement has become dominant, the associated management and evaluation practices have been applied in almost every setting without regard to whether the tools ideally are suited to the situation.

A telltale buzzword is “simple.” The accountability movement thrives on the notion that everything can be reduced to a simple, linear approach. In general for the business world, the term “complexity” equates to bloated, wasteful operating inefficiencies (Ashkenas, 2007). There is no acknowledgment of the need to differentiate behavior to suit the complexity of the system in which an individual is working. Complexity is viewed as a problem to remedy, not as a characteristic of a system.

Some people have challenged this conventional wisdom. University of Michigan business professor Karl Weick (Coutu, 2003) urges leaders to be skeptical of ideas presented as simple. This is because we are operating in a world that is quite complex, and “leaders must complicate themselves in order to keep their organizations in touch with the realities of the business world. My worry when executives say, ‘Keep it simple, stupid,’ is that they’re underestimating the complexity of their own organizations and environments.”

Comparing dominoes to pick-up sticks helps make the distinction. Constructing a traditional logic model is akin to setting up a string of dominoes, where knocking over a single domino sets off a chain reaction and topples every subsequent domino in succession. The string may be long and may even be elaborately designed. But the system is always linear and simple. And as with any simple logic chain, we can always pinpoint which input (i.e., which domino) must topple to produce a specific output (i.e., the next domino in the line). Each output then becomes the input for the next connection in the line of falling dominoes. Here accountability rules may fit well.

On the other hand, consider the game of pick-up sticks, where each player in the game attempts to remove a single stick from the pile without moving any of the other sticks. This is closer to real life, where interconnections make it difficult to touch one aspect of public life without affecting another. In fact, we can make the game even more real. Imagine the game where every player is moving sticks at the same time. Now it is nearly impossible to predict exactly how the pile will move. We are left with little or no horizon for

planning, as each player in the game moves sticks whenever they want.

In this volatile setting where linear planning approaches do not carry much value, what we need are robust tools that emphasize a continuous cycle of feedback and rapid adjustment.

Dominoes places a premium on a very precise plan. Winners are those who do the best job of designing and sticking — exactly — to the plan. Dominoes greatly reduces, and may even eliminate, uncertainty.

Precise plans do not carry much value in the real-world version of pick-up sticks. Here winners are those who can most rapidly harvest, process, and *act on* feedback, in real time. To achieve great results in complex situations where the world around us is messy and unpredictable, we have to adapt as we go. In complex situations there is no chance of eliminating uncertainty. The best we can do is reduce uncertainty.

Foundations investing through an accountability model are essentially limiting themselves only to problem-solving opportunities that line up neatly like a chain of dominoes. Or they ignore the complexities of the issues confronting them and force an overlay of simple tools and interventions. In either case, the effect of the accountability model is to limit the benefits that accrue to society.

What Works? Versus What Works, for Whom, Under What Circumstances, for How Long?

The Obama administration is contributing to the myth that we can impose simple processes on complex situations. First Lady Michelle Obama reinforced the bias in announcing the \$50 million White House Social Innovation Fund: “The idea is simple: to find the most effective programs out there and then provide the capital needed to replicate their success in communities around the country that are facing similar challenges” (Hrywna, 2009).

The truth is it is not so simple. In response to President Obama’s recommendation to put

federal funding behind a single program (presumably, in the First Lady's words, the one program deemed "the most effective") to help mothers with young children, four experts published a letter to the White House arguing for a more flexible approach. The authors are Deborah Daro, research fellow at University of Chicago's Chapin Hall; Ken Dodge, Duke University professor of public policy; Heather Weiss, founder of the Harvard Family Research Project; and Edward Zigler, emeritus professor of psychology at Yale.

In truth, there are no universal "what works" answers, at least when the programs involve human beings.

They are evaluators focused on the impact of government and foundation interventions. These experts are not soft touches. Their public position serves as proof that acknowledgment of complex systems does not equate to a post hoc, wait-and-see approach ready to declare whatever happens as a relevant outcome. The experts here argue for an approach that they believe would touch more children and families with more good outcomes:

The President's decision to invest in home visitation for newborns represents an important public policy choice and one we fully support and applaud. The proposed strategy for building this new system, however, gives us cause for concern. We do not believe the current structure of this policy will achieve maximum impacts and benefits for the next generation of young Americans (Daro et al., 2009).

The quartet's central objection is one that needs to be raised across government policy and philanthropy: That there are no one-size-fits-all solutions to complex social issues. The accountability movement's inclination to codify "what works" and then require the approved protocol to be applied universally has led to the misuse of results of randomized trials and resulted in the failure to maximize impacts.

In truth, there are no universal "what works" answers, at least when the programs involve human beings. Instead, sophisticated experimental designs yield evidence on what tends to work for whom, for how long, under what circumstances, and so on. Every first-year methods student knows this. Yet these qualifications are dropped when accountability champions, including the White House, look to use evaluation results to select a single "what works" program to implement across the board. Those seduced by the apparent simplicity of the accountability movement somehow dismiss Campbell's (Cook and Campbell, 1979) questions of external validity in their entirety.

Here is how the Daro-Dodge-Weiss-Zigler letter explains problem:

An underlying requirement of the President's home visitation initiative is that it contains only models that have been proven effective through randomized clinical trials. We do not disagree with the unique role such trials can play in identifying promising interventions and new policy opportunities. Randomized control trials offer some assurance that investments made in a given policy direction will have desired impacts on program participants and generate short- and long-term savings. However, knowing that a program is capable of achieving effects under ideal conditions is not the same as knowing it will achieve effects when broadly implemented with more challenged populations or in more poorly resourced communities. In the real world, the success of a home visitation program will depend on how local parents from all points on the risk continuum view early intervention services, what service and provider characteristics will attract new parents into these programs, and the relation between these efforts and other elements within a community's existing service continuum.

In many respects the core features of a well-done randomized trial—a highly specified intervention, consistent implementation, and a specific target population—limit the ability to generalize its findings to diverse populations and diverse contexts. Contrary to what some have suggested, we know of no evaluation scholar who would conclude that randomized

trials provide sufficient insight for making program replication decisions.

The single model reflected in the President's proposal (Nurse Family Partnership) is an outstanding home visitation program and has been proven effective in achieving several outcomes suggestive of potential cost savings. These outcomes, however, have been largely limited to young women who enrolled in services early in their first pregnancy. Although some of the participants present significant risks that work against good parenting practices including young maternal age, low income, single parent status, and a range of mental health challenges, as a group they do not fully represent populations often found to dominate public welfare caseloads nor those reported for maltreatment, areas where the current proposal hopes to achieve significant long-term savings. Based on the 2006 birth data available from CDC, a unique focus on first-time parents would leave about 62 percent of newborns ineligible for service (about 2.7 million births). Further, infants in the foster care system are eight times more likely than other infants to have mothers who received no prenatal care—a reality that would have precluded these women from accessing Nurse Family Partnership.

Fortunately, other well-researched home-based interventions have been found effective in achieving positive outcomes with these more troubling populations, and they have accomplished their objectives without employing nurses. The concept of early home visitation is being taken to scale in many communities across the country, some of which have invested in a model similar to that proposed by the President while some follow other models such as Parents as Teachers, Healthy Families America, Home Instruction for Parents of Preschool Youngsters, Parent-Child Home program and Early Head Start. These efforts have been studied by multiple investigators, and the findings have been used to improve program structure, staff training, and outcome documentation. In developing the parameters for the President's proposal, we would encourage you to give careful attention to this full body of research.

Another telling example comes from Dr. Jerome Groopman and Dr. Pamela Hartzband (2009), both Harvard Medical School faculty and on

staff at Beth Israel Deaconess Medical Center in Boston. They describe the misuse and overuse of bright-line accountability practices. As they explain in the *Wall Street Journal*, tools suited to minimum standards do not work when applied to complex medical situations where every patient needs to be considered individually.

At the outset, the doctors acknowledge the positive power of accountability metrics. Recent federal projects designed to improve Medicare quality have had positive effects. Imagining a hospital “as a large factory where systems needed to be standardized to prevent avoidable errors,” new protocols and guidelines regarding hand washing have reduced patient infections.

As part of the regulatory oversight, doctors who not comply with the new guidelines draw negative public recognition and suffer financially. This is an important and relentless force within the accountability movement: Not following the plan is what is punished.

However, what worked with hand washing may not work so clearly with more complex clinical care decisions. Groopman and Hartzband (2009) explain that regulators reviewed competing approaches to medical care and decided to mandate that hospitals keep blood-sugar levels within a very narrow range among ICU patients. This became a metric on report cards that judge performance of hospitals and doctors—and thus one criterion by which punishments and rewards were meted out. Keeping the blood-sugar levels within the narrow range would be rewarded. Varying the care would draw serious punishment.

Yet now studies show more patients are dying when their blood-sugar levels are tightly controlled. Patients receiving more flexible care do better. The “what works,” it turns out, depends on the conditions of the patient. Cookie-cutter accountability solutions that work well on minimum standards, such as with hand washing, are not robust enough to work with more complex issues.

The case shows “why rigid and punitive rules to broadly standardize care for all patients often

break down,” Groopman and Hartzband write, pointing to Orwell and Kafka as models for today’s accountability regulators:

Human beings are not uniform in their biology. A disease with many effects on multiple organs, like diabetes, acts differently in different people. Medicine is an imperfect science and its study is also imperfect. Information evolves and changes. Rather than rigidity, flexibility is appropriate in applying evidence from clinical trials. ... And what is best sometimes deviates from the norms. Yet too often quality metrics coerce doctors into rigid and ill-advised procedures.

Differentiating Between Simple, Complicated, and Complex Systems

Foundations and nonprofits are continually urged to adopt leading practices from the business sector. So philanthropy should take note when *Harvard Business Review* publishes an article (Hall and Johnson, 2009) asserting that “the movement to standardize processes has gone overboard.”

Joseph Hall and Eric Johnson, faculty members at Dartmouth’s Tuck School of Business, have figured out what many foundation and nonprofit staff have felt in their gut: In many settings flexible processes yield better results than standardized processes. Moreover, “process standardization can undermine the very performance it’s meant to optimize.”

This runs counter to more than 100 years of practice, largely driven by the Industrial Revolution’s launch of modern manufacturing technology. As Hall and Johnson report, by tradition “management calls for blindly reducing variability.” Six Sigma programs and management training worldwide teach and reward process standardization. Yet the authors contend that individuals should be able to recognize when there is volatility in the environment and volatility with inputs or outputs (or both). In these cases, processes require more “art” than science.

Consider the authors’ story about Ritz-Carlton executives, who traditionally required staff to follow precisely a 20-point checklist on customer

service. That is classic accountability, with the focus on compliance—following the checklist—and not the ultimate satisfaction of customers. In 2006 Ritz-Carlton executives figured out this rigid approach was no longer sufficient. The nature of their guests, and thus their guest needs, was varied. Expanding the customer-service script was deemed impractical. In its place, staff was encouraged to “use their judgment and improvise,” basing their service on a unifying set of corporate values.

Instead of rewarding employees for following a predetermined script (“always carry a guest’s luggage”), the new system rewarded employees for their ability to figure out, understand, and respond to guests’ specific needs. In the end, Ritz-Carlton guests gave higher marks for customer service.

Hall and Johnson zero in on exactly what is required to move beyond the accountability movement’s minimums to take on more ambitious social change efforts: recognizing that different types of systems require different approaches.

The traditional simple, linear system is based on a manufacturing model. If we change a gear or alter the speed of an assembly line, it is easy to predict how these changes will ripple throughout the entire system. And if we run widgets through the same system under the same conditions, we are pretty certain of observing the same effects on all of the widgets. Obsessive adherence to process standardization is how companies such as Coca-Cola can efficiently produce millions of bottles and cans of soda every day, with no unintended variation.

Trouble arises when we transfer this approach to systems involving human beings, each of whom is different. As obvious as this may seem, philanthropy has struggled to move beyond simple, linear logic models that assume social programs are driven by precise gears and belts and that individuals running programs and receiving interventions react as predictably as aluminum or steel.

In their 2006 book *Getting to Maybe*, Westley, Zimmerman, and Patton (2006) force us to con-

TABLE 1 Simple, Complicated and Complex Problems

Following a Recipe	Sending a Rocket to the Moon	Raising a Child
The recipe is essential	Formulas are critical and necessary	Formulas have a limited application
Recipes are tested to ensure easy replication	Sending one rocket increases assurance that the next will be OK	Raising one child provides experience but no assurance of success with the next
No particular expertise is required, but cooking expertise increases success rate	High levels of expertise in a variety of fields are necessary for success	Expertise can contribute but is neither necessary nor sufficient to ensure success
Recipes produce standardized products	Rockets are similar in critical ways	Every child is unique and must be understood as an individual
The best recipes give good results every time	There is a high degree of certainty of outcome	Uncertainty of outcome remains
Notes the quantity and nature of "parts" needed	Separate into parts and then reassemble and coordinate	Cannot separate the parts from the whole; essence exists "in the between" in the relationship between elements

front this distinction by articulating the characteristics of simple, complicated and complex systems (see Table 1).

The key here is the following: The more complex the system, the less likely we can simply follow a plan and expect any certainty in the outcome. The more complex the system, the more important it is to identify the right feedback sources and the optimal frequency and nature of feedback.

It challenges our tradition of problem solving to consider that success could ultimately depend less on following a plan and more on tracking and reacting to the right feedback systems. Traditional problem-solving models tend to start with the task of defining the problem and then hinge on implementing the agreed-upon course of action. Block (2008) explains that these models do not easily accommodate ongoing feedback or changes in the system during implementation.

To move beyond this, we should reconsider even our everyday understanding of feedback. We tend to imagine positive feedback as praise, negative feedback as a scold. With a systems approach, however, feedback is the mechanism by

which we affirm a decision or make a course correction. Negative feedback becomes an essential tool in our effort to increase the likelihood of a positive outcome. Negative feedback is neither summative nor punitive. It is only an indication that a change is needed. In the most basic example, consider a self-organizing, self-correcting system, such as your home thermostat. This system is based on negative feedback. The thermostat is set to react to extra warm air by making things cooler. It reacts to extra cool air by making things warmer.

To recognize the importance of feedback, we first must be able to recognize the nature of the system in which we are working. The key is to differentiate between simple and complex systems, as the *Getting to Maybe* team urges. *New York Times* columnist David Brooks, commenting on findings from a longitudinal study tracking two groups of men for nearly 70 years, recognizes this distinction. In describing how the study explains variations in the lives of the men, Brooks marveled at the "stream of suggestive correlations" from variables linked to health, happiness, and general well-being. But more interesting to Brooks was what the research *could not simplify*: "It's the

baffling variety of their lives that strikes one the most. ... There is a *complexity* to human affairs before which science and analysis simple stands mute” (emphasis added).

In much the same way, in his 2007 commencement speech at Harvard, Bill Gates (2007) was explicit about the need to adjust our approaches to suit the complexity of the issues before us:

All of us here in this Yard, at one time or another, have seen human tragedies that broke our hearts, and yet we did nothing—not because we didn’t care, but because we didn’t know what to do. If we had known how to help, we would have acted.

The barrier to change is not too little caring; it is too much complexity.

To turn caring into action, we need to see a problem, see a solution, and see the impact. But complexity blocks all three steps.

This central argument—that “the barrier to change is not too little caring; it is too much complexity”—represents a shift in American thought.

A generation ago, Martin Luther King, Jr.’s (1966) observation about our response to global population growth is notable for the differing point of view. The problem “is soluble by means we have discovered and with resources we possess. What is lacking is not sufficient knowledge of the solution but universal consciousness of the gravity of the problem and education of the billions who are its victims.” This shift is critical to understanding the value of the accountability movement, then and now.

In situations where solutions are evident and resources are at hand, where there is sufficient knowledge, then bright-line accountability tools are highly useful. When all that is required is caring or consciousness, it is inexcusable not to implement the remedy. This minimum standard of performance can be enforced through accountability. Failing to meet the standard should draw punishment. As an example, all that was required

to decrease hospital infections was to mandate hand washing.

Now, today, Gates is encountering a set of global problems where the situation has flipped. No longer is caring enough. Solutions are not evident. As Gates and other leaders shift their attention to these types of complex social dilemmas, the accountability tools suited to problem solving a generation ago are less useful. Success cannot be realized through following a plan or implementing the agreed-upon set of activities. Uncertainty cannot be eliminated. Success will depend on identifying and reacting to the most powerful feedback variables — and making adjustments at each decision point that increase the chances of contributing to some social change.

Education: A Case Study in Evaluation, Accountability, and Social Change

Education, especially high school reform, is one of the complex issues in the sight lines of the Gates Foundation.

The challenge before the country—educating millions of school children so that all reach their potential—is more like our messy game of pick-up sticks than toppling a chain of prearranged dominoes. It is a volatile environment, where everything is changing at once. Individuals themselves arrive at school with diverse needs. Individual students react differently to different stimuli over time. Teachers, principals, and parents are in continual states of change. Students are constantly subjected to changing outside influences.

Despite these multiple moving targets, education has traditionally relied on an accountability approach focused on minimum standards and punishments for failing to clear the bar. High-stakes tests tend to be set up as pass-fail markers for kids and schools. In some cases, students who do not meet the minimum standard do not pass a grade or graduate. With other tests, a predetermined share of student test takers must meet a minimum score or else teachers, or perhaps the entire school, suffer some punishment.

Most notable is that all of these evaluation tools provide feedback after the intervention has come to a close. In short, the evaluation results arrive at a moment when it is too late to make any improvements, to make any changes. This is the case with end-of-year report cards. Once schools or students receive their final report cards, there is very little anyone can do to act on the information. As a result, education's accountability approach does a terrific job of producing snapshots of low performers, and then punishing them. Yet the approach contributes little or nothing to bringing about large-scale social change that would alter the course for these students and schools.

In education, the accountability movement has not produced social change. The deficiencies that launched recent education reform models, and the prior reform models, remain.

What we are left with are stories like the ones Jay Matthews (2009) tells. Matthews is the education reporter for the *Washington Post* and author of a 2009 book on the KIPP charter school model and its founders, David Levin and Mike Feinberg.

As new college graduates who entered the classroom through Teach for America and not through a traditional university education program, Levin and Feinberg were essentially insurgents challenging the existing bureaucracy—and thus challengers to the process standards of the accountability system. Levin and Feinberg were focused on authentic outcomes. For them, this meant helping low-income elementary school students in Texas learn and achieve so they could improve their lives. This meant the two teachers were highly flexible in what they would do to bring about this result. In fact, they were willing to try most anything — and did — to help their students.

Their commitment to social change recalls the value articulated by Franklin Roosevelt (1932), who in a college commencement address made the case for action, however flawed: “The country needs and, unless I mistake its temper, the coun-

try demands bold, persistent experimentation. It is common sense to take a method and try it. If it fails, admit it frankly and try another. But above all, try something.”

Social change efforts, in contrast with the accountability movement, greatly favor trying something, anything, to disrupt the status quo. Too often accountability models end up reinforcing the status quo.

We see the same admonition from evaluator Donald Campbell, whose essay “The Experimenting Society” (Campbell and Russo, 1999) makes the case for what he calls action research: “Faced with a choice between innovating a new program or commissioning a thorough study of the problem as a prelude to action, the bias would be toward innovating.” Leaping in to act, of course, raised the importance of feedback, which Campbell recognized through his call for “trial-and-error.”

Social change efforts, in contrast with the accountability movement, greatly favor trying something, anything, to disrupt the status quo. Too often accountability models end up reinforcing the status quo.

Accountability rewards following the plan or script. If the outcomes are not realized, it is not the fault of the individuals following the plan. They followed instructions. Individuals are accountable for following the rules established by whatever authority is managing the accountability system. Thus, Levin was poorly received because he would not agree to teach the same way all the others were teaching. Levin found some faculty “were openly hostile to him,” and once in the school parking lot his tires were slashed (Matthews, 2009).

At the time in Texas (1994), the driving force in the accountability movement was the Texas Assessment of Academic Skills test. The accountability movement set a minimum standard: at each school, a predetermined percentage of students from every racial-ethnic subgroup must pass the test. If not, the school would receive a negative rating, which would reflect poorly on the principal. Levin had 11 low-performing students in his class that were of particular concern to his principal.

As Matthews tells the story: “The best way to avoid an unsatisfactory rating,” Levin was told by school staff, “was to exempt those 11 students from taking the test”:

This could be done if their teacher or their parents signed a statement saying that their language skills were not adequate to take the test or that they had learning disabilities that would make it unfair to judge their progress by that exam. Levin was told to fill out and sign the exemption forms. Other teachers were doing the same. ...

Levin refused to sign. ... He wanted his students to qualify for one of the magnet middle schools, which he hoped would challenge them in the same way that he had been challenging them. Some of the magnet middle schools started in the fifth grade. If his fourth graders took the TAAS exam and received good scores, they had a chance to move to the magnet right away.

Levin persuaded the parents of his students that the children were ready to perform. So when school officials bypassed Levin and asked the parents to sign the form required to keep the kids from participating in the TAAS exam, the parents refused.

The results showed Levin had been correct to focus on preparing his students, not on finagling a way to avoid putting the school at risk of not meeting the state standard. Levin wanted more than loopholes to ensure minimum compliance. He wanted social change. In the end, his students performed wonderfully on the test, which kept the school off the watch list.

The reaction of the principal? He fired Levin. He did so by making a big show of the firing during class one day. The explanation: insubordination.

Insubordination may be the iconic symbol of the accountability movement, as it articulates a punishment for not following the process standard even when, in Levin’s case, an alternative approach yields the best result.

Although Levin’s response put him in the minority of professionals willing to challenge directly the accountability bureaucracy, this characteristic is a key indicator of an individual committed to bringing about change. Foundations interested in social change must invest in these types of individuals, these “transformed nonconformists.”

At the opening plenary of the 2009 annual conference of the Council on Foundations, Morehouse College President Dr. Robert Franklin (2009) reminded several hundred foundation staff in attendance that philanthropy cannot advance social change by supporting current practices. Drawing on quotes from Dr. King, he said: “This hour in history needs a dedicated circle of transformed nonconformists. The saving of our world from pending doom will come, not from the action of a conforming majority, but from the creative maladjustment of a transformed minority. Dr. King maintained there was a difference between the social consequences of conforming majorities and the creative and redemptive maladjustment of the few who see possibilities in every crisis.”

In the accountability movement, the “conforming majorities” are those who put process compliance ahead of social change. The idea of a “transformed nonconformist” defining success as compliance with accountability rules tilts toward oxymoronic. The accountability movement, almost by definition, depends on an “conforming majority” unwilling to question the orthodoxy.

Education Innovation: Evaluation for Social Change

In education, blind compliance to process has resulted in numerous examples of cheating, fraud,

and data manipulation (Popham, 2006), always in an attempt to avoid the punishments built into accountability systems as the response for not meeting a minimum standard.³ One of the most recent examples comes from Georgia (Judd 2009), where administrators at five schools are under investigating for changing students' answers on tests to improve overall school performance. Another high-profile case has surfaced in Chicago, where high schools were found to be counting students as present, day after day, long after the students had dropped out.

The reason for the cheating? The accountability model punishes schools for truancies and absences. The *Chicago Tribune* (Ahmed, 2009) reports that city schools already lose \$18 to \$20 million each year because students are missing days. As with most accountability models, the system is set up to report year-end statistics, which affect annual funding decisions, but it is not set up to improve or change the situation. The accountability system is intended to count students who drop out, not help teachers and schools prevent students from dropping out. "When you have that kind of pressure, human beings tend to cheat," says William Gerstein, a high school principal. "There is a lot of incentive ... to report students present on days they missed."

Unlike a thermostat, which responds to heat or cold by providing resources to change the status quo once the current situation becomes unacceptable, accountability systems simply identify and punish performance outside the desired range. Further, the model focuses only on a minimum standard: getting kids in the building long enough to count them as present. The model does not support a strategy to change the lives of the students.

³ The Houston scandal over fraudulent dropout data was one of the most high-profile examples of the last decade. For details, visit *60 Minutes II* story from Aug. 25, 2004, "The 'Texas Miracle,'" <http://www.cbsnews.com/stories/2004/01/06/60II/main591676.shtml>, or NPR's March 25, 2009, story, "Scandal Surrounds Houston's High School Dropout Rate," <http://www.npr.org/templates/story/story.php?storyId=4565125>.

Now consider a social change approach, one that uses evaluation information in entirely different ways. In Montgomery County, Maryland, public school officials have adopted an evaluation approach to maximize social change. The approach acknowledges education is a complex system requiring robust feedback mechanisms during the period of the intervention. This is because the intervention itself may need to be modified while is unfolding.

Unlike a thermostat, which responds to heat or cold by providing resources to change the status quo once the current situation becomes unacceptable, accountability systems simply identify and punish performance outside the desired range.

Instead of waiting until the end of the school year to report on a pass-fail statistic—at a time when it is too late to change the course for the student—Montgomery County is investing nearly \$50 million a year on Edline, a data system that coordinates input from multiple sources and sends real-time reports to parents and other stakeholders to quickly match kids to interventions.

In describing the system, the *Wall Street Journal* (Hechinger, 2009) tells the story of Duane Wilson, a high school senior in Bethesda. During the first semester of his senior year, Duane was receiving low marks in several advanced placement courses. With the Edline system, as soon as the teacher recorded one of Duane's low grades, the data system automatically sent an e-mail message to Duane's mother. This was enough for Duane's mother to rally her son. Duane picked up his grades and gained acceptance to college.

“Like a smoke detector sensing fire, a school computer sounded an alarm,” the *Journal* article explains. More than three dozen staffers continually look for emerging risks across the student population. “Alerts of flagging performance come from Edline and another data-driven tracking system modeled after the one used by the New York City Police. The warnings, often sent via email, can spark immediate action, such as after-school tutoring, study sessions and meetings with families.”

These tools put a maximum on the hard thinking that goes along with identifying flaws and hazards in the proposed strategy, as well as the willingness to set up contingency plans—triggered by tracking a set of variables that can serve as warning bells—and use them as conditions warrant.

Unlike the traditional accountability models, the system in Montgomery County does not punish early signs of low performance. The system reacts to these signals by supplying additional resources in an attempt to change the system, to reverse the current direction. This type of real-time feedback is already the underpinning of logistics-sector leaders FedEx, UPS, and WalMart. Perhaps the closest parallel to public schools, in fact, is the Harrah’s casino chain (Levinson, 2001), which several years ago adopted a technology solution to providing sweeping, real-time feedback about every individual customer engaged.

For the Montgomery schools, over the course of an academic year there is no way to predict which students will encounter which problems at any particular moment, so education officials

constantly monitor all key variables in the way that meteorologists monitor variables on a hurricane. They do not know which intervention to implement, for whom or when, until the real-time feedback systems yield the relevant information.

Eight Tools for Advancing Social Change

There is much to be said for the notion that success hinges on fitting the right tool to the right job. That applies here. Much of the frustration foundations experience with evaluation efforts can be attributed to the wrong tools.

One of the most common complaints, that the evaluation reports arrive long after an initiative has ended and too late to be actionable, goes to the heart of the need to produce rolling, real-time updates. A commitment to this type of feedback is the only way to manage work occurring within complex systems. Likewise, foundations seeking to launch ambitious social change agendas are frustrated when they are repeatedly confronted with simple, linear logic models. Shoehorning the information into the boxes on the page satisfies a compliance step, but filling out the document does nothing to strengthen the underlying logic.

Assuming a static environment, these models also assume the foundation is the only player moving pieces on the game board. Thus, the models hinge more on an ability to predict the future than on ability to receive and act on input in real time. Anything that deviates from the predetermined critical path—a rifle shot—essentially eliminates any chance of success.

Below is a list of eight tools that enable foundations to work in a very different way. These tools put a maximum on the hard thinking that goes along with identifying flaws and hazards in the proposed strategy, as well as the willingness to set up contingency plans—triggered by tracking a set of variables that can serve as warning bells—and use them as conditions warrant. These tools allow foundations to hedge, essentially creating multiple opportunities for success.

The list is in no way exhaustive. There are numerous tools and services that support social change

agendas. To identify these resources, foundations should be aggressive in looking for practices outside the narrow field of nonprofit evaluation.

1. Risk Analysis

The basic questions in risk analysis are the following: (1) What is most likely to go wrong? (2) For each potential hazard, what are the potential consequences? (3) Considering both the probability of occurrence and the potential harm, what are the highest priority risks to worry about? (4) If these risks materialize, what will we do in response?

Working through these questions, with foundation staff and grantees, allows for an honest conversation about the greatest threats to a strategy. It is not acceptable to say that nothing will go wrong. Things always go wrong. Always. Success hinges on an ability to track variables that will sound alarms—and agreeing on threshold levels that will trigger the alarm—as risks emerge, enabling foundations to activate contingency plans or alternative approaches waiting in the wings (Bare, 2001).

Foundation and nonprofit staff are rarely asked to be up front and candid about what is most likely to go wrong. Making risk analysis a central part of the work of philanthropy is the quickest way to move from an accountability model to a social change model.

2. Systems Approaches

Another quick way to move from an accountability model to a social change model is to use the “Getting to Maybe” tool (shown in Figure 2) to differentiate the system in which foundations and nonprofits are working.

To fit the right tool to the right situation, it is necessary to agree on the nature of the system in question. When foundations and nonprofits are implementing projects in a simple system, the use of simple, linear accountability tools is the right move.

When working in complicated or complex systems, however, foundations and nonprofits must

select their tools accordingly. Before selecting evaluation and management tools to deploy in a given setting, determining the nature of the system is the correct first step.

3. Testing Assumptions

Just like everyone tells the dentist they brush and floss regularly, every foundation and nonprofit staffer will report to an evaluator that they have considered the assumptions on which their strategy is based.

In most cases this consideration has been carried out superficially, if at all. Assumptions are our blind spots. By definition, we cannot see our blind spots. It requires extraordinary effort, and challenging input from multiple and varied sources, to surface assumptions in an authentic way.

We all see the arrows between the boxes in a simple logic model. In this kind of “if, then ...” logic, each assertion within a chain of logic rests on all kinds of assumptions not depicted literally in the model. If the assumptions hold, of course, the supposed logical connection between the boxes is more likely to play out as planned.

For example, the logic model for a high school math intervention may show a simple, causal connection between the time students spend in an optional after-school tutoring and the students’ performance on the math portion of the school’s standardized tests.

This rests on any number of assumptions: that students have time to attend the tutoring session (and do not have to be present at part-time jobs); that students who sample the tutoring will like it and come back; that students who stay after school for tutoring (and miss the school bus) can get transportation home; that the tutoring curriculum lines up with the content of the standardized test; that there is a sufficient supply of capable tutors; that students most likely to perform poorly on the standardized test will elect to consume the tutoring; and so on. Perhaps the most basic assumption is that the absence of tutoring itself is the barrier to students improving their performance on the test. If underlying physi-

cal or mental illness is diminishing the students' learning experience, then the tutoring may miss the mark.

Despite the fact that all of these assumptions do not show up anywhere in the project's logic model, the foundation may want to build in real-time feedback systems to track variables associated with these questions. Even though these are indirectly related to the initiative, a failed assumption will do more to damage chances of success than anything else. Repeated, ongoing testing of these assumptions can provide the critical feedback to monitor the needs of the initiative.

4. *The Outside View*

Dan Lovallo, a professor with the Australian School of Business, and Daniel Kahneman, a psychology professor at Princeton University and Nobel laureate, published a 2003 article in the *Harvard Business Review* that described a demon for many foundations: "delusional optimism."

The authors identify a set of organizational pressures and cognitive biases that routinely lead executives to inflate their forecasts, especially with their own projects: "We overemphasize projects' potential benefits and underestimate likely costs, spinning success scenarios while ignoring the possibility of mistakes."

Because the source of the bias is embedded in the way our brains process information, this is a difficult demon to overcome. Lovallo and Kahneman (2003) deliver the remedy, what they call "The Outside View." This external view is necessary to push back against the internal view, which is the source of the delusional optimism.

Pursuing The Outside View requires us to do the following:

- Select a set of past projects to serve as the reference class.
- Assess the distribution of outcomes from projects in this reference class.
- Predict your project's positioning the distribution.
- Assess the reliability of your past predictions.

- Correct your intuitive estimate based on your track record of forecasting.

5. *Sensemaking*

Karl Weick defines "sensemaking" as "the transformation of raw experience into intelligible world views":

What I've repeatedly noticed is that the people who really get in trouble during these crises are those who try to think everything through before taking any action. The problem with defining and redefining your hypotheses without testing them is that the world keeps changing, and your analyses get further and further behind. ...

Action, tempered by reflection, is the critical component. ... Once you start to act, you can flesh out your interpretations and rework them. But it's the action itself that gets you moving again. That's why I advise leaders to leap in order to look, or to leap while looking. ... In crises especially, leaders have to act in order to think—and not the other way around. ...

I usually urge executives to fight their tendency to want to plan everything. Most plans are too specific, and the details create the illusion that the plan grasps everything that is going on and therefore can be trusted. As a result, when you have a plan, you tend not to look for things that disconfirm it. ... The worst aspect of plans is that they heighten the tendency to postpone action when something unexpected happens (Coutu, 2003).

As would be expected, there is a precise list of steps to take to undertake sensemaking. In general, Weick encourages executives to adopt the practices of "high-reliability organizations," or HROs, which have to find ways to perform at high levels in volatile environments where unpredictable events make most planning processes useless. These practices include a constant focus on signs of potential failure, even signs staff consider as weak. Another practice: refusing to try to simplify complex situations.

6. *Game Theory*

Foundations are in the decision-making business. Game theory is the science of strategic decision

making. It follows that game theory is a useful discipline for foundation professionals seeking to increase the impact of their work.

A decision tree is a basic tool in game theory. In many situations, foundation staff can use a decision tree to depict all the possible decisions, considering the moves the foundation might make as well as moves others may make. This step alone is useful in identifying all of the multiple routes to success. While the foundation may be heavily invested in a single critical path, a decision tree can make explicit all of the decision points along the way. This helps illustrate all the courses of action that could help, or hurt, the foundation's chances of success.

In a more sophisticated approach, the foundation may assign probabilities to each branch of the decision tree, as the model depicts future occurrences. This effort not only helps users identify contingency plans and alternative courses of action, it helps users prioritize needs and resources.

7. Scenario Planning

Scenario planning is perhaps the polar opposite of linear, logic models where success relies on future events playing out exactly as planned.

Scenario planning requires participants to imagine and construct a variety of possible futures. This is not an effort to predict which future will actually materialize. It is not a forecasting competition. Scenario planning requires an explicit acknowledgment that the only thing certain about the future is that it will be different from today.

Scenario planning forces us to imagine what resources and strategies we may need to increase our chances of success in a future that will be something other than what we are experiencing today. Whereas traditional planning processes have users focusing on short-term variables within their control, scenario planning shifts the focus to macro-level variables outside of users' control. Once users establish a number of variations on what the future may look like at a future date—say, 10 or 15 years out—the question becomes:

What do we need to do to be prepared to succeed in these different conditions?

The energy company Royal Dutch Shell has been a high-profile user of scenarios. On the firm's Web site, the company publishes narratives and video versions of energy scenarios for the year 2050 (http://www.shell.com/home/content/aboutshell/our_strategy/shell_global_scenarios/dir_global_scenarios_07112006.html). In one scenario, named "Scramble," the firm imagines a rapid expansion of the coal industry as energy needs explode globally without a coordinated response by industrialized countries. With nations pursuing their own independent approaches to developing energy resources, global climate change needs are ignored. By 2020, growing transportation needs drive up interest in biofuels, which leads to a sharp rise in food prices. By 2030, coal production has reached its limit globally. By 2040, nations turn to nuclear power, but the technology does not deliver as expected. By 2050, consumers finally demand serious energy efficiency measures.

In another scenario, named "Blueprints," the firm imagines a future where consumers and advocacy groups organize to pressure governments to pursue alternative energy strategies. By 2020, this slows demand for coal and pushes governments to establishing and exploiting new, market-based strategies to take alternative energy industries to scale. By 2030, electric vehicles enter the mass market. By 2050, the world is using about 26 percent less energy than if it had stayed with the old course.

Again, these are not forecasts. These are imagined versions of how the world may unfold between now and 2050. Whether the future materializes in line with either of these scenarios is not the point. It is not a contest based on forecasting accuracy. Instead, the exercise challenges executives to imagine how dramatically their firms must change to succeed in the decades to come. Again, the only certainty is that the future will not be like today. What is required to succeed today is not sufficient to succeed in the future. The status quo will not hold. Here is how Shell describes its use of scenario planning:

Shell uses scenarios to explore the future. Our scenarios are not mechanical forecasts. They recognise that people hold beliefs and make choices that can lead down different paths. They reveal different possible futures that are plausible and challenging. Our latest energy scenarios look at the world in the next half century, linking the uncertainties we hold about the future to the decisions we must make today.

8. *Documentary Methods*

Through photography, written narratives, audio, or video, the documentary method serves as a powerful way to explain complex systems that do not lend themselves to simple reports. Further, the documentary products themselves provide feedback in extraordinary ways. The information carries great affect and influence, much more so than a written report.

Further, because documentary products often receive wide circulation, the delivery method ensures that individuals who object to the information cannot ignore it or keep it secret. The distribution systems help ensure documentary products are broadly accessible. This empowers consumers to invent and reinvent users for the products. In effect, documentary products set things in motion.

StoryCorps is a prominent producer of radio documentaries today. StoryCorps (<http://www.storycorps.org/about>) specializes in harvesting oral histories from everyday people, and the organization trains users to collect the stories. Through StoryBooths and MobileBooths, the organization has created fixed sites for individuals to use around the country, as well as mobile resources that travel from city to city collecting stories. Within this general mission, there are specific initiatives such as StoryCorps Griot, dedicated to preserving the stories of African Americans. The stories are archived with the National Museum of African American History & Culture.

In recent years, documentarians Michael Moore and Ken Burns have helped raise the profile of the medium. Moore's documentaries draw audiences to movie theaters. By shining light on issues other

people often would rather not discuss, his projects have affected the national political debates on gun control, health care, and economic inequities. The Burns documentaries, which are distributed primarily through public television, have helped shape the cultural conversations on the Civil War, baseball, and jazz.

When former eBay president Jeff Skoll launched his philanthropic efforts, he started a firm, Participant Media, to advance his social change agenda. Through feature films, documentaries and other media products, he is able to engage millions of users and advance complex ideas. Other methods are not robust enough to reach such a large audience with such ambitious ideas.

Participant Media has managed several high-profile media projects to completion, including *Food Inc.* and *An Inconvenient Truth*. Other efforts, such as a documentary about the effect of AIDS on a South African village, or *Darfur Now*, which examines genocide, are included in the firm's push to move individuals to "social action" (http://www.participantmedia.com/social_action.php).

The firm's Web site profiles a long list of projects. For each, the "social action" content provides a list of "5 Things You Can Do Now." Here is how Participant Media describes its mission:

The company seeks to entertain audiences first, then to invite them to participate in making a difference. To facilitate this, Participant creates specific social action campaigns for each film and documentary designed to give a voice to issues that resonate in the films. Participant teams with social sector organizations, non-profits and corporations who are committed to creating an open forum for discussion, education and who can, with Participant, offer specific ways for audience members to get involved. These include action kits, screening programs, educational curriculums and classes, house parties, seminars, panels and other activities and are ongoing "legacy" programs that are updated and revised to continue beyond the film's domestic and international theatrical, DVD and television windows. To date, Participant has developed active, working relationships with 83 non-profits reaching over 20 million people.

Participant Media's approach blurs the line between for-profit Hollywood projects and do-gooder nonprofit efforts. It blurs the line between entertainment and documentary. As would be expected, it does not concern itself with accountability questions of whether its offerings fit neatly into any traditional category. It is much more concerned with authentic social change than about short-term, tactical process metrics. In the end, Participant Media's approach is very much like that of insurgent teacher David Levin. Participant Media will do whatever it takes to move individuals to action.

Conclusion: A Call to Action

For foundations and nonprofits to increase the impact of their collective investments, philanthropy's leaders must expand their imaginations and their toolkits.

In reimagining the work, philanthropy's leaders have an opportunity to take on complex social change strategies without apologizing for the lack of pound-your-fist-on-the-table certainty. The field can leave that kind of certainty to the accountability movement's compliance tools, which use bright-line measures to describe dollars invested, dollars spent, the number of units of services delivered, and the efficiency of implementing inputs. It is important to monitor adherence to minimums. It is just not enough.

Reimagining the work requires something else: It means acknowledging that philanthropy cannot deconstruct complex problems into a series of small, linear projects and expect the sum of the parts to add up to large-scale social change. Seeking to extend basic human rights to more individuals around the world, seeking to reduce racism in a given city, or seeking to change public-health norms in small town — all of these aspirations require first a willingness to take on challenges that defy short-term, causal, quantifiable results attributable to a best practice. It is back to admitting foundations are playing a game of pick-up sticks, where everyone in the game is moving their sticks at the same time. In volatile settings, with short time horizons, linear plans that connect inputs to outputs lose value as the underlying assumptions

and conditions change. Foundations cannot generate outcomes by easily setting off a string of chain reactions, like toppling dominoes.

Willing to reimagine the work in this way, foundation and nonprofit executives will discover a new set of approaches and tools to support them through the experience of "getting to maybe." The tools described here, and other tools that may be adapted to complex settings, breathe new energy into the work of philanthropy. Just as the imagination can redefine the work, new tools enable philanthropy's leaders to pursue, manage, and evaluate their efforts in powerful ways.

Reimagining the work requires something else: It means acknowledging that philanthropy cannot deconstruct complex problems into a series of small, linear projects and expect the sum of the parts to add up to large-scale social change.

The revelation, across all of the tools, is that foundation and nonprofit executives are freed from making precise predictions about how the future will play out, with success hinging on forecasting exactly what future will unfold. Instead, emphasis is placed on tools that inform and support decisions, at each step along the way, that create the greatest opportunities for impact across many of the possible futures. These tools assume a fixed and unwavering commitment to the outcome and a readiness to adapt and adjust in terms of what it will take to produce the desired impact.

References

AHMED, A. (2009). *In Chicago high school: class attendance or numbers game?* Chicago: *Chicago Tribune*.

- ASHKENAS, R. (2007). *Simplicity-minded management: a practical guide to stripping complexity out of your organization*. Cambridge: Harvard Business Review.
- BARE, J. (2002). *Risk*. Cambridge: Evaluation Exchange.
- BLOCK, P. (2008). *Community: the structure of belonging*. San Francisco: Berrett-Koehler Publishers.
- BROOKS, D. (2009). *They had it made*. New York: *The New York Times*.
- CAMPBELL, D., RUSSO, M. (1999). *Social Experimentation*. Thousand Oaks: Sage.
- CLINTON, W. (1993). *Remarks by president Clinton announcing the initiative to streamline government*. Available at: <http://govinfo.library.unt.edu/npr/library/speeches/030393.html>.
- COOK, T., AND CAMPBELL, D. (1979). *Quasi-Experimentation: Design & Analysis for Field Settings*. Boston: Houghton Mifflin Company.
- COUTU, D. (2003). *Sense and reliability: a conversation with celebrated psychologist Karl E. Weick*. Cambridge: Harvard Business Review.
- DARO, D., DODGE, K., WEISS, H., AND ZIGLER, E. (2009). *Correspondence to President Obama*. Available at: <http://preventchildabuse.org/pdf/CommentsonHomeVisitingProposal.pdf>.
- DE VISE, D. (2007). *Montgomery schools chief says federal mandate lowering standards*. Washington, D.C.: *The Washington Post*.
- FRANKLIN, R. (2009). *Council on Foundations opening plenary speech, Atlanta*.
- GATES, B. (2009). *Harvard commencement speech*. Available at: <http://www.scribd.com/doc/97979/Bill-Gates-Harvard-Commencement-transcript>.
- GROOPMAN, J., AND HARTZMAN, P. (2009). *Why "quality" care is dangerous*. New York: *The Wall Street Journal*.
- GRUBER, A. (2003). *True confessions*. Washington, D.C.: *Government Executive*.
- HALL J., AND JOHNSON, E. (2009). *When should a process be art, not science?* Cambridge: Harvard Business Review.
- HEAD, S. (2003). *The new ruthless economy: work and power in the digital age*. Oxford: Oxford University Press.
- HECHINGER, J. (2009). *Data-driven schools see rising scores*. New York: *The Wall Street Journal*.
- HRYWNA, M. (2009). *\$50 million social innovation fund to target education, health, economy*. Morris Plains, N.J.: *The Nonprofit Times*.
- HYMOWITZ, C. (2008a). *When meeting targets becomes the strategy, CEO is on the wrong path*. New York: *The Wall Street Journal*.
- HYMOWITZ, C. (2008b). *Readers share tales of jobs where strategy became meeting target*. New York: *The Wall Street Journal*.
- JENSEN, MICHAEL C. (2003). *Paying people to lie: the truth about the budgeting process*. European Financial Management. Available at: <http://ssrn.com/abstract=423635>.
- JUDD, A. (2009). *Cheating on CRCT signals test's high stakes*. Atlanta: *Atlanta Journal-Constitution*.
- KING, M. (1966). *Acceptance speech for the Margaret Sanger Award in Human Rights*.
- LEVINSON, M. (2001). *Harrah's entertainment — jackpot! Using IT to manage customer information*. Framingham, Mass: *CIO Magazine*.
- LOVALLO, D., AND KAHNEMAN, D. (2003). *Delusions of success: how optimism undermines executives' decisions*. Cambridge: Harvard Business Review.
- MATTHEWS, J. (2009). *Work Hard. Be Nice. How Two Inspired Teachers Created the Most Promising Schools in America*. Chapel Hill: Algonquin Books.
- POPHAM, W. (2006). *Educator cheating on No Child Left Behind tests: can we stop it?* Washington, D.C.: *Education Week*.
- ROOSEVELT, F. (1932). *Oglethorpe University commencement address*.
- STEVENSON, T. (2006). *The best coaching inspires without punishing*. Raleigh, N.C.: *The News & Observer*.
- STURROCK, C. (2006). *States distort school test scores, researchers say*. San Francisco: *San Francisco Chronicle*.
- SUROWIECKI, J. (2005). *The wisdom of crowds*. New York: First Anchor Books.
- WALLACE, K. (2007). *No child left behind: state tests vary*. New York: CBS News.
- WESTLEY, F., ZIMMERMAN, B., AND PATTON, M. (2006). *Getting to Maybe*. Toronto: Random House Canada.

John Bare is the vice-president for Sports Philanthropy and Affiliated Funds at The Arthur M. Blank Family Foundation. He can be reached at the Arthur M. Blank Family Foundation, 3223 Howell Mill Road, NW, Atlanta, GA 30327.