

Daniel and the rhinoceros

David E.K. Hunter *

Director of Evaluation, The Edna McConnell Clark Foundation, 415 Madison Avenue, 10th Floor, New York, NY 10017, USA

Received 13 September 2005; received in revised form 24 October 2005; accepted 24 October 2005

Abstract

This paper examines ways in which funders often do harm in the name of good by focusing on randomized control experiments over all other evaluation methods when helping not-for-profit organizations improve the effectiveness of their programs. It offers a critique of current practice and suggests ways in which foundations might work usefully and productively with grantees on evaluation-related capacity-building. Using a biblical example of an early evaluation, it notes how even simple evaluations that fall short of meeting the criteria of the randomized experiment can be really meaningful, useful and cost-effective for both grantee organizations and funders.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Capacity-building; Randomized experiment; Evaluation methods

1. An early evaluation and what we can learn from it

Sometimes, when confronting contemporary challenges, it is useful to consider the lessons of history. Our cultural heritage includes the story of the focused, high-stakes evaluation of a dietary program undertaken some 2600 years ago in 605 BCE. That was the year in which Nebuchadnezzar, King of Babylon, conquered Jerusalem and carried off into servitude thousands of the people of Israel. It seems that Nebuchadnezzar was a progressive ruler who wanted to assimilate his captives and turn them into productive citizens (but not too productive—he had the men castrated). He therefore commanded Ashpenaz, the master of his eunuchs, to select a group of Israelite aristocrats who were free from physical blemish, skillful, wise, and educated in science, and therefore able to comport themselves appropriately and in an aesthetically pleasing manner in his palace—so that they might be taught the language, culture, and knowledge of their hosts and then serve the king as mid-level bureaucrats. To ensure the success of this undertaking, Nebuchadnezzar ordered that they receive a daily provision of the king's own meat and wine, intending thus to nourish them for 3 years so that, at the end of this period, they might appear before him healthy, well-educated, and ready to earn their keep.

Among the selected group was Daniel as well as three others—Hananiah, Mishael, and Azariah. Whereas other

Israelites agreed to Nebuchadnezzar's terms, Daniel and his friends could not bring themselves to eat what Jewish dietary laws considered to be unclean food. He therefore demurred, threatening a hunger strike, and thereby put Ashpenaz, the newly appointed Director of the Babylonian Government Acculturation and Dietary Program for Israelites, in a terrible bind. As Ashpenaz said to Daniel, "The buck stops with me and the king is one unforgiving boss...". He went on to explain that if Daniel and his companions were to go on a hunger strike and stand before Nebuchadnezzar looking sick or emaciated, "The king'll have my head on a platter".

But Daniel, being, after all, a prophet and hence presumably able to see some 2600 years into the future, contemplated the 'Government Performance and Results Act of 1993' (about which there will be more to say later) and proposed a dietary program evaluation, suggesting that he and his friends be given 10 days to eat a porridge of legumes (thus avoiding unclean meat) and drink water according to Jewish law. And he made it a high-stakes evaluation, saying in good King James argot, "Then let our countenances be looked upon before thee, and the countenance of the children that eat of the portion of the king's meat: and as thou seest, deal with thy servants". Upon hearing this plan, Ashpenaz, through his Assistant Director Melzar, rather courageously agreed.

What happened? Well, after 10 days Daniel and his three friends looked healthier and chubbier than their compatriots who ate the king's dishes. Ashpenaz and Melzar, accepting these outcome data, allowed Daniel and his friends to follow Jewish dietary law and gave them a steady diet of pulse,

* Tel.: +1 212 551 9122; fax: +1 212 421 9325.

E-mail address: dhunter@emcf.org

a leguminous porridge. And further, we can assume that Ashpenaz and Melzar did not leave matters at that, and instituted a performance-tracking system through which the ongoing health of Daniel and his friends was monitored. In any event, as we all know, they thrived, and Daniel, at least, was a star performer—a standout when compared with the other Israelites who continued to eat the king's meat.

Let's take a look at this evaluation. Even though we are dealing here with Daniel, who gets the benefit of the doubt by virtue of being blessed by God and a biblical prophet, the methodology from a research perspective seems pretty weak. First of all, not only are four subjects a ridiculously small number to use for the purpose of comparison, the four that participated in this evaluation were self-selected. Furthermore, the intervention dosage was minimal; the time trial was ludicrously short; and the indicators were ambiguous. When we also consider that the outcomes were, at best, impressionistic and the measures purely subjective, it becomes clear that this evaluation was not what we would call scientifically sound. So, any evaluation novice will recognize that this evaluation lacks credibility on virtually all fronts: it has statistically useless data because of the tiny number of participants, is rife with built-in selection bias, features a weak and ridiculously short intervention, and is woeful in its lack of precision in measurement. In short, it is a mess.

But on closer inspection, we have to acknowledge that this evaluation has some virtues, too. It assessed a clearly articulated theory of change (kosher diet leads to good health—better than the king's diet), it was participatory in that program participants had a full voice in the evaluation design (actually, they proposed it), and it was led by a program director who accepted the research methodology as dispositive even though the outcomes were qualitative in nature. When we also consider that, subsequent to the evaluation and based on its findings, the program (dietary course of action) that it assessed was implemented, that an ongoing evaluation capacity was built within the program so that it could be used for continuous performance monitoring and quality improvement, and finally that it allowed for some spectacular long-term performance outcomes (Daniel, after all, correctly prophesized the fall of Babylon to the Persians), we must acknowledge that this evaluation had some noteworthy qualities.

So, here we have an evaluation that, at first glance, is problematic at best, and useless, at worst. Nevertheless, it was designed to be useful, was inexpensive, built local evaluation capacity, supported high-stakes decision-making, was used to develop an ongoing program performance monitoring and quality management system, accomplished what it was designed to do, and consequently, in fact, was useful to all involved. That is not bad at all—and worth keeping in mind as we consider the ways in which evaluation can be meaningful and useful both to not-for-profit grantee organizations and to funders—criteria that present funder-driven evaluation practice mostly honors with their absence.

2. The 'gold standard' of program evaluation: a rhinoceros in the living room

Increasingly, foundations are accepting the importance of evaluating their grants, programs, and initiatives. They do so in order to accomplish various purposes, including: (a) learning about their undertakings and what it takes to get them done, (b) assessing what they have accomplished, (c) holding their grantees and themselves accountable, and (d) developing and disseminating knowledge to their fields of interest. Yet, in spite of these perfectly good intentions, it is not unusual for significant tensions to mar relationships not only between foundation evaluators and program officers, but, even more problematically, between funders and the grantee organizations whose programs are being evaluated.

It is this latter situation that I want to think about here—because we need to get past it in order to strengthen the effectiveness and credibility of the non-profit sector and the funders who support its organizations. The stakes are high and getting higher as the federal government continues to retrench and devolve to the not-for-profit (or so-called 'social') sector its basic safety net functions. Yet, sadly, more often than not the non-profit sector is being hurt more than helped by funders' approaches to evaluation. This happens for at least two reasons. First, some funders—with more than a bit of self-interested complicity on the part of professional evaluators—have been pushed toward a very narrow focus equating evaluation with accountability and quantified outcomes, and they are imposing this view wholesale on their grantees. Second, other funders, rejecting the procrustean reductionism of the former group, are colluding with grantees, either in rejecting the value of evaluation totally or settling for such 'soft' evaluations and/or outcome measures, that their meaning is obscure at best. Both groups are, in my view, missing the most fundamental point of all, namely, that to be really meaningful, useful and cost-effective for both grantee organizations and funders, evaluation should not be thought about, nor implemented, outside or even alongside an organization's operations. Rather, evaluation should be integrated fully into organizations' operational capacities—just as Ashpenaz and Melzar did.

Consider two important developments that—though well-intentioned—helped get us into the mess we are in at the present time.

The first was the passage of the 'Government Performance and Results Act' ([Government Accountability Office \(GAO\), 1993](#)). This was an act "(t)o provide for the establishment of strategic planning and performance measurement in the Federal Government, and for other purposes". GPRA, as the act is known colloquially, was intended to address several very real problems, including waste and inefficiency in Federal programs, and inadequate information about their performance. The purposes of the act were, among other things, to improve the confidence of the American people in the capability of the federal government, in federal program effectiveness, and in public accountability. By promoting a new focus on results, service quality, and customer satisfaction,

Download English Version:

<https://daneshyari.com/en/article/322649>

Download Persian Version:

<https://daneshyari.com/article/322649>

[Daneshyari.com](https://daneshyari.com)