Check for updates

Article

American Journal of Evaluation
2016, Vol. 37(4) 463-485
© The Author(s) 2016
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1098214015623634
aje.sagepub.com
SAGE

Data and Decision Making: Same Organization, Different Perceptions; Different Organizations, Different Perceptions

Nan L. Maxwell[1], Dana Rotz[2], and Christina Garcia[3]

Abstract

This study examines the perceptions of data-driven decision making (DDDM) activities and culture in organizations driven by a social mission. Analysis of survey information from multiple stakeholders in each of eight social enterprises highlights the wide divergence in views of DDDM. Within an organization, managerial and nonmanagerial staff working for the organization and staff from a prominent funder all expressed different perceptions of the same organization's DDDM activities and culture. Study findings also provide insights into how to improve an organization's capacity to build and use performance management systems, which include building a common understanding about what activities are—or are not—being undertaken. Finally, findings provide insights about structuring research on DDDM, which indicate that information from only one respondent in an organization or only one organization might not be reliable or generalizable.

Keywords

data-driven decision making, performance measurement, nonprofits

Using verifiable data to make decisions can be a valuable business strategy. Research suggests that data-driven decision making (DDDM) increases performance (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2010), output, and productivity (Brynjolfsson, Hitt, & Kim, 2011) in for-profit firms and the effectiveness of management decisions (LeRoux & Wright, 2010) in nonprofit organizations. Using data to make decisions effectively requires at least two key elements. First, an organization's culture must support DDDM such that staff value and embrace using data to develop and implement strategies (Preskill & Boyle, 2008; Preskill & Torres, 1999). Such a culture can influence

[1] Mathematica Policy Research, Oakland, CA, USA
[2] Mathematica Policy Research, Cambridge, MA, USA
[3] The James Irvine Foundation, San Francisco, CA, USA

Corresponding Author:
Nan L. Maxwell, Mathematica Policy Research, 505 14th Street, Suite 800, Oakland, CA 94612, USA.
Email: nmaxwell@mathematica-mpr.com

a staff's perceptions of and inclination toward using data and lead to fundamental changes in the decision-making process (Boyle, Lemaire, & Rist, 1999; Huffman, Lawrenz, Thomas, & Clarkson, 2006; Kiernan & Alter, 2004; McDonald, Rogers, & Kefford, 2003). Second, it logically requires a coordinated process of sequential DDDM activities in which an organization *collects, analyzes,* and *uses* data. Each stage is important. If data are not collected, organizations cannot analyze information to draw conclusions. If data are not analyzed consistently and correctly, staff might use it to draw incorrect conclusions.

Until recently, nonprofits had relatively few incentives to adopt an effective DDDM strategy (Nonprofit Technology Network, 2012).[1] The tide is shifting; however, as some funders now provide financial incentives or technical assistance to develop and institutionalize DDDM in the nonprofits they assist. The goal is to better position the social sector for impact and innovation beyond the limits of any one funder's engagement.[2]

This research examines DDDM in nonprofits by surveying multiple stakeholders associated with each of eight social enterprises. Social enterprises are organizations that seek to achieve both a social mission to improve their communities and a business mission of financial viability. Thus, they face both business and funder-provided incentives to ground decisions in data and provide a unique opportunity to study DDDM. Our study assesses DDDM from the perspective of organization staff in a position to use data to make decisions about the social enterprise and staff of a funder working to increase DDDM in an organization.

Our research findings highlight the large differences in the perceptions of and support for DDDM both across and within the social enterprises. Such large variation in perceptions has implications for practitioners looking to build an organization's reliance on data when making decisions and for researchers looking to expand the knowledge base on DDDM. Practitioners—including funders and executives of mission-driven organizations—who are seeking to increase DDDM might be hindered by a lack of consensus on what it means to use data to make decisions. One potential way to build consensus might be to use the survey administered for this study as a baseline assessment of stakeholders' perceptions about DDDM. For example, tabulations from the survey could be used to identify discrepancies between the beliefs and perceptions of a change agent hoping to build a learning culture and organization staff. These tabulations could also serve as a basis for discussion and building alignment about what it means to use verifiable data to make decisions.

Our findings also have implications for researchers. Researchers seeking to understand or develop principles to guide organizations trying to increase DDDM as a management strategy should be cautioned against extrapolating conclusions about data use that are drawn from research based on information from a single organization or single individual in multiple organizations. Case study research cannot account for the divergence of views about DDDM across organizations, even when organizations are seemingly similar, and single-respondent research cannot account for the divergence of views within an organization.

DDDM and Organizations With a Social Mission

Although for-profit businesses often have complex and multifaceted goals, they can only remain in business if they generate enough revenue to cover their costs. Indeed, even though for-profit firms use a variety of metrics to gauge performance, investors typically measure long-run success by the difference between revenues and costs and use this relatively easily quantified and interpreted measure of profits to capture whether the business is achieving its mission. Nonprofits have no such widely recognized, simple performance measure: A multitude of accepted generalizations about nonprofit behavior exists, each of which implies a slightly different measure of its success (Malani, Philipson, & David, 2003). As a result, it is often difficult for a funder to quantify or assess whether an individual nonprofit organization achieves its mission or the comparative impact across a grant portfolio or initiative.

Funders increasingly see DDDM as a way to make this assessment (Kaplan, 2001; Speckbacher, 2003) and have devoted resources to developing and using evidence as a basis for funding (Carman & Fredericks, 2008; Davenport, 2006).[3] Perhaps as a result, a plethora of tools have arisen that nonprofit leaders can adopt in developing and implementing performance management systems.[4] In addition, case studies have identified best practices that nonprofit managers can adopt when implementing a performance management system (Hendricks, Plantz, & Pritchard, 2008; Kaplan, 2001).[5]

Despite the availability of tools and case study–based best practices, research suggests that some nonprofits still struggle to adequately undertake the three components necessary to implement an effective performance management system: collecting, analyzing, and using data to make decisions (Carman & Fredericks, 2010). Although nonprofits often collect a proliferation of data, many lack (1) sufficient knowledge or the resources to plan and implement the technical side of data collection and management and (2) the ability to think strategically about the data they should collect (Carman, 2007; Carman & Fredericks, 2008, 2010). They often understand that data can be used to improve decision making and programs (Alaimo, 2008; Behn, 2003), better manage human resources, and achieve goals (Carman & Fredericks, 2008; Zimmerman & Stevens, 2006). But when faced with competing priorities and limited time and resources, nonprofit managers often only use data to meet and report against short-term funding requirements and do not invest in DDDM capacity over the long term (Carman, 2007; Carrilio, Packard, & Clapp, 2003).

The struggles nonprofits face in developing effective performance management systems suggest funders or executives looking to build an organization's capacity to use data to make decisions could benefit from research that identifies challenges in building the activities and supporting culture needed to sustain an organization's DDDM. Unfortunately, much of the current research on DDDM in nonprofits is based on information from a single organization or research grounded in surveys from a single respondent in different organizations. Both types of research would have limited validity if perceptions of DDDM implementation vary across or within organizations. Findings from case study research might not generalize to other organizations if DDDM activities and cultures vary across organizations. Cross-organization research drawing information from a single individual in each organization might not have internal validity if staff at the same organization hold different views of DDDM. Selden and Sowa (2011) hinted at difficulties in the latter by showing differences between staff and managers in understanding performance management.

This study examines the perceptions of DDDM activities and culture in organizations driven by a social mission. By drawing information from multiple stakeholders in each of eight social enterprises, it highlights a wide divergence in views both across and within organizations in the activities and culture surrounding DDDM. Its findings provide funders with insights into how to improve an organization's capacity to build and use performance management systems over the long term and researchers with insights about the generalizability of findings from case studies and the validity findings from survey research on DDDM.

The study draws information from multiple individuals each associated with eight different social enterprises. Social enterprises provide a unique opportunity to learn about using data to make decisions in a nonprofit setting. By selling goods and services in competitive markets, such organizations have an incentive to use DDDM to make business decisions. By receiving funding tied to use performance measurements, they have an incentive to use DDDM to make decisions to achieve their social mission. Despite the strong set of incentives to use DDDM that social enterprises share, our research findings show variation in perceptions about it both within and across organizations. Such variations highlight difficulties nonprofits face integrating data into decision making and researchers face designing studies that can be used to develop or understand DDDM in nonprofit organizations.

We conceptualize and capture DDDM activities as sequential events in which organizations (1) collect the data needed to make decisions that enhance service delivery and business operations, (2) analyze the data collected to produce actionable information, and (3) use data systematically to drive decision making. We conceptualize and capture an organization's DDDM culture as a shared belief in the value of using data in decision making and a shared understanding about an organization's DDDM processes, activities, and supports. We examine congruency in the perceptions about DDDM among all organization staff in a position to use data to make decisions and between organization staff and staff at a funder that is actively engaged in building DDDM in the organization.

Social Enterprises in the Study

All eight social enterprises in this study were selected by a venture philanthropy funder in an open competition to receive funding and technical assistance to develop or expand social enterprises that (1) use transitional employment and social supports to help people with substantial barriers to employment to succeed in the labor market (social mission) and (2) generate revenue to achieve financial viability (business mission).

When the venture philanthropy funder selected organizations, one of its criteria was that data guide enterprise expansion decisions. To strengthen the use of data in decision making, the funder convened meetings with each organization monthly to (1) review business financials, performance, and social outcomes and (2) allow the funder's research and evaluation staff to strengthen performance management systems and influence the use of data to fulfill a social mission. The technical assistance and selection of enterprises predisposed to use data in decision making affords an opportunity to observe the behavior of nonprofits likely to be relying on data in decisions. But this also means that patterns identified should be considered preliminary and confirmed in future research.

All eight social enterprises shared a vision of achieving a social and a business mission and stated a commitment to using data in making decisions. However, the specifics of each organization's operations differed (Table 1).[6] As a group, the social enterprises operated in nine different industries, including street cleaning, lobby services, cleaning services, groundskeeping, pest control, and retail stores. Some focused on multiple business lines, with employees either assigned to one line or moving between them, as dictated by demand. They varied greatly in size, with the smallest employing 10 workers per year and the largest employing 500 workers per year. Two started more than 20 years before, this research project began (in 2013); others started with grant money from the funder in 2012.

The enterprises all used transitional jobs to achieve a social mission of moving individuals with severe and multiple employment barriers into jobs but differed in the types of individuals they sought to serve. Target populations included (1) individuals with mental health disabilities, (2) low-income individuals, (3) homeless individuals, (4) at-risk youth, and (5) parolees or exoffenders. Evidence suggests that individuals hired at all organizations faced a multitude of employment barriers. Only 63% of employees had worked in the year before starting the social enterprise job and about 25% never had a job. More than 80% had been arrested during their life, and only 16% had stable housing throughout the year before starting social enterprise employment. Only 23% of their income came from work.

Social enterprises faced a variety of operational challenges, the largest being balancing their business and social missions. All struggled to simultaneously balance investing in employees' skills development and providing employee supports with maintaining a productive workforce. Organizations also often struggled to impose a limit on the duration of employment and to hire supervisory staff with the ability to both operate the business competently and relate to transitional employees. Despite these challenges, evidence suggests that these social enterprises improved the economic self-sufficiency and life stability of workers. One year after starting social enterprise employment, individuals were 181% more likely to be employed and 250% more likely to live in stable housing (Rotz, Maxwell, & Dunn, 2015).

Table 1. Description of Social Enterprises in the Study.

Organization	Business Lines	Employees Hired Annually	Date Started	Target Population
A	Retail	36	1986	Homeless
B	Groundskeeping	100	2012	At-risk youth
C	Lobby services and temporary staffing	51	2011	Formerly incarcerated and low income
D	Construction	10	1991	Formerly incarcerated
E	Pest control	12	2007	Mental health disabilities
F	Cafes and street cleaning	500	2010	At-risk youth
G	Building maintenance	55	2012	Homeless
H	Cleaning service	12	2007	Homeless, mental health disabilities, and at-risk youth

Note. Although we changed some details of the organizations to protect their identity, we maintained an accurate description of enterprises across all organizations.

Method

The venture philanthropy funder that provided funding to the organizations also supported a mixed-methods evaluation of the social enterprises it selected. That evaluation included a survey about the DDDM activities and culture in the social enterprises (available in an online Appendix).[7] The survey was fielded as part of site visits conducted in April 2013. Site visitors asked all staff involved in social enterprise decision making to complete a questionnaire using 5-item Likert-type scales to rate statements about DDDM activities and culture, including:

- how often the social enterprise collects different types of data,
- how often the social enterprise assesses different types of data,
- how the social enterprise uses data,
- the social enterprise's resources for and commitment to DDDM, and
- his or her beliefs about using data to make decisions.

Thirty-six organization staff members completed the questionnaire. These included 17 individuals who considered themselves to be management of the nonprofit hosting the social enterprise, 14 who considered themselves social enterprise management, and 18 who considered themselves front-line or support staff (individuals often reported multiple roles).

In May 2013, staff at the venture philanthropy organization who provided funding and technical assistance around DDDM to the social enterprises were asked to complete the same instrument, which provides an opportunity to understand whether funder staff held the same perceptions about the DDDM as the organization's staff. These eight funder staff independently completed 32 questionnaires, one for each of the organizations to whom they provided technical assistance.

Measures of DDDM

We used information from these questionnaires to develop three summary indices of perceptions of DDDM: one index of DDDM activities and two of culture.[8] The activities index captures the extent to which organizations collect, analyze, and use data in decision making. The two culture indices include one that captures the organization's culture around DDDM and one that describes individuals' beliefs about using data in decision making (used for organization staff only). We quantified an organization's DDDM culture as one in which it committed the resources needed to collecting and analyzing data to facilitate informed data-driven decisions and quantified an individualized DDDM culture as one in which the staff believe that using data furthers the social and business missions of the social enterprise and are comfortable using data for those purposes.

We used a three-step process to develop each index. We first mapped answers (other than *don't know*) to each question onto a 5-point scale, with higher numbers indicating a greater inclination toward DDDM.[9] Answers of *don't know* or otherwise missing responses were assigned to the average across all organizations and respondents (i.e., imputed using the overall mean value).[10] We normalized each item to have a mean of 0 and a standard deviation of 1. We then summed the normalized scales across items and renormalized each sum, so each index had a mean of 0 and standard deviation of 1. A higher value of an index indicates a higher perceived level of DDDM. Cronbach's α was .91 for the activities index (range of .89 to .97 when computed for each of the eight organizations individually), .89 for the organization culture index (range of .61 to .92), and .80 for the staff culture index (range of .71 to .98).

Analytic Methods

We used several different types of quantitative analysis to examine differences in perceptions of DDDM activities or culture. Initially, we compared average levels of DDDM activities and culture across all organizations using each component of the summary indices. This analysis builds an overall understanding of the general perceptions about DDDM in the organizations in our sample and provides a cursory assessment of differences between organization and funder staff.

We built on this description with an analysis of variance (ANOVA) of the three indices and assessed variation between and among staff at different organizations in their perceptions of DDDM. We performed an ANOVA for organization staff (only), funder staff (only), and both organization and funder staff (together) to assess whether within-organization or between-organization differences drive variation in the indices. Because the F statistic generated from the ANOVA captures the ratio of between-organization variation and within-organization variation, its significance ($p \leq .05$) indicates that differences in perceptions between organizations are greater than differences within organizations. We would expect such differences if organization staff held similar views about DDDM (assuming variation in perceptions across organizations). An F statistic of less than 1 implies greater variation in perceptions within an organization than between organizations, and a statistic greater than 1 implies more variation between organizations than within organizations.

We also used Fleiss' κ to more formally assess the similarity of staff views of an organization's DDDM. Fleiss' κ is a statistic typically used to analyze interrater reliability when more than two individuals provide ratings. It captures how the observed agreement in ratings compares with the expected amount of agreement if ratings occurred at random. Because the κ statistic relies on individuals rating the same issues, it is appropriate in assessing similarity of views of an organization's DDDM. We followed Landis and Koch's (1977) method and rated κ values less than .20 as indicating very limited agreement. We would expect higher κ levels if staff agree about an organization's DDDM activities or culture.

Finally, we ran ordinary least squares regressions for all organizations to understand what might drive differences in perceptions of DDDM activities and culture within an organization. For this analysis, we classified respondents as funder staff, nonmanagerial staff at the organization, and managerial staff at the organization. We specified two regressions, both of which included organization fixed effects. The first specification included a single indicator for being a funder staff. A significant coefficient on this indicator implies that the average funder staff held different views on DDDM than the average organization staff. The second specification included indicators for being a funder staff and being a nonmanagerial staff at the organization. A significant coefficient on the indicator for a funder staff implies that the average funder staff held different views on DDDM than managerial staff at the organization, and a significant coefficient on the indicator variable for nonmanagerial staff suggests that managerial and nonmanagerial staff at the organization held different views. Because we included organization fixed effects in the regressions, coefficients show within-organization differences in perceptions about DDDM.

We augmented the quantitative analysis with qualitative information on DDDM activities and culture in Organization A and Organization E. These organizations showcase the variety in DDDM activities across organizations and different ways in which staff believe DDDM could be integrated into the organizations' culture. The qualitative discussion of these organizations was gleaned from analysis of qualitative information we obtained during site visit interviews.[11]

Findings

We discuss results from our analyses along three lines: in the aggregate, within organizations, and across organizations. The aggregate analysis of perceptions in DDDM in all social enterprises in the study provides a context for the subsequent discussions of differences in perceptions within an organization and across organizations.

Perceptions of DDDM Among All Organizations

Differences in the components that make up the DDDM indices foreshadow the variations in perceptions of DDDM across and within organizations (Table 2). We found that most organization staff see their organization as undertaking some DDDM activities, but the extent of consensus about the undertaking varies considerably by activity. For example, more than 90% of staff reported their organization collects data on an employee's job performance, but only about 55% reported that their organization collects data on the life circumstances of workers after social enterprise employment (even though a central goal of all organizations is to improve workers' long-run circumstances).

Perceptions of organization staff suggest that data are often collected but less often analyzed. This might be expected, given the sequential nature of DDDM (i.e., data must be collected if it is to be analyzed). For example, although more than 90% said the organization collects data on job performance, only about 81% said someone in the organization analyzes it; and although about 64% said the organization collects information on the demand for business, only about 44% say someone analyzes it. Conversely, staff are generally more likely to report that their organization uses data than analyzes data: Percentages of staff reporting that their organization uses data are closer to those percentages of staff reporting that data are collected than the percentages of staff reporting that their organization analyzes data. For example, about 89% of staff believed their organization uses data to improve job performance, and 67% believed it uses it to identify business opportunities. Importantly, about 70% believed discussions of data are translated into actions. One potential explanation for this difference is lack of clarity in what it means to analyze data.[12]

In general, organization staff reported that they believe in using data to make decisions but reported less confidence in their organization's ability to do so (Table 2). All organization staff said they believed that using data could improve services provided to employees, and 83% said they believed using data builds an understanding of how the enterprise operates. Only 14% believed using data takes away from spending time helping employees (the population they desire to help). Somewhat fewer staff believed the organization's culture supports DDDM: About 72% believed that using data is part of their organization's culture, 69% said that their organization uses data well, 50% said that the organization has sufficient resources to collect data, and 44% said that the organization has an efficient data collection system in place.

Table 2. Measures of DDDM Activities and Culture.

	Organization Staff		Funder Staff	
Number of Observations	36		32	
Activities				
	Very Often or *Often*	*Don't Know*	*Very Often* or *Often*	*Don't Know*
Collecting data				
Before an employee starts work, we collect data on . . .				
Work skills	86.1	2.8	53.1*	34.4*
Need for job supports	83.3	0.0	46.9*	40.6*
Need for life supports	66.7	5.6	43.8	40.6*
While an individual is working, we collect data on . . .				
Job performance	91.7	5.6	53.1*	28.1*
Job development or job placement	88.9	5.6	34.4*	31.3*
Work assignments	86.1	8.3	59.4*	34.4*
Work or life stability supports	69.4	5.6	18.8*	34.4*
After an individual leaves work, we collect data on . . .				
Employment status	77.8	0.0	37.5*	21.9*
Life circumstances	55.6	2.8	15.6*	37.5*
We collect data on . . .				
Demand for business	63.9	8.3	28.1*	56.3*
Customer satisfaction	55.6	13.9	12.5*	56.3*
Assessing data				
We assess data on . . .				
Performance during social enterprise employment	80.6	8.3	37.5*	28.1*
Employment after social enterprise employment	77.8	2.8	31.3*	34.4*
Skills and needs before social enterprise employment	72.2	2.8	34.4*	46.9*
Skills developed during social enterprise employment	69.4	8.3	28.1*	21.9
Supports during social enterprise employment	66.7	8.3	12.5*	40.6*
Demand for business	44.4	19.4	15.6*	56.3*
	Strongly Agree or *Agree*	*Don't Know*	*Strongly Agree* or *Agree*	*Don't Know*
Using data				
Discussions of data are translated into actions	69.4	5.6	37.5*	31.3*
We use data to . . .				
Help improve job performance	88.9	2.8	28.1*	25*

(continued)

Table 2. (continued)

	Organization Staff		Funder Staff	
Number of Observations	36		32	
Activities				
	Strongly Agree or *Agree*	*Don't Know*	*Strongly Agree* or *Agree*	*Don't Know*
Make the social enterprise more productive	86.1	8.3	31.3*	37.5*
Identify and develop needed supports	77.8	2.8	28.1*	34.4*
Increase efficiency	77.8	13.9	34.4*	40.6*
Help develop life skills	75.0	2.8	15.6*	31.3*
Identify and develop training programs	66.7	2.8	28.1*	25*
Identify business opportunities	66.7	13.9	18.8*	50*
Improve employment after social enterprise	61.1	8.3	18.8*	34.4*
Improve life circumstances after social enterprise	50.0	5.6	9.4*	40.6*
Culture				
	Very Often or *Often*	*Don't Know*	*Very Often* or *Often*	*Don't Know*
Organization culture				
I believe using data to make decisions is part of the organization's culture	72.2	0.0	31.3*	21.9*
I believe using data in this organization is not done well (inverse for index)	30.6	2.8	31.3	25.0*
In my organization, we . . .				
Have staff with expertise in data analysis	61.1	0.0	28.1*	28.1*
Have sufficient resources to collect data	50.0	0.0	28.1	28.1*
Have an efficient data collection system in place	44.4	0.0	21.9*	25.0*
Individual beliefs				
I believe that using data . . .				
Can improve services provided to employees	100.0	0.0	n.a.	n.a.
Benefits the work we do with employees	97.2	0.0	n.a.	n.a.
Makes me uncomfortable (inverse used for index)	8.3	0.0	n.a.	n.a.
Is not how to help our population (inverse used for index)	8.3	2.8	n.a.	n.a.
Builds an understanding of how the enterprise operates	83.3	0.0	n.a.	n.a.
Takes away time spent helping employees (inverse for index)	13.9	2.8	n.a.	n.a.

Note. n.a. indicates that the measure was not used for the population. DDDM = data-driven decision making.
*Statistically significant difference between organization and funder staff at the $p \leq .05$ level.

Different Perceptions, Same Organization

Figure 1 demonstrates that staff at the funder typically thought DDDM was less common and less culturally supported than organization staff reported. This figure shows the average value of each index in each organization as well as the range of index values. It shows that funder staff frequently rated DDDM activities and culture one or more standard deviations below the average rating provided by organization staff.[13] The vast majority of the funder staff rated DDDM activities or organization culture below the lowest rating of any organization staff member, except in Organization B. In half the organizations (C, E, F, and H), all of the funder staff rated levels of activities below the lowest rating provided by organization staff. Fewer discrepancies seem to exist between funder and organization staff in perceptions of organization culture, although in one organization (H), all funder staff rated DDDM culture lower than did all organization staff as did all but one funder staff member in another organization (E).

Descriptive analysis of individual survey items confirms the disagreement between funder and organization staff about an organization's DDDM activities and culture (Table 2). For example, 86% of organization staff believed their organization collects data on an employee's work skills, and 83% believed their organization collects data on an employee's need for job supports (its social mission). In contrast, only 53% of funder staff believed the organization collects data on work skills, and only 47% believed it collects data on the need for job supports. Such differences occur on every item examined, with perceptions on using data to make business decisions being starkest. Furthermore, about 56% of organization staff said their organization collects data on customer satisfaction, compared with 13% of funder staff. About 44% of organization staff and 16% of funder staff said the organization analyzes data on business demand, whereas about 67% of organization staff but only 19% of funder staff said the organization uses data to identify business opportunities. This discrepancy and the relatively large percentage of funder staff who say they don't know about DDDM in organizations are particularly striking in light of the focus and assistance the funder has provided the organization to build or strengthen their use of data in decision making.

Large differences also exist between funder and organization staff when evaluating the organizations' DDDM culture (Table 2). Less than one third of funder staff but 72% of organization staff believed the organization has a culture that emphasizes using data to make decisions, with about 70% of both funder and organization staff believing DDDM is done well. Both populations identified insufficient resources for DDDM as an issue. For example, only about half of organization staff and 28% of funder staff believed that the organizations have sufficient resources to collect data for use in decision making.

The differences in the perceptions of an organization's DDDM activities and culture are not confined to differences between funder and organization staff: Organization staff hold disparate views of DDDM activities and culture in their organization (Figure 1A and B), and their beliefs about DDDM differ widely (Figure 1C). The indexed responses of organization staff perceptions of DDDM activities and culture often have a range of more than 1.5 standard deviations *within an organization* (see Table 3 for precise ranges). Particularly inconsistent perceptions exist in Organizations B, C, D, and G when we define inconsistency as the prevalence of organization staff rating DDDM more than 1 standard deviation above or below the organization mean. Such disparities exist with respect to DDDM activities in Organization B, culture in Organization D, and beliefs in C and G (Table 3).

Perceptions of DDDM culture were more disparate than the perceptions of activities in most organizations, suggesting a particularly strong lack of consensus among these measures. The exceptions are Organization B, in which all staff perceived a high level of organization culture, and organization C, in which all rated culture as below average (i.e., below the mean of 0). In three of eight organizations, staff perceptions of organization culture fall both above and below the average of all organizations as was true in six of eight organizations regarding beliefs about the value of DDDM.

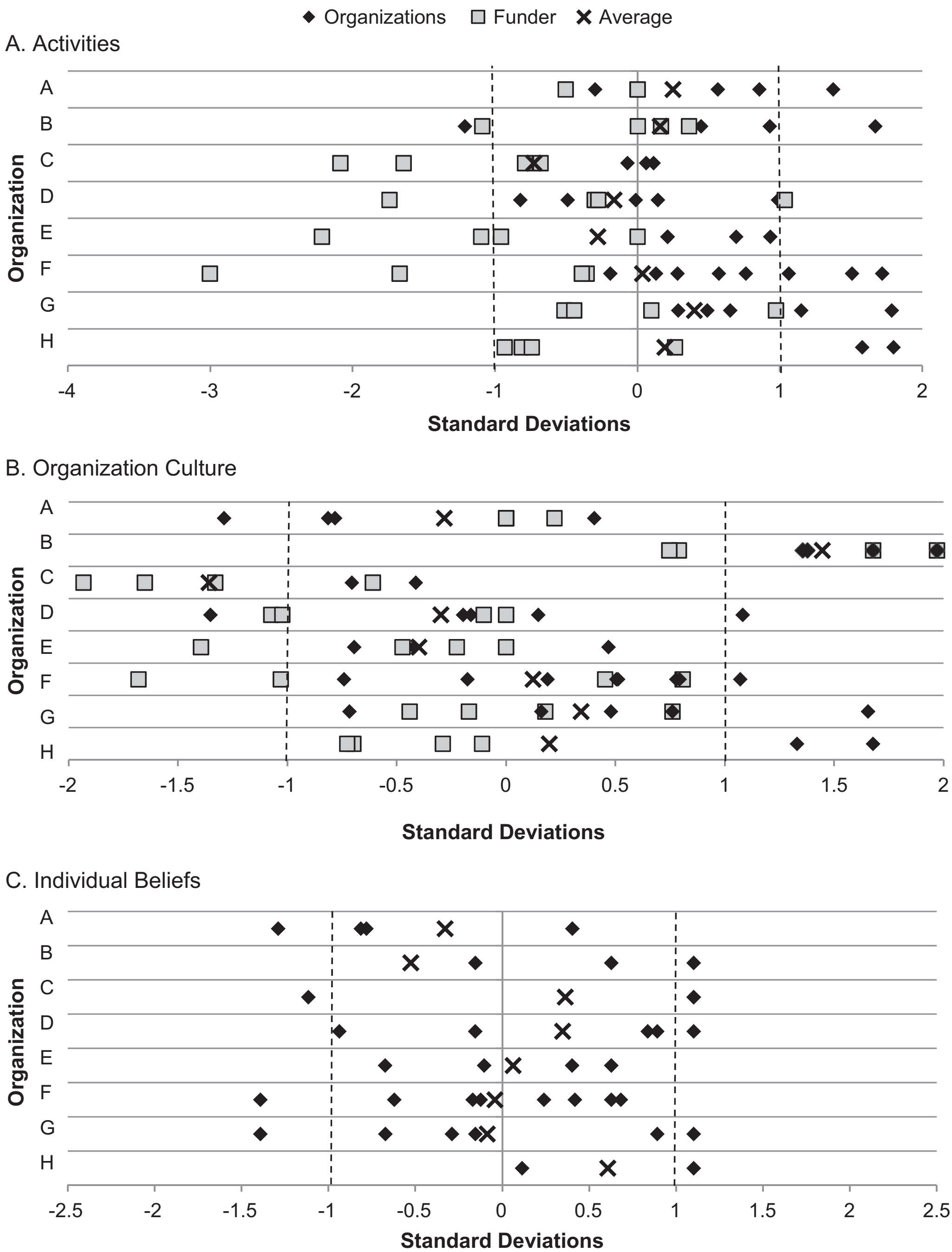

Figure 1. Within- and between-organization indices of data-driven decision making. Indices have a mean of 0 and standard deviation of 1 across respondents. The *x*-axis measures standard deviations from the mean. Missing values are replaced with overall mean. Calculations are based on 36 surveys from organization staff and 32 surveys from funder staff.

Table 3. Within- and Between-Organization Variation in DDDM Activities and Culture: Analysis of Variance (ANOVA).

	Average Level Within an Organization			Range in Standard Deviations		
		Culture			Culture	
Organization	Activities	Organization	Individual Beliefs	Activities	Organization	Individual Beliefs
Total (organization and funder staff)						
Number of observations	68	68	68	68	68	68
Organization						
A	0.25	−0.28	−0.33	1.88	1.69	1.95
B	0.16	1.45	−0.53	2.87	1.22	4.78
C	−0.73	−1.36	0.36	2.19	2.45	2.22
D	−0.16	−0.30	0.35	2.77	2.43	2.04
E	−0.28	−0.40	0.06	3.14	1.86	1.30
F	0.03	0.12	−0.04	4.71	2.75	2.07
G	0.40	0.34	−0.09	2.30	2.37	2.49
H	0.19	0.20	0.61	2.73	2.40	0.99
Across all organizations	0.00	0.00	0.00	4.80	4.84	4.78
ANOVA *F* statistic	1.02	8.80	0.42	—	—	—
p value	.430	.000	.879	—	—	—
Organization staff						
Number of observations	36	36	36	36	36	36
Organization						
A	0.62	−0.62	−0.33	1.67	1.69	1.95
B	0.46	1.60	−0.53	2.88	0.61	4.78
C	0.03	−1.33	0.36	0.18	2.45	2.22
D	−0.04	−0.10	0.35	1.81	2.43	2.04
E	0.51	−0.27	0.06	0.72	1.16	1.30
F	0.73	0.37	−0.04	1.91	1.81	2.07
G	0.65	0.52	−0.09	2.27	2.37	2.49
H	1.69	1.50	0.61	0.22	0.35	0.99
Across all organizations	0.54	0.20	0.00	3.01	4.84	4.78
ANOVA *F* statistic	1.53	6.23	0.42	—	—	—
p value	.199	.000	.879	—	—	—

(continued)

Table 3. (continued)

	Average Level Within an Organization			Range in Standard Deviations		
		Culture			Culture	
Organization	Activities	Organization	Individual Beliefs	Activities	Organization	Individual Beliefs
Funder staff						
Number of observations	32	32	n.a.	32	32	n.a.
Organization						
A	−0.13	0.06	n.a.	0.50	0.22	n.a.
B	−0.14	1.30	n.a.	1.45	1.22	n.a.
C	−1.30	−1.38	n.a.	1.41	1.32	n.a.
D	−0.32	−0.55	n.a.	2.78	1.07	n.a.
E	−1.07	−0.52	n.a.	2.22	1.40	n.a.
F	−1.36	−0.36	n.a.	2.65	2.49	n.a.
G	0.03	0.08	n.a.	1.49	1.20	n.a.
H	−0.56	−0.46	n.a.	1.20	0.62	n.a.
Across all organizations	−0.60	−0.23	n.a.	4.04	3.90	n.a.
ANOVA *F* statistic	1.85	5.84	n.a.	—	—	—
p Value	.124	.001	n.a.	—	—	—

Note. Level is measured as the average value of the index, with each index having a mean of 0 and a standard deviation of 1. DDDM = data-driven decision making; n.a. = indicates that the measure was not used for the population.

Table 4. Within- and Between-Organization Variation in DDDM Activities and Culture: Fleiss' κ.

	Average Level Within an Organization		
		Culture	
Organization	Activities	Organization	Individual Beliefs
Total (organization and funder staff)			
Number of observations	68	68	68
Organization			
A	−.05	−.06	.00
B	.11	.28	.03
C	.04	.11	.09
D	.05	−.03	.29
E	−.06	−.07	.07
F	.04	−.02	.20
G	.01	.09	.15
H	−.03	−.10	.33
Organization staff			
Number of observations	36	36	36
Organization			
A	.10	−.08	.00
B	.07	.37	.03
C	.29	−.14	.09
D	.08	−.08	.29
E	.01	.01	.07
F	.06	.01	.20
G	−.03	.07	.15
H	−.08	.05	.33
Funder staff			
Number of observations	32	32	n.a.
Organization			
A	−.05	−.13	n.a.
B	.17	.11	n.a.
C	.02	.37	n.a.
D	−.07	−.06	n.a.
E	−.12	−.08	n.a.
F	.04	−.08	n.a.
G	.01	.00	n.a.
H	.05	−.09	n.a.

Note. n.a. indicates that the measure was not used for the population. DDDM = data-driven decision making.

The Fleiss' κ analysis (Table 4) also confirms the dissimilarity of staff views of DDDM activities, culture, and beliefs: The statistic is generally below .2, our cutoff for slight agreement. Only one organization (B) shows evidence of staff agreeing on the nature of the organization's culture ($\kappa = .37$). Additionally, staff in only one organization (C) have slight agreement ($\kappa = .29$) on which activities occur in support of DDDM. Individual beliefs about DDDM align somewhat more in three organizations (D, F, and H), for which the Fleiss' κ scores suggest slight agreement ($\kappa = .29$, $\kappa = .20$, and $\kappa = .33$, respectively).

Our regression analysis supports the findings that both differences in perceptions between funder and organization staff and among different types of organization staff contribute to the divergence in perceptions about DDDM within an organization (Table 5). Coefficients indicate statistically significant differences between funder staff and organization staff in all specification (1) estimations. On

Table 5. Within-Organization Differences in DDDM: Regression Analysis.

Respondent Type	Measure of DDDM: Activities (1)	Activities (2)	Organization Culture (1)	Organization Culture (2)	Individual Beliefs (2)
Funder staff	−1.151*** (0.210)	−0.945*** (0.227)	−0.395** (0.189)	−0.188 (0.208)	n.a.
Nonmanagerial staff at organization		0.644** (0.250)		0.647** (0.246)	−0.143 (0.432)
Observations	68	68	68	68	36
R^2	.429	.474	.544	.590	.121

Note. Numbers show estimated coefficients and (standard errors). Regressions also include controls for organization fixed effects. All outcome variables are normalized to have mean of 0 and standard deviation of 1. Robust standard errors are in parentheses. n.a. indicates that the measure was not used (because funder staff were not included in the estimation). DDDM = data-driven decision making.
*Statistically significant difference between organization and funder staff at the $p \leq .05$ level.

average, respondents from the funder rated an organization's DDDM activities 1.15 standard deviations lower than organization staff, which suggests a very different view of what DDDM is occurring in an organization. Differences are smaller when assessing an organization's culture, with a statistically significant 0.40 standard deviation difference. Coefficients also indicate statistically significant differences between managerial and nonmanagerial staff in two of the three specification (2) estimations. Nonmanagerial staff rated both DDDM activities and culture higher than managerial staff: 0.64 standard deviations for activities and 0.65 standard deviations for culture. Both differences are statistically significant. Differences in individual beliefs, however, were smaller (0.14 standard deviations) and statistically insignificant.

Different Perceptions, Different Organizations

Formal statistical tests confirm differences in staff perceptions about their organizations' DDDM activities and culture but also highlight differences in perceptions across organizations. Our ANOVA suggests that staff perceptions about DDDM activities and their beliefs about using data to make decisions are as different across organizations as they are within an organization: The only *F* statistic significantly other than 1 is for organization culture, which suggests that perceptions about DDDM activity and staff beliefs across organizations are as different within an organization as across an organization (Table 3). Such patterns hold for both organization staff and funder staff perceptions of DDDM activities.

Case study analysis supports cross-organization differences in DDDM activities and culture, even though organizations were selected, in part, because of their predisposition toward DDDM and were subject to the same funder pressures to build performance measurement systems. Indeed, discussions of DDDM with managerial and nonmanagerial staff during site visits illustrate the rather stark differences in DDDM activities, culture, and beliefs across organizations.

Organization A. Organization A operates a long-established retail social enterprise. It hires about 36 people each year, focusing on employing homeless individuals. Funder and organization staff perceptions suggest DDDM activities are slightly above the average for all organizations studied. Organization staff—but not funder staff—assess its culture as below average (Figure 1B). Funder staff rated DDDM activities of the organization about 0.75 standard deviations lower than organization staff but rated the culture higher by about 0.68 standard deviations. Although perceptions of DDDM were more cohesive than in any other organization studied, the variance in these was still large.

During semistructured interviews, both organization and funder staff cited examples of how the organization used data in decision making. Both mentioned the organization had a tracking system to monitor work readiness, through which data on work attendance, productivity, hygiene, and willingness to learn were collected and assessed by counselors. Staff reported that these assessments are shared with employees regularly and used to help employees transition to employment. The organization also used data to track business activities; other analysis demonstrated that the social enterprise was currently solvent with operating revenues and subsidies covering costs (Rotz et al., 2015). Despite perceptions of success in assessing work readiness of current employees, staff reported deficiencies in measuring whether the organization achieved its social mission of moving individuals into employment. One staff member noted that the social enterprise "needs to develop better tracking systems on worker outcomes" after social enterprise employment ends and another lamented that he or she was "not sure what was happening with transitions for clients out of the social enterprise ... [but] sensed that outcomes were not good."

Organization E. Organization E operates a pest control enterprise that formed shortly before receiving financial assistance from the philanthropic funder. It hires about 12 individuals each year and targets those with mental health disabilities. Overall perceptions of DDDM activities lie below the average of all enterprises studied, although funder and organization staff disagreed strongly about them. Organization staff perceived DDDM activities as above average for all organizations but funder staff perceived them as lower. Agreement about the organization's DDDM culture was closer with all but one funder staff and one organization staff rating culture as below average. Still, organization staff expressed near-average beliefs about the value of DDDM.

Discussions with staff confirmed a lack of DDDM in the organization. Several mentioned that they obtained detailed information on individuals before their placement in social enterprise employment but had only limited information about them after they were hired. Indeed, they expressed a lack of knowledge about the number of individuals served through transitional employment, with one staff noting, "It would be helpful to have a spreadsheet of some kind to help us track all the jobs we've created." Staff also spoke of a lack in DDDM to advance its business mission. One staff indicated that the need to act swiftly and take advantage of business opportunities meant decisions were made quickly and without thorough thinking or use of data. Other staff, both at the organization and funder, expressed concerns that the organization lacked the capacity to process and use basic financial data. Irrespective of cause, the result was a disagreement about whether the social enterprise was financially viable: Some said financial losses were large, one suggested the enterprise was breaking even, and others could not or would not estimate viability.

Discussion

This study examined differences in perceptions about DDDM among staff within and across eight organizations, capturing DDDM activities (collecting, analyzing, and using data to make decisions) and two dimensions of DDDM culture (organization culture and staff beliefs about the value of DDDM). Perceptions were quantified through a survey administered to all staff in eight social enterprises who are in a position to use data to make decisions and staff at a funding organization who worked directly to build the organization's use of DDDM.

Analysis showed that perceptions about DDDM not only differed across organizations but also within an organization, both among staff at the organization and between funder and organization staff. Perceptions of DDDM activities and culture that organization staff held often had a range of more than 1.5 standard deviations *within an organization*, and staff perceptions about DDDM activities and their beliefs about using data to make decisions were as different within an organization as they were across organizations. Some of these differences can be explained by respondent type. Managerial staff at organizations rated DDDM activities and culture lower than nonmanagerial staff, and funder staff rated activities and culture lower than organization staff, irrespective of the perceived level of an organization's activity and culture.

One potential explanation for different perceptions within an organization is different definitions of data. Frontline staff in the organizations often had a social work background, and discussions with them during site visits suggested that they define data as qualitative information (Maxwell, Rotz, Dunn, Rosenberg, & Berman, 2013). In contrast, discussions with staff at the funder and some higher level organization staff suggest these individuals view data as quantitative information. These findings are consistent with findings from Carman (2007) and Hwang and Powell (2009), who demonstrate that nonprofit staff often believe that measured outcomes do not capture important aspects of the work they do with clients. Benjamin (2012) further demonstrates that existing performance measurement frameworks focus on program implementation, rather than the process through which staff work with clients, which might explain some differences in perspective.

Building a culture to support DDDM requires an organization to have a common understanding of what it means to use data in decision making. A lack of common understanding among organization staff and between funders and organizations might limit efforts to promote DDDM both directly through goal setting and assessing progress and indirectly by impeding the development of a DDDM culture.

Results of this study can help identify three areas that might facilitate building DDDM systems within a mission-driven organization. First, the incongruity in perceptions about an organization's DDDM suggests that funders or executives wanting to increase reliance on data in decision making should begin by developing a common, organizational understanding of what data and DDDM means, which DDDM activities the organization should be and are—or are not—currently undertaking, and the value of using data to make decisions. Developing this understanding could take place before or while establishing an organization's culture, but one cannot assume the understanding to exist. Before developing a DDDM investment or capacity-building plan, a champion of DDDM should clearly articulate his or her expectations and parameters.[14] Such a plan should include an orientation to organization principles, an emphasis on the value DDDM can have on an organization's ability to execute its mission, a commitment to data transparency, and a clear description of an organization's vision for its target beneficiaries. The DDDM survey used in this study provides a basic structure for assessing DDDM in an organization and is easy for an organization's line staff, managers, and executives and funder staff to complete. Organizations could then use the results of the survey to initiate a discussion of what DDDM means.

Second, funders must not take for granted that an organization will have a culture amenable to DDDM, even if the (one or two) staff members they interact with believe this is the case. Ultimately, funders and the organizations they support want to accomplish a similar social mission. Although some organizations see DDDM as an option to increase progress toward their goals and to provide funders with information to more effectively target their resources, others might not. Many of the organizations we studied did not have the culture in place to support DDDM, even though the funder was explicit in its requirement that data be used for decisions to build or expand enterprises for which they sought funding (and perhaps even though the organization's representative to the funder believed this was so). In absence of a culture of grounding decisions in verifiable data, financial investments in DDDM might be unproductive, making it as important for funders and executives seeking change to build a foundation for an organizational learning culture, as it is to provide financial resources for building DDDM activities. Although funders, in particular, can often easily finance an organization's new data system, they might struggle with the more complex and nuanced task of influencing the organization's culture to use data. In this regard, they must be conscious of the power dynamics at play in the funder–grantee–beneficiary relationship; understand that cultivating a strong, stable DDDM culture is a process that can be marked by incremental milestones; and ensure that DDDM is supported by a learning organization culture of using *the right* data to make *the right* decisions that support *the right* outcomes in *the right* moment. When data are used to make high-stakes decisions that affect not just an organization's capacity to carry out its mission but also the people who depend on the essential services that an organization provides, funders and organizations must work closely together to ensure that data are defined and used properly in decision making.

Finally, findings from this study have important implications for researchers seeking to advance DDDM by identifying general principles to guide mission-driven organizations in developing effective performance management systems. The large variation in perceptions of an organization's DDDM shown in this study suggests that any analysis based on information from a single respondent in an organization might produce different conclusions with a new draw from the distribution of respondents in the organization. Indeed, results from this study suggest that research based on surveys about DDDM collected from a single individual at each organization might suffer from potentially large biases due to the mismeasurement of critical variables (i.e., attenuation biases). In addition, the variation in perceptions of DDDM across organizations suggests that research results from case studies might not generalize to other organizations. Taken together, the within- and between-organization differences in perceptions of DDDM suggest that researchers attempting to assess the use of data in decision making or to establish impacts of DDDM should demonstrate the robustness of their results to different samples of organizations or provide evidence of the robustness of their results to different respondents in an organization.

Appendix

Robustness of Analysis to Alternative Treatment of Missing Data

Text Table 2 shows that 12% to 38% of respondents answered don't know to our questionnaire items, with funder staff more likely to respond in this way than organization staff. Respondents were also more likely to report an answer of don't know in certain domains, especially when asked about DDDM as it relates to the business mission of the social enterprise. Many potential approaches are available to deal with a large number of don't know responses. In this study's main analyses, we imputed all missing data based on the overall sample mean. However, the large number of valid options for imputation suggests that we need to cross-validate our results using alternative methods. Table A1 contains such alternate estimates for the calculations of Cronbach's α (used to validate our creation of DDDM indices) and the ANOVA tests of within- versus between-organization variations (used to demonstrate the lack of cohesion in an organization around DDDM). Estimates of Fleiss' κ are not included in this analysis, as the statistic is calculated by treating missing values as their own categorical response.

The columns of Table A1 each contain the results obtained by using a different strategy to handle missing data:

1. Original imputation: Missing data are replaced with the average response across all respondents.
2. Imputation alternative 1: Missing data are replaced with the average response within organization.
3. Imputation alternative 2: Missing data are replaced with the average response by respondent type (staff at funder or organization).
4. Imputation alternative 3: Missing data are replaced with the average response within organization by respondent type.
5. Omit missing: Missing data are omitted from the analysis.

6. Set missing to neutral: Missing data are replaced with the most neutral response option (e.g., neither *agree* nor *disagree*).

As the table shows, our results are robust to the imputation method used. Cronbach's α changes little across methods of handling missing data and always implies the elements of our indices demonstrate internal consistency. The ANOVA suggests that staff beliefs about DDDM are as different within organizations as they are across organizations. Furthermore, the analysis always suggests higher variation in ratings of DDDM culture across organizations than within organizations. The precise imputation measure matters for only one conclusion: Some imputations enable us to conclude greater agreement among funder staff on DDDM activities within an organization than across organizations, implying that funder staff have more consistent beliefs about DDDM activities than our main analysis suggests. However, because such agreement still does not exist under any imputation method when data from organization and funder staff are pooled, our overall conclusions do not change with different methods of handling missing data.

Table A1. Robustness of Results to Alternative Methods of Handling Missing Data.

	Original Imputation	Imputation Alternative 1	Imputation Alternative 2	Imputation Alternative 3	Omit Missing	Set Missing to Neutral
Cronbach's α						
Activities	0.91	0.91	0.93	0.93	0.92	0.93
Organization culture	0.89	0.89	0.89	0.89	0.90	0.89
Individual beliefs	0.80	0.80	0.80	0.80	0.80	0.80
ANOVA *F* statistics (*p* value)						
All respondents						
Activities	1.02 (0.430)	1.76 (0.112)	0.84 (0.562)	1.44 (0.208)	1.01 (0.435)	0.75 (0.627)
Organization culture	8.80 (0.000)	10.66 (0.000)	9.14 (0.000)	8.78 (0.000)	7.79 (0.000)	9.15 (0.000)
Individual beliefs	0.42 (0.879)	0.42 (0.880)	0.42 (0.879)	0.42 (0.880)	0.41 (0.887)	0.41 (0.888)
Organization staff						
Activities	1.53 (0.199)	1.57 (0.187)	1.40 (0.213)	1.61 (0.173)	1.51 (0.205)	1.53 (0.198)
Organization culture	6.23 (0.000)	6.13 (0.000)	6.21 (0.000)	6.07 (0.000)	6.20 (0.000)	6.28 (0.000)
Individual beliefs	0.42 (0.879)	0.42 (0.880)	0.42 (0.879)	0.42 (0.880)	0.41 (0.887)	0.41 (0.888)
Funder staff						
Activities	1.85 (0.124)	3.08 (0.019)	1.53 (0.205)	4.38 (0.003)	1.84 (0.134)	1.44 (0.234)
Organization culture	5.84 (0.001)	7.36 (0.000)	6.39 (0.000)	8.79 (0.000)	5.72 (0.001)	6.30 (0.000)
Individual beliefs	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Note. Level is measured as the average value of the index, with each index having a mean of 0 and a standard deviation of 1. n.a. indicates that the measure was not used for the population. ANOVA = analysis of variance.

Acknowledgment

The authors thank Jacqueline Berman, Tracy Lam-Hine, and Viki Rasmussen for work from which the study emerged; Lindsay Cattell, Adam Dunn, Chrissie Grover-Roybal, Mindy Hu, and Jessica McElroy for research assistance; and Ron D'Amico, Josh Haimson, Kristin Hallgren, and anonymous reviewers for comments on previous drafts. The views expressed are those of the authors and should not be attributed to any organization.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research is based on work supported by REDF with funds from the Corporation for National and Community Service under Social Innovation Fund Grant No. 10SIHCA001. Opinions or points of view expressed are those of the authors and do not necessarily reflect the official position of, or a position that is endorsed by, the Corporation or the Social Innovation Fund program.

Notes

1. In contrast, for-profit firms face market pressures to adopt data-driven decision making (DDDM) strategies to provide a competitive edge (Economist Intelligence Unit, 2012) and public sector organizations face requirements to meet performance measures (Ikemoto & Marsh, 2007; Julnes & Holzer, 2001; Marsh, Pane, & Hamilton, 2006). In the federal government, the Office of Budget and Management issued guidance to strengthen "agencies' abilities to continually improve program performance by applying existing evidence about what works, generating new knowledge, and using experimentation and innovation to test new approaches to program delivery" (Burwell, Munoz, Holdren, & Krueger, 2013, p. 1).
2. Such a strategy has been used by funders such as the Bill & Melinda Gates Foundation (http://www.gatesfoundation.org/How-We-Work/General-Information/Evaluation-Policy) and the Edna McConnell Clark Foundation (http://www.emcf.org/our-strategy/our-selection-process/evidence/).
3. We differentiate between DDDM and external evaluations, which can provide or draw on information typically used in a DDDM but are not ongoing processes of continuously and systematically collecting information and integrating it into decision making (Liket, Rey-Garcia, & Maas, 2014). Indeed, more externally driven evaluations can leave an organization "drowning" in data that neither contribute to nor coincide with strategic decision making (Snibbe, 2006).
4. Available tools include scorecards, logic models, formal evaluations, participant observations, and comprehensive guides that provide concrete steps for developing and implementing DDDM (Behn, 2004; Carman & Fredericks, 2008; Hatry, 2014; Kaplan, 2001; Speckbacher, 2003).
5. Studies suggest that performance management systems should (1) focus on improving programs and not meeting the funding requirements (Carman, 2007), (2) have support from both organizational leaders (Alaimo, 2008; Hoole & Patterson, 2008) and frontline staff (Carrilio et al., 2003; Julnes & Holzer, 2001; Moynihan, Pandey, & Wright, 2012), and (3) promote an organizational culture of learning and improvement (Hoole & Patterson, 2008).
6. Information on the social enterprises presented in Table 1 and the text is taken from Maxwell et al. (2013) and Rotz et al. (2014).
7. The funder worked with researchers to structure the survey to reflect DDDM activities and culture measures in areas that data were expected to be used in making decisions about business operations and provision of work and supports to meet social goals.
8. Table 2 maps the survey questions to each index. We developed only one index for DDDM activities because they are mostly cumulative: To assess data, you must collect it, and to use it, you must collect and analyze it.
9. For questions asking about a positive orientation toward DDDM, 5 = *strongly agree* or *very often*, 4 = *agree* or *often*, 3 = *neither agree nor disagree* or *neither often nor rarely*, 2 = *disagree* or *rarely*, 1 = *strongly disagree* or *very rarely*. For questions asking about a negative orientation, the scale is reversed.
10. The Appendix discusses the robustness of our results to several different assumptions about how to handle responses of don't know.
11. Maxwell et al. (2013) describes the site visits and methods to analyze information from them. We preserve the anonymity of organizations in this discussion by changing operational details not associated with DDDM.
12. The instrument tried to minimize differences in defining what it means to analyze data by using the word *assess* when phrasing the questions and defining it as

systematically counting up the characteristics of employees (such as the number enrolled or number working), or looking at trends and patterns in the information you have (such as attendance or most frequently needed supports). Data assessment is one step beyond data collection. It means that the organization has some way of organizing and examining the data collected.

13. All respondents were pooled to estimate this standard deviation.
14. See King and Volkov (2005); Cairns, Harris, Hutchison, and Tricker (2005); and Taut (2007) on the importance of a DDDM champion.

Supplemental Material

The online data supplements are available at http://aje.sagepub.com/supplemental.

References

Alaimo, S. P. (2008). Nonprofits and evaluation: Managing expectations from the leader's perspective. *New Directions for Evaluation*, *119*, 73–92.

Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, *63*, 586–606.

Behn, R. D. (2004). *Performance leadership: 11 better practices that can ratchet up performance*. Washington, DC: IBM Center for the Business of Government.

Benjamin, L. M. (2012). Nonprofit organizations and outcome measurement from tracking program activities to focusing on frontline work. *American Journal of Evaluation*, *33*, 431–447.

Boyle, R., Lemaire, D., & Rist, R. C. (1999). Introduction: Building evaluation capacity. In R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity* (pp. 1–19). New Brunswick, NJ: Transaction.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011, April 22). *Strength in numbers: How does data-driven decisionmaking affect firm performance?* Retrieved June 13, 2013, from http://dx.doi.org/10.2139/ssrn.1819486

Burwell, S. M., Munoz, C., Holdren, J., & Krueger, A. (2013). *Next steps in the evidence and innovation agenda*. Retrieved July 22, 2015, from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-17.pdf

Cairns, B., Harris, M., Hutchison, R., & Tricker, M. (2005). Improving performance? The adoption and implementation of quality systems in UK nonprofits. *Nonprofit Management and Leadership*, *16*, 135.

Carman, J. G. (2007). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation*, *28*, 60–75.

Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*, *119*, 51–71.

Carman, J. G., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, *31*, 84–104.

Carrilio, T. E., Packard, T., & Clapp, J. D. (2003). Nothing in—nothing out: Barriers to the use of performance data in social service programs. *Administration in Social Work*, *27*, 61–75.

Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*, *84*, 98–107.

Economist Intelligence Unit. (2012). *The deciding factor: Big data & decision making*. Retrieved October 18, 2013, from http://www.capgemini.com/resources/the-deciding-factor-big-data-decision-making

Hatry, H. P. (2014, July). *Transforming performance measurement for the 21st century*. Washington, DC: The Urban Institute.

Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way funded programs: Expectations and reality. *New Directions for Evaluation*, *119*, 13–35.

Hoole, E., & Patterson, T. E. (2008). Voices from the field: Evaluation as part of a learning culture. *New Directions for Evaluation*, *119*, 93–113.

Huffman, D., Lawrenz, F., Thomas, K., & Clarkson, L. (2006). Collaborative evaluation communities in urban schools: A model of evaluation capacity building for STEM education. *New Directions for Evaluation*, *109*, 73–85.

Hwang, H., & Powell, W. W. (2009). The rationalization of charity: The influences of professionalism in the nonprofit sector. *Administrative Science Quarterly*, *54*, 268–298.

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the ''data-driven'' mantra: Different conceptions of data-driven decision making. Retrieved October 18, 2013, from http://www.rand.org/pubs/reprints/RP1372.html

Julnes, P. D., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, *61*, 693–708.

Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management & Leadership*, *11*, 353–370.

Kiernan, N. E., & Alter, T. R. (2004). Can a web site be used as a vehicle for organizational learning about evaluation? *Journal of Higher Education Outreach and Engagement*, *10*, 121–134.

King, J. A., & Volkov, B. (2005). A framework for building evaluation capacity based on the experiences of three organizations. *CURA Reporter*, *35*, 10–16.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*, 21–31.

LeRoux, K., & Wright, N. S. (2010). Does performance measurement improve strategic decision making? Findings from a national survey of nonprofit social service agencies. *Nonprofit and Voluntary Sector Quarterly*, *39*, 571–587.

Liket, K. C., Rey-Garcia, M., & Maas, K. E. H. (2014). Why aren't evaluations working and what to do about it: A framework for negotiating meaningful evaluation in nonprofits. *American Journal of Evaluation*, *35*, 171–188.

Malani, A., Philipson, T., & David, G. (2003). Theories of firm behavior in the nonprofit sector. A synthesis and empirical evaluation. In E. Glaeser (Ed.), *The governance of not-for-profit organizations* (pp. 181–216). Chicago, IL: University of Chicago Press.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education*. Retrieved February 20, 2013, from http://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf

Maxwell, N., Rotz, D., Dunn, A., Rosenberg, L., & Berman, J. (2013). *The structure and operations of social enterprises in REDF's social innovation fund portfolio: Interim report*. Oakland, CA: Mathematica Policy Research. Retrieved October 16, 2014, from Mathematica Policy Research website: http://www.mathematica-mpr.com/our-publications-and-findings/publications/the-structure-and-operations-of-social-enterprises-in-redfs-social-innovation-fund-portfolio-interim-report

McDonald, B., Rogers, P., & Kefford, B. (2003). Teaching people to fish? Building evaluation capacity of public sector organizations. *Evaluation*, *9*, 9–29.

Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2012). Prosocial values and performance management theory: Linking perceived social impact and performance information use. *Governance*, *25*, 463–483.

Nonprofit Technology Network. (2012). *The state of nonprofit data*. Retrieved October 18, 2013, from http://www.nten.org/sites/default/files/data_report.pdf

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*, 443–459.

Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.

Rotz, D., Maxwell, N., & Dunn, A. (2015). *Economic self-sufficiency and life stability one year after starting a social enterprise job*. Retrieved December 22, 2015, from Mathematica Policy Research website: http://www.mathematica-mpr.com/~/media/publications/pdfs/labor/redf_yearone_rpt.pdf

Selden, S., & Sowa, J. E. (2011). Performance management and appraisal in human service organizations: Management and staff perspectives. *Public Personnel Management*, *40*, 251–264.

Snibbe, A. C. (2006, Fall). Drowning in data. *Stanford Social Innovation Review*, pp. 39–45. Retrieved April 29, 2014, from http://www.ssireview.org/articles/entry/drowning_in_data

Speckbacher, G. (2003). The economics of performance management in nonprofit organizations. *Nonprofit Management & Leadership*, *13*, 267–281.

Taut, S. (2007). Studying evaluation capacity building in a large international development organization. *American Journal of Evaluation*, *28*, 45–59.

Zimmerman, J. A. M., & Stevens, B. W. (2006). The use of performance measurement in South Carolina nonprofits. *Nonprofit Management & Leadership*, *16*, 315–327.