Performance measurement—documenting whether an organization is reaching its goals—has become a growth industry in the United States. But it is not clear what the current vogue for performance measurement has actually produced, especially for municipal governments. In fact, our research suggests that performance measurement rarely leads to improved government performance or more efficient and accountable municipal management.

Calls for the demonstration of governmental performance have become ubiquitous in the United States in recent years (Kelman, 2007; Marr, 2009; Radin, 2006; Van Dooren & Van de Walle, 2008). Citizens are demanding better results from government at a time when resource constraints are increasing, and the level of trust in government at all levels is at an historic low (National Performance Management Advisory Commission, 2010). The recent downturn of the U.S. economy and the decline in revenue for cities and states have only increased the demand for more efficient governance.

Performance measurement in government begins when public executives identify important public purposes, specify the main goals of the administration, and select the best ways to measure the accomplishment of these goals (Marr, 2009). One common motivation for measuring performance is to learn and to improve, and one of the ways an organization can improve continuously is to use performance information to learn. Continuous learning, however, requires an organization...
that is geared to learning, and also a culture that not only measures performance, but also uses the information to improve administrative performance (Behn, 2003).

The hope is that collecting performance data will create a virtuous circle: performance measurement will bolster and promote efficient performance management, by improving how an organization accomplishes its tasks and strives to achieve its goals. Better information ought to enable elected officials and managers to demonstrate success, identify poor performance, and apply the lessons to improving municipal management.

For example, if a major performance outcome of a city’s department of finance is to maximize the collection of taxes, an annual measure of the proportion of taxes due that were actually collected would be an important measure of the agency’s performance. The department can compare annual measurements and report on how the agency is doing. They can also compare that performance with the finance departments of other cities to benchmark their performance. But knowing how they are doing does not always tell them how they could be doing better. For example, suppose there is a unit that is responsible for sending out tax bills. If those bills go to bad addresses, are incorrect, or otherwise confusing to taxpayers, recipients may fail to pay the correct amount at the right time.

A good performance-managed system would measure the bill preparation and mailing units on the multiple dimensions that influence the agency’s outcome. In order to use data to improve performance, one would need to measure continuously over the year to evaluate the impact that operational changes (more timely mailings, clearer bills, improved address checking) have on the timeliness, accuracy, and proportion of payments that are collected. This would require relentless follow-up. Measuring aggregate agency performance once a year for overall accountability would never capture the right data over the appropriate timeframe to help with this kind of learning (Behn, 2003, 2006). Learning requires access to frequent, timely, operational measures and regular meetings where leadership can ask questions and follow up on the results of previous management decisions. Thus, organizational learning requires both structural and cultural elements that are mutually supportive (Moynihan & Landuyt, 2009).

Committed leadership is key (Behn, 1991; Khademian, 2002; King, Zeckhauser, & Kim, 2004; Kotter & Heskett, 1992; Kouzes & Posner, 1987; Levin & Sanger, 1994; Moynihan & Ingraham, 2004). Leaders are able to make informed decisions, develop strategy, and communicate key ideas to an organization’s members. Ultimately leadership can coordinate organizational components for cultural change necessary for creating a performance-managed operation (Behn, 1991, 2006; Kotter, 2001; Levin & Sanger, 2004; Moynihan, 2008, p. 78; Spitzer, 2007, p. 125). Leadership needs to reinforce the value of learning and setting goals, and to emphasize their relationship to enhancing performance and targeting performance deficits (Behn, 2006). An effective performance management effort requires resources, training, support, and rewards for improvement (Marr, 2009). Leadership must develop a culture where organizational units see the value of measurement

---

1 A supportive culture for a learning organization is one that is based on shared experiences, norms, and understandings that foster intelligent behavior (Senge, 1990). Specifically, motivated leaders provide strong messages about the organizational values, purposes, and support behaviors that use performance data to learn about what drives outcomes and what operation changes might improve them (Behn, 2003, 2006; Sanger, 2008b).

2 Behn (2011) sees the identification and targeting of performance deficits as a key leadership challenge for performance leadership.
and the relationship between their work and organizational outcomes. The culture needs to provide the opportunity for managerial discretion, risk-taking and some tolerance for well-conceived failures to encourage learning, innovation, and problem solving (Behn, 1991, 2006; Levin & Sanger, 1994; Moynihan & Pandey, 2010; Sanger, 2008a).

But elected leaders and their appointees come and go while public employees remain for years. Unelected officials may pay lip service to the aspirations of a mayor or a commissioner, while waiting for the next election (Larkin & Larkin, 1996; Sanger, 2008b, p. 625). For performance measurement to have a sustained impact on performance management, a municipal organization has to be forged that takes learning seriously. For a performance culture to endure, it must create and imbed values and routines that reinforce common purpose, accountability for results, and truth about performance results sufficient to produce trust. Organizational members need a clear definition of performance-driven culture (Marr, 2009, p. 241). Performance measurement alone says little about what is good or bad performance, or how to make it better.

At the same time, advocates of performance measurement assume that the collected information will facilitate accountability and transparency, by telling the public, in principle, how well its government officials are doing in accomplishing their goals and meeting public demands. As one expert writes, performance measurement promises “more sophisticated systems [that] will undergird management processes, better inform resource allocation decisions, enhance legislative oversight and increase accountability”3 (Ammons, 1995, p. 37), but it also plays a political role (de Lancer Julnes, 2008; Johnsen, 2008).

While most public officials who collect performance data claim to use it in practice, recent research that has looked beyond self-reports has raised doubts on this score (Behn, 2002; Berman & Wang, 2000; de Lancer Julnes & Holzer, 2001; Melkers & Willoughby, 2005; Moynihan, 2008). Recent research has raised similar doubts about the real impact of measurement on public opinion. Without effective outlets for active oversight,4 citizens have to rely on the professional judgment of elected officials and managers to evaluate the results of performance measurements (Ho, 2008, p. 195; Sanger, 2004). Selective reporting is a constant temptation (Dixit, 2002; Lynn, Heinrich, & Hill, 2001; Marr, 2009; Moynihan, 2008, p. 5; Propper & Wilson, 2003), and fear of exposure provides incentives for gaming, manipulating, or misrepresenting results (Marr, 2009; Radnor, 2008; Smith, 1995; Van Dooren, Bouckaert, & Halligan, 2010, p. 158).

In fact, our research shows that the current use of performance measurement for improving management performance is surprisingly limited, even in cities that have invested a great deal of time and energy in gathering the relevant information (Behn, 2002; de Lancer Julnes, 2008; de Lancer Julnes & Holzer, 2001). Instead of integrating measurement with management strategies designed to improve administrative performance, most cities use performance measurement in a limited way. Rarely do citizens demand that officials take heed of data the city has collected. As a result, performance measurement is perversely disconnected from actual decisionmaking (de Lancer Julnes, 2008).

Performance measurement continues to be widely used in a large number of American municipalities. Most cities that systematically measure performance see

---

3 Hatry (2008, p. 231) provides a complete taxonomy of all the uses of performance measurement.
4 Public participation and citizen involvement in performance measurement has become an important standard of best practice for many of the regulatory and advocacy organizations to whom we talked. See footnote 6.
it as a mark of “good government.” In this way, measurement fulfills an essential political purpose, even if the data are generally ignored by officials in practice.

Performance measurement ought to change behavior. Data should lead to questions about what drives performance and an understanding of the value of measurement for organizational learning. Results from hypothesis testing ought to alter managerial operations and improve performance. Thus, performance management should be promoted by a culture of performance measurement and best practice. This research is an effort to investigate that relationship.

**ANALYTIC APPROACH**

The research began by searching for a comprehensive list of U.S. cities likely to measure the performance of their service delivery. We generated the list of study cities by, first, a search of the literature where research had revealed cities that measure their service delivery performance, and second, contact with multiple organizations that track, support, and reward U.S. cities for their efforts. This search revealed 198 jurisdictions, for which we were able to locate data on performance measurement efforts in 190 (See list of cities in Appendix A).

For each of the 198 cities, we undertook a search for public documents that would reveal evidence of citywide performance data for any of four service areas. We chose service areas where we expected the greatest probability of measurement: police, fire, parks and recreation, and public works. We reviewed all the city documents we found on the web to uncover whether performance data were visible and reported, where it appeared, and the nature of performance measures used. We sought performance reports, city and agency budgets, strategic plans, annual reports on service efforts and accomplishments, and other public documents. From our measures, we ranked cities on the character and quality of their reporting and use of performance data, according to the best practice commonly identified by a number of organizations that promote performance measurement (See endnote 5 and Hatry, 2006).

We characterized the nature of their performance measurement effort by the quality of their measurement and distinguished cities with more developed measures of performance as more mature. We reviewed the relevant documents and collected data on the character of their measures. We then ranked the cities on the kinds of performance measures we found: whether they measured outcomes and efficiency (as well as outputs, inputs, or workload); whether the measures were benchmarked by time period, compared to comparable cities, and subdivided by precincts or other subunits; and whether they set performance targets against which reported outcomes were compared. All of these characteristics are associated with best practice

---

5 This includes the Government Accounting Standards Board (GASB), the International City/County Management Association (ICMA), the Urban Institute, the Public Performance and Measurement Research Network and the National Center for Public Performance at Rutgers, the Fund for the City of New York's Center on Government Performance and their awardees, and the Association of Governmental Accountants Service Efforts and Accomplishments Award Program.

6 All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

7 These are common service areas where the output and outcome measures are relatively straightforward and least controversial. They top the list for ICMA's completed surveys reported by jurisdictions.

8 We use the concept of maturity because measuring performance well takes time, resources, and experience, and cities that are engaging often begin with more limited efforts and add to them over time. We identify as exemplary the cities and agencies within them that have mature measurement efforts, because they exhibit important dimensions of measurement recognized as best practice.
Cities with evidence of all of these characteristics were considered mature or exemplary, but there was a great deal of variation on these and other characteristics that we tracked.

Using this approach clearly has some limitations. For example, our goal was in part to see if cities that are measuring performance well would ultimately see its value for management. Some cities are well known for managing for performance, either for a few or for many service areas (e.g., Baltimore and New York). Several of these cities did not appear on our list, most likely because they were not identified by the sources we used, or because their performance data were not visible in documents we could find on their websites. So we might have missed cities that actually do measure their performance, manage for performance, or both. Further, if the performance measurement system we could observe in any of the public documents online did not exhibit all the dimensions we thought would be most important, they did not appear in our list of the best measurement cities, and thus were not considered when evaluating which were managing for performance. As a result, we may have missed some cities with good measurement systems because we could not find the evidence online. We may also have missed some cities that manage for performance because their measurement systems did not exhibit online all the dimensions we sought.

RELATIONSHIPS BETWEEN A CITY’S CHARACTERISTICS AND ITS PERFORMANCE MEASUREMENT SYSTEM

Twenty-seven of the 190 cities we studied met all the characteristics of exemplary performance measurement systems. These were the cities we selected to study separately in our qualitative analysis. Initially, we explored many bivariate relationships for the 190 cities to see whether there were other factors that might explain variation in the character and quality of their performance measurement efforts. The vast majority of these relationships did not generate statistical significance. Those relationships we did observe were weak. As a result, we concluded that our originally planned multiple regression analysis was unnecessary. Even so, we were able to rank cities on key factors described above and their measurement efforts, and compare those that were more robust and mature to less developed systems.

We sought to learn more from semistructured interviews with city officials in those 27 cities. We requested three types of interviews: with city and agency leaders, data collectors, and operators. First, we sought mayoral or city manager leadership responsible for performance, and those charged with leadership for collecting or

9 Use of citizen surveys and regular public reporting are additional characteristics of best practice, and though important for accountability and transparency, they are not necessarily required for performance management. Thus, though we tracked these characteristics, we did not include them here.

10 We know from other research (Behn, in press; Bratton & Smith, 2001; Sanger 2008a, 2008b) that Baltimore and New York are two such cities where leadership manages for performance, but they did not appear on our list of exemplary measurement cities.

11 We found statistical significance in the relationships between household income and whether the city measured outcomes and compared itself to other jurisdictions; population size and whether it benchmarked; share of the city that is black and jurisdiction type, and whether it surveyed their citizens; and share of the city that is Latino/Hispanic and benchmarks, and a positive relationship between the share of the city that was Latino/Hispanic and whether it had performance targets. With one exception, the volunteerism and civic engagement variables we used were only found to be statistically related to whether the jurisdiction undertook a citizen survey. That finding was expected. The one exception occurred when examining the share of the city that attended a public meeting and whether that city measured outcomes; however, the difference in cities that employ outcomes versus those that do not only differed by less than 2 percentage points based on the share of the city reporting public meeting attendance.
analyzing the data, or both. In cities where particular agency efforts rather than citywide efforts were observed, we interviewed agency heads or performance measurement leadership within them, using a semistructured instrument to determine their initial motivation for measuring performance and the support for organization, resource commitments, and measurement practices of their efforts. Finally, we sought managers or operators in the agencies that used measurement to manage. From them, we sought to understand the impact of their measurement efforts on management and their operations. Specifically, we asked if and how they were managing for performance, the distribution of the results, and the impact of the performance management effort. Finally, we did follow-up interviews on the impact of the recession on their commitments. The range of interviewees in each jurisdiction allowed for a comprehensive and contextual understanding that would have been absent if only one role was sought. Leadership, for example, could seldom explain operations practices in the agencies. Further, the development, location, and management of the effort varied by jurisdiction, and the relevant players would not have been identified if there had been strict adherence in the interview design to contacting the same job titles in each city.

A semistructured instrument was used in one-hour telephone interviews with the city officials. Of the 27 cities with highly ranked performance measurement efforts, we were able to interview leadership in 24 (see list of cities in Appendix A). We repeated interviews in 10 cities one year later, after the effects of the recession began to be felt, to understand how measurement efforts were influenced by the budget environment and to follow up if we had any missing data.

COMPARING CITIES THAT MEASURE THEIR PERFORMANCE

The 190 cities that we found to be measuring performance varied by size, region, type of government, dominant party, and population composition, but the differences were only statistically significant for race, with exemplary cities having a slightly higher proportion of whites and a lower proportion of blacks. We provide descriptions below (see Table 1).

Exemplary cities were not distinguished by size or income, and while they were less likely to be in the Atlantic or Southern region, regional differences were not meaningful. We hypothesized that nonpartisan cities and cities with a professional manager might be more likely to characterize exemplary cities, but none of these differences were statistically significant. Only race mattered. Exemplary cities were more likely to have a smaller proportion of blacks than all cities. But while significant, the proportional differences are still small.

Cities also varied significantly by the characteristics and maturity of their performance measurement systems and transparent publishing of data. Some cities have citywide efforts, but for many more we found evidence for the use of performance measurement only in particular service areas, especially police (see Table 2). These

---

12 While we asked them if they were using performance data to manage, we also asked them how they used it. Where their responses did not conform to commonly understood practices central to performance management, we did not evaluate them as managing for performance, consistent with their response.

13 Similar job titles can have different responsibilities in different cities.

14 All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher’s Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

15 States, cities, and counties generally comply with federal reporting standards (Uniform Crime Reports—UCR) for crime for statistics established voluntarily, by the International Association of Chiefs of Police (http://www.fbi.gov/about-us/cjis/ucr/ucr). In releasing the information, police
Table 1. Characteristics of cities that measure performance.

| Characteristics                      | All cities  
|                                      | (n = 190) | Exemplary cities  
|                                      | (n = 27) | Pearson's chi-square or T-statistic |
| Mean population                     | 276,484  | 271,468 | 1.748 |
| Median household income              | $48,691 | $51,578 |
| Region                               |          |          |       |
| Atlantic states                      | 6%       | 0%       | 1.748 |
| Midwest                              | 20%      | 19%      | 0.018 |
| Mountain states                      | 10%      | 15%      | 0.453 |
| Pacific                              | 19%      | 30%      | 1.599 |
| South                                | 27%      | 15%      | 1.820 |
| Southwest                            | 18%      | 22%      | 0.303 |
| Political environment                |          |          |       |
| Democratic                           | 27%      | 22%      | 0.228 |
| Republican                           | 14%      | 19%      | 0.391 |
| Nonpartisan                          | 60%      | 59%      | 0.001 |
| Racial characteristics               |          |          |       |
| White                                | 71%      | 77%      | 2.128* |
| Black                                | 15%      | 7%       | -4.931** |
| Indian                               | 0.9%     | 1%       | 0.245 |
| Asian                                | 5%       | 7%       | 1.322 |
| Hispanic/Latino                      | 15%      | 17%      | 0.533 |
| Form of government                   |          |          |       |
| Council-Manager                      | 66%      | 78%      | 1.410 |
| Mayor Council                        | 32%      | 22%      | 1.023 |
| Commission                           | 2%       | 0%       | 0.505 |


a Data could not be located for all cities on every variable.
b *Significance at the 0.05 level. **Significance at the 0.01 level. Pearson’s chi-square test was used for region, political environment, and form of government. One-sample T-tests were used for racial characteristics.
c Artificial data was not available for Centennial, CO as the city was incorporated in 2001.
d d 185. Political environment could not be determined for Farmington, NM; Highland Park, IL; Maryland Heights, MO; and Raytown, MO.
e This represents the average percentage of each race across cities, not the cumulative racial percentages of all cities. Means were compared using the T-test.

differences were expected since the variables we selected to do our ranking included many of these best practices specifically.

Table 2 shows the differences between all performance measurement cities and those we identified as exemplary based on the characteristics of their measurement systems. When we could not find evidence of citywide performance measurement, we looked at four common service areas: police, fire, public works, and parks and recreation. Exemplary cities were selected by whether they exhibited the characteristics that we thought would predispose them to using data to manage. But we were also interested in seeing whether those that had such characteristics would be

departments follow guidelines set by the Office of Management and Budget and the Department of Justice (http://www.fbi.gov/about-us/cjis/ucr/ data_quality_guidelines). For the most part, agencies submit monthly crime reports using uniform offense definitions to a centralized repository within their state. The state UCR Program then forwards the data to the FBI’s national UCR Program. Thus, data collection has long been the norm for police departments. Finally, the remarkable success in New York City during the Giuliani administration (a performance management system called Compstat) has captured attention and been adopted in many cities. We thus expected that cities would be most likely to systematically collect performance data in police departments, of all service areas.
Table 2. Performance measurement maturity measures.\textsuperscript{a}

<table>
<thead>
<tr>
<th>By agency\textsuperscript{b}</th>
<th>All</th>
<th>Exemplary</th>
<th>Pearson chi-square\textsuperscript{c}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fire</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance report</td>
<td>44%</td>
<td>63%</td>
<td>0.311</td>
</tr>
<tr>
<td>Budget</td>
<td>48%</td>
<td>80%</td>
<td>0.190</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>28%</td>
<td>67%</td>
<td>0.055</td>
</tr>
<tr>
<td>Citizens survey</td>
<td>9%</td>
<td>9%</td>
<td>0.988</td>
</tr>
<tr>
<td><strong>Police</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance report</td>
<td>52%</td>
<td>80%</td>
<td>0.095</td>
</tr>
<tr>
<td>Budget</td>
<td>48%</td>
<td>80%</td>
<td>0.190</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>21%</td>
<td>88%</td>
<td>0.000**</td>
</tr>
<tr>
<td>Citizens survey</td>
<td>13%</td>
<td>0%</td>
<td>0.229</td>
</tr>
<tr>
<td><strong>Public works</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance report</td>
<td>28%</td>
<td>43%</td>
<td>0.427</td>
</tr>
<tr>
<td>Budget</td>
<td>47%</td>
<td>80%</td>
<td>0.167</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>29%</td>
<td>67%</td>
<td>0.062</td>
</tr>
<tr>
<td>Citizens survey</td>
<td>9%</td>
<td>0%</td>
<td>0.332</td>
</tr>
<tr>
<td><strong>Parks</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance report</td>
<td>26%</td>
<td>29%</td>
<td>0.878</td>
</tr>
<tr>
<td>Budget</td>
<td>45%</td>
<td>60%</td>
<td>0.530</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>40%</td>
<td>80%</td>
<td>0.018*</td>
</tr>
<tr>
<td>Citizens survey</td>
<td>12%</td>
<td>0%</td>
<td>0.252</td>
</tr>
<tr>
<td><strong>Citywide</strong>\textsuperscript{d}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance report</td>
<td>68%</td>
<td>88%</td>
<td>0.066</td>
</tr>
<tr>
<td>Budget</td>
<td>96%</td>
<td>100%</td>
<td>0.273</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>70%</td>
<td>96%</td>
<td>0.017*</td>
</tr>
<tr>
<td>Citizens survey</td>
<td>48%</td>
<td>58%</td>
<td>0.212</td>
</tr>
<tr>
<td>Has 311</td>
<td>14%</td>
<td>8%</td>
<td>0.306</td>
</tr>
<tr>
<td><strong>Type of Measure</strong>\textsuperscript{c}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inputs</td>
<td>99%</td>
<td>100%</td>
<td>0.563</td>
</tr>
<tr>
<td>Outputs</td>
<td>88%</td>
<td>100%</td>
<td>0.054</td>
</tr>
<tr>
<td>Outcomes</td>
<td>69%</td>
<td>100%</td>
<td>0.001**</td>
</tr>
<tr>
<td>Efficiency</td>
<td>34%</td>
<td>100%</td>
<td>0.000**</td>
</tr>
<tr>
<td>Quality indicators</td>
<td>57%</td>
<td>92%</td>
<td>0.002**</td>
</tr>
<tr>
<td><strong>Benchmarking</strong>\textsuperscript{f}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmarks are used</td>
<td>68%</td>
<td>88%</td>
<td>0.109</td>
</tr>
<tr>
<td>Against other time periods</td>
<td>96%</td>
<td>100%</td>
<td>0.254</td>
</tr>
<tr>
<td>Against other jurisdictions</td>
<td>70%</td>
<td>96%</td>
<td>0.017*</td>
</tr>
<tr>
<td>Use subjurisdictions</td>
<td>48%</td>
<td>58%</td>
<td>0.038*</td>
</tr>
<tr>
<td><strong>Targets</strong>\textsuperscript{g}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Targets are used</td>
<td>52%</td>
<td>100%</td>
<td>0.000**</td>
</tr>
<tr>
<td>No. of service areas</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most</td>
<td>37%</td>
<td>71%</td>
<td>0.000**</td>
</tr>
<tr>
<td>Some</td>
<td>7%</td>
<td>21%</td>
<td>0.038*</td>
</tr>
<tr>
<td>Few</td>
<td>13%</td>
<td>8%</td>
<td>0.047</td>
</tr>
<tr>
<td>None</td>
<td>44%</td>
<td>0%</td>
<td>0.000**</td>
</tr>
</tbody>
</table>

more likely to report their performance to citizens through performance reports, and elicit citizen’s preferences and needs through the use of citizen surveys and 311 reporting efforts. With the exception of police, exemplary cities were more likely to have a strategic plan. This is not surprising since strategic planning identifies goals and targets and presumably measures their achievement over time.

Exemplary cities differ from all cities on the character of the measures they use, whether they benchmark them, set targets for performance, and report on them...
Table 2. Continued.

<table>
<thead>
<tr>
<th>By agency&lt;sup&gt;b&lt;/sup&gt;</th>
<th>All</th>
<th>Exemplary</th>
<th>Pearson chi-square&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frequent reporting&lt;sup&gt;h&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance report</td>
<td>5%</td>
<td>5%</td>
<td>.830</td>
</tr>
<tr>
<td>Strategic plan</td>
<td>0%</td>
<td>0%</td>
<td>-</td>
</tr>
</tbody>
</table>


<sup>a</sup>Total possible of \( n = 190 \) for “all” cities, \( n = 27 \) for “exemplary” cities.

<sup>b</sup>All: Performance report: fire \( n = 70 \); police \( n = 83 \); public works \( n = 67 \); parks \( n = 65 \). Budget: fire \( n = 32 \); police \( n = 32 \); public works \( n = 35 \); parks \( n = 34 \). Strategic plan: fire \( n = 56 \); police \( n = 56 \); public works \( n = 58 \); parks \( n = 68 \). Citizens survey: fire \( n = 78 \); police \( n = 80 \); public works \( n = 79 \); parks \( n = 78 \). Exemplary: performance report: fire \( n = 8 \); police \( n = 10 \); public works \( n = 7 \); parks \( n = 7 \). Budget: fire \( n = 4 \); police \( n = 5 \); public works \( n = 5 \); parks \( n = 5 \). Strategic plan: fire \( n = 6 \); police \( n = 8 \); public works \( n = 6 \); parks \( n = 10 \). Citizens survey: fire \( n = 11 \); police, public works, and parks \( n = 10 \).

<sup>c</sup>Categories with * are significant at the 0.05 level, ** at the 0.01 level.

<sup>d</sup>All: Performance report \( n = 188 \); Budget \( n = 190 \); Strategic plan \( n = 188 \); Citizens survey \( n = 189 \); Has 311 \( n = 190 \). Exemplary: \( n = 27 \) on all measures.

<sup>e</sup>All: Inputs, Outputs, and Outcomes \( n = 190 \); Efficiency and Quality \( n = 189 \). Exemplary: \( n = 27 \) on all measures.

<sup>f</sup>All: Benchmarks are used \( n = 186 \); Against other time periods and jurisdictions \( n = 178 \); Use subjurisdictions \( n = 179 \). Exemplary: \( n = 27 \) on all measures.

<sup>g</sup>All: Targets are used \( n = 189 \); No. of service areas \( n = 180 \). Exemplary: \( n = 27 \) on all measures.

<sup>h</sup>Represents the percent of public reports conducted more frequently than annually. All: Performance report \( n = 139 \); Strategic plan \( n = 75 \). Exemplary: Performance report \( n = 24 \); Strategic plan \( n = 16 \).

frequently. Most of these are statistically significant. This was expected since they were ranked and selected because their measurement practices were characterized by best practice.

**MATURE PERFORMANCE MEASUREMENT SYSTEMS CAN CREATE THE CAPACITY FOR PERFORMANCE MANAGEMENT**

We hypothesized that cities whose performance measurement systems were more mature would be more likely to use performance data to manage. Mature measurement efforts attempt to define metrics that reflect on desired service outcomes such as safety, street cleanliness, and road condition. Further, they seek to determine service efficiency by evaluating the units of services they produce for the resources expended. Because outcomes often take longer to be influenced by changes in operations than, for example, outputs, mature systems seek to measure intermediate progress toward outcomes, and those usually involve measuring outputs, too. As we reported earlier, many jurisdictions we examined simply report on workload inputs (numbers of workers or person-hours expended). Mature systems seek to benchmark performance against other jurisdictions and time periods as well as set targets from period to period. More importantly, the potential to learn improves immensely by collecting performance data on the subjurisdictional level, by neighborhood, precinct, office, or service area. Looking for outliers allows managers to hypothesize what explains variation within their operations. Looking at operations at this level and comparing good performers to bad performers often reveals best practices that can be adopted across the agency citywide. Timely operational data are key to learning what management changes lead to changes in performance. We also saw the value of target setting in mature systems as a potential motivator...
consistent with the requirements for performance management.\textsuperscript{16} We were interested in the potential that any performance measurement effort had to use data in these ways. And if they did, we recognized the increased likelihood that they would be managing for performance, or at least have more capacity to do so.

Cities varied significantly on the characteristics of their performance measurement efforts. Most cities (68 percent) were found to measure citywide performance efforts, even if most of them did not exhibit all of the characteristics we associated with a mature system (See Table 2). Even for those whose performance measurement was mature, few report their performance publicly more regularly than annually or collect and report data on a subjurisdictional level. The great majority of cities do not collect data often or disaggregate what they do collect in ways likely to induce hypotheses about what drives differential performance. Indeed, most cities simply do not measure the differential performance of operating units within their agencies (See Table 2).

\textbf{WHAT CHARACTERIZED THE CITIES WITH MATURE SYSTEMS?}

Cities with mature systems were, in general, not very different from the larger sample (see Table 1). They were more likely to have a manager-council form of government and more likely to have a nonpartisan system, but these differences are not statistically significant. The cities where we conducted interviews were slightly more likely to be from the Southwest and the Pacific coast, but again, these differences were not statistically significant. One statistically significant difference, however, is that mature cities or their agencies were more likely to have a strategic plan. Strategic plans call for an agency to identify its important public purposes and to lay out a plan to accomplish them. Performance information that is generated through strategic planning and resulting performance measurement routines can connect information to decision venues (Moynihan, 2008, p. 5). Thus, we would expect the existence of strategic planning to be associated with mature performance measurement efforts and perhaps be more likely to support performance management.

In almost all the cities with the most mature performance measurement efforts, collection and distribution of performance data reside in the city’s budget office or is formalized as part of the budget process; in a few cities, including Minneapolis, Minnesota; Tacoma, Washington; and Westminster, Colorado, the effort resides instead, or in addition, in the city manager’s office, and in one city, in the auditor’s office. Most cities said they were motivated to invest in performance measurement for accountability, that is, to demonstrate their openness to scrutiny about their operations. Some were motivated to report to the public (to report to citizens in Alexandria, Virginia, and Ankeny, Iowa), and others for internal purposes. They also frequently reported that they were motivated by a desire to adopt a best practice or to use a cutting edge approach.\textsuperscript{17} As an official from Westminster, Colorado, reported, “sophisticated governments are engaged in performance measurement.” An official in Corvallis, Oregon, said they began because “everybody was doing it.” Some, however, saw the potential for improved management at the outset. The

\textsuperscript{16} Behn (2011) identifies recognizing performance deficits as critical for improving performance. Targets are an explicit recognition that an operation can do better.

\textsuperscript{17} Performance measurement is promoted in professional organizations for public managers, where some reported taking workshops or hearing presentations; consultants abound and articles and examples of promising practices are reported in widely read publications such as \textit{Governing Magazine} or \textit{Public Executive}. 

\textit{Journal of Policy Analysis and Management}  DOI: 10.1002/pam
Published on behalf of the Association for Public Policy Analysis and Management
promotion of Compstat in New York City and Citistat in Baltimore, Maryland, had popularized the notion and methods of managing for performance, and that motivated some jurisdictions to explore measurement’s value. A councilman in Long Beach, California, we learned, persuaded the council and city manager to adopt a performance measurement system on that basis.

As in Long Beach, California, performance measurement was typically the brainchild of a key city leader, usually the city manager or mayor, but sometimes the city council or some champion within it. In a few cases, citizen efforts supported adoption, such as in Ankeny, Iowa, where a number of cities in the state had funding from the Sloan Foundation to undertake citizen-centric performance measurement. The budget office, however, remained the champion of the effort inside city government.

A variety of factors are thought to influence the development, maintenance, and use of performance measurement (de Lancer Julnes, 2008; Kelly, 2002; King, Zeckhauser, & Kim, 2004; Moynihan, 2008; Moynihan & Landuyt, 2009). Willingness and the ability to undertake the effort generally derives from leadership that understands and communicates the value, allocates the resources, and insures system integrity. Organizational culture is also important to create and support a learning organization that invests in performance data and can see the value of its use (Khademian, 2002; Moynihan, 2008; Sanger, 2008b). While a citywide culture of measurement and evidence-based decisionmaking is occasionally apparent—and we observed a few examples, such as Charlotte, North Carolina, and Tacoma, Washington—rarely is the effort uniform across agencies. More often, we observed pockets of real energy and other areas of city government with little enthusiasm.

Given the importance of leadership in developing robust performance measurement systems, we expected to find that active and enduring leadership that supports performance measurement efforts would be associated with maintenance of effort; the loss of that leader correlates with waning investment. Further, in cities with existing performance measurement systems, changes in leadership were often associated with changes in investment in performance measurement. Long Beach, California, is instructive. With a long history of performance measurement, a committed city council and city manager, and a performance management office that originally resided in the city manager’s office, a former budget official and performance management staffer we interviewed reported a declining commitment. The effort has been rolled into the budget office and “the new city manager is not as fervent as the last one, and the city council is not as ‘sold’ either,” she said.

Finally, the fiscal crisis of U.S. cities associated with the recent national recession has depleted local treasuries around the country, resulting in reductions in staff and personnel. In annual surveys by the National League of Cities, city finance officers reporting on the condition of their cities confirmed increasing rates of hiring freezes, salary and wage reductions, layoffs among municipal employees, and service cuts in their jurisdictions from 2009 through 2010 (National League of Cities, 2009, 2010). Indeed, in 2009, nine of 10 reported that they were making spending cuts and predicted further cuts in 2010.

While we may theorize that performance measurement is particularly well suited to a resource-constrained environment when the need to demonstrate that the results of investments is emphasized, our interviews painted a very different picture. Indeed, in our follow-up interviews, officials in cities with a history of investment in measurement frequently reported reduced efforts. The program management

---

18 A citywide, data-driven performance management system for all city agencies based on the New York City police model including frequent performance reporting, accountability for results, setting targets based on identified performance deficits, regular meetings, and relentless follow-up and feedback (see Behn, 2008).
coordi-80329ator in the Office of Management and Budget in Alexandria, Virginia, for example, reported that they “dropped their quarterly performance report last year because of other pressures, and cancelled this year’s performance report entirely.”

The internal auditor in Reno, Nevada, reported that interest had dwindled, particularly once the prior city manager left, and they “dropped out of ICMA [International City/County Management Association] to save the annual $5,500 fee.” Scottsdale, Arizona, reported that they also dropped out of ICMA; a senior advisor for results tracking and reporting remarked that performance measurement “could be the casualty of the budget climate.” A follow-up interview with the city manager of Corvallis, Oregon, revealed pressure to cut $2.5 million from a $118 million budget. He reported delays in posting “this year’s report card . . . I would not say that there have been cuts to monies allocated to performance measurement, but simply that we have fewer staff resources to absorb ongoing performance measurement work on a timely basis.” Urbandale, Iowa, reported that “the data generated had become overwhelming. We are moving even more toward an annual report format.”

The experience of Lynnwood, Washington, also illustrates larger trends we observed. Early in 2000, a councilman proposed and spearheaded passage of a “performance management results act” that mandated the collection and dissemination of performance-related data. Consultants hired to help in system development began with two service areas: parks and recreation and courts. Agencies collected and analyzed their own data, which was challenged by the budget office only when it aroused curiosity.19 While expanding the system to additional agencies, the budget coordinator began to deemphasize its use. Ironically, the councilman who originally spearheaded the effort was elected mayor and now pays it little attention. While we can only speculate, his neglect of a strongly held commitment following his election may suggest his lack of appetite for the increased scrutiny and the political risks that performance measurement and management would pose. Now, leadership has waned almost entirely, the effort has no teeth, and the value and potential of its use for management has been largely abandoned. The only support for the effort is carried by a part-time budget coordinator, who reported that she is charged largely with collecting ICMA benchmarking data. “The departments,” she reported, “don’t even have their own budget analysts.” We have found that 10 of 24 cities with initially robust efforts now report that efforts are waning.

FROM MEASUREMENT TO MANAGEMENT

Performance measurement efforts, however initially popular, are inherently fragile. Even in places where the performance measurement system was robust, continued support is subject to changing leadership and priorities; further, these efforts sometimes fall victim to the politics of transparency where the risk of exposure to failure can create internal resistance to accountability. Without a strong culture of measurement where employees deep in the organization see and value the contribution of measuring performance, internal sabotage is possible and often predictable (Radnor, 2008; Sanger, 2008a; Smith, 1995; Van Dooren, Bouckaert, & Halligan, 2010, p. 158).

Fear of exposure to failure may create pressures to distort or even corrupt the data, so auditing of data is crucial. However, when we asked how data reliability is ensured, only two cities actually reported that they audit (Scottsdale, Arizona, and Sunnyvale, California). Another reported that they did “spot checks” (Olathe,

19 It appeared that no formal auditing effort was in place. This would seem to encourage data manipulation to insure that nothing reported looked curious.
Kansas). An internal auditor in Reno, Nevada, when asked if the city's performance data were audited, reported that several years ago, a deputy at the fire department who had been responsible for submitting their ICMA data was fired. They went in and checked his files to pick up his methodology and discovered that he had simply been making up his numbers all along.\textsuperscript{20}

Due to these inherent liabilities, we see the survival of robust performance measurement systems in far fewer jurisdictions than we expected. Even so, we sought to determine how many of them had moved from performance measurement to performance management. Here, we will distinguish in our discussion those who claimed to be using measures to manage, from those who we consider to be truly employing performance management. There were more cities whose officials said they use performance data to manage than those for whom we could find evidence that they did. For example, if they reported that they did not collect, review, and analyze data more frequently than quarterly, it was our judgment that they were not managing for performance, because they could not be using timely operational data to test hypotheses about what is driving changes in performance. Only eight of 24 cities claimed to be analyzing data as often as monthly; only three reported that they met weekly to analyze the data. Finally, only five of 24 thought they had sufficient resources committed to performance management.

We were also interested in exploring whether there was a learning orientation at the city or agency level. This was a condition we thought was essential for building and sustaining a performance management system. We explored in our interviews how data were used by managers and executives to promote learning in the organization. Typically, a learning environment would be reflected in what happened in regular meetings where managers are required to report on their performance and the impacts of operational changes they make in response to it from meeting to meeting.

We asked all 24 jurisdictions if they used data to manage. While eight cities reported that they analyzed data monthly, only seven of 24 definitively responded that they used performance data to manage. Several others claimed to be moving in that direction, but among our high-performing cities, only Charlotte, North Carolina; Corpus Christi, Texas; Minneapolis, Minnesota; and perhaps Long Beach, California, have the necessary culture and preconditions to support performance management according to our parameters. Some jurisdictions talked about performance management as an “evolutionary process.” Others, such as Las Cruces, New Mexico, and Fort Collins, Colorado, see the value of managing for performance and have it as a goal, but are not there yet. As the budget analyst we interviewed in Fort Collins reported when asked if anyone is sitting down on a weekly or monthly basis to look at the data, “we are trying; it’s our goal, but I do not know anyone who really does it.”

A line manager we interviewed in the street maintenance division in Charlotte, North Carolina, captured the necessary learning and management process, but we failed to see this in most of the other cities. When we asked the deputy street superintendent about their process for using performance data to manage, he described it this way

We meet once a month on a regular basis . . . in the meeting we talk and compare our three maintenance facilities with each other and to other standard outputs in the field.

\textsuperscript{20} Marr (2009) provides data from the largest performance management study internationally conducted by the Advanced Performance Institute that found evidence that 68 percent of public sector organizations occasionally make up or fabricate performance data. Thus without audits, the value of the data always remains suspect.
We are looking for data that stands up and then looking at why it stands up. It can be on both sides of the coin, areas that we are doing something better or worse; we compare our facilities and between them with standards in the industry. We aren’t looking to find failures in each facility; we want to improve. We have a corporate view approach; we are looking if someone is doing better and trying to learn why, how, what are they doing? We use this as a learning process and we try to capitalize on success.

Corpus Christi, Texas, another city that manages for performance, is using a balanced scorecard (Kaplan & Norton, 1996) and developing a performance management effort where each agency’s performance scorecard rolls up into a city scorecard. They began their measurement effort more than 10 years ago, and there has been enormous personal investment by the city manager, use of outside consultants, and even retreats with the city council. Support is reported to be strong for both performance measurement and management. When queried about how he uses performance measures to manage, the assistant city manager described an electronic data system that uploads data on a daily basis, with reminders placed on managers’ computer screens to check their data daily. Reports are run monthly, and the city manager has monthly meetings with agencies to review the data. Agency heads meet with their supervisors prior to the monthly meetings. All measures have set targets, but they are realistic and set in collaboration with the departments. “The environment is supportive of learning, so there are no sanctions used for missing a target . . . , where instead, missing a target provokes questions and a learning opportunity,” said the interim assistant city manager. She reported that the data are disseminated only internally, although the aggregate agency performance findings appear in the annual report and budget documents. This level of scrutiny and use of performance data to manage performance, she reported, has only developed over the last two years, even though the effort began over 10 years ago. A culture of measurement and evaluation reflects itself in the data collection, analysis practices, and expectations of managers and line workers throughout the organization. Further, learning cultures subject these data to scrutiny by regularly asking questions. These cultures are more likely to generate hypotheses about what drives performance and to encourage hypothesis testing and innovation in operations to improve it.

When asked about the lessons from their efforts, most of our high-performing cities emphasized the need to go slow and to move toward performance management stepwise. But few, as we have seen, are using performance measurement to manage, and even in efforts with a long history, leadership, predisposing cultures, and significant investment, systemic and imbedded performance management routines were hard to find.

**DISCUSSION AND CONCLUSION**

Our findings were disappointing but instructive. We hypothesized that having a more mature measurement system might lead to efforts to manage for performance. Our analysis found no clear link between demographic and governance characteristics of cities and the robustness of performance measurement systems. More troubling, even among jurisdictions with exemplary performance measurement efforts, we found only a few who used performance data to manage.

Few analysts have evaluated these relationships. Many, however, have inquired about the use of performance measurement (de Lancer Julnes, 2008). But whether exemplary performance measurement inevitably leads to performance management and performance improvement has not been studied directly. It is an assumption of many of the performance measurement proponents that better measurement will
lead to better management. Many researchers, good government groups, and advocates see the value of measuring performance as a desirable end in itself, promoting accountability and improving transparency to citizens. If outcomes and efficiency are also measured, public officials have some evidence of program performance that is useful to justify budget requests, and the literature has shown increasing use of performance data for that purpose. But performance data, even appropriate and sophisticated outcome measures, do not tell why the outcome occurred. A performance management system is necessary to generate the management structures and the kind of useful data that allows organizations to learn what operational elements drive performance. We have seen that this requires leadership, a learning organization, and a culture that both supports innovation and experimentation, and tolerates well-conceived failures (Sanger 2008b). These organizations collect the right kind of data in a timely fashion, analyze the data, meet regularly, and follow-up relentlessly. As we have seen in Charlotte, North Carolina, and Long Beach, California, for example, managers have the right kind of data and management structures to do this.

Why do good performance measurement systems so rarely lead to robust performance management systems? For the most part, cities do not use performance measurement to manage, and those that do are vulnerable to obstacles from within agencies and from external stakeholders. The effects of shifting political and agency leadership, unstable resources, the changing political fortunes of champions, resistant public employees, and varying levels of citizen engagement tend to alter forces that sustain efforts (Bourdeaux, 2006; de Lancer Julnes & Holtzer, 2001; Moynihan & Ingraham, 2004; Sanger 2008a; Van Dooren, Bouckaert, & Halligan, 2010). Interviews with those we thought most likely to have embraced performance management provide support for these views.

This research has looked at cities that measure their performance; the evidence came directly from what we found in their public documents. We looked at citywide efforts as documented in city reports and budgets, and where they were absent, at agency documents. This approach may have missed everything happening on the ground and perhaps our methodology biased our results. What is publicly reported or available may not fully capture what managers collect and use. Some of the cities where we did not find all the elements of a mature system on the web may be using performance management approaches in their agencies and may be managing for results, but these may not have been online or derived from an exemplary performance reporting system. Further, some have argued that the use and value of performance measurement may be more indirect and thus less observable (de Lancer Julnes, 2008).

There can be mandates for performance reporting, but “the other factors are difficult to legislate (culture, leadership, worker beliefs, the nature of clients and stakeholders)” (Moynihan, 2008). The vigor with which performance measurement is embraced and exploited for management is subject to a complex set of conditions, many of which are beyond the control of legislatures, city mayors, managers, and agency heads (Moynihan & Pandey, 2005). Even so, evidence continues to support the notion that success is most evident at the agency level when leadership is strong and stable, a learning culture is built with imbedded routines, and managerial discretion and external political support is available. Even within our most exemplary jurisdictions, the level of political, resource, and managerial investment varies by agency, over time, and with

---

21 These are key elements of the successful PerformanceStat models that are growing in popularity (Behn, 2008).
changing agency and city leadership. The ability to bring about performance-driven change is thus highly constrained, subject to multiple and powerful obstacles, and more likely to be successful where more organizational autonomy is possible.

Great expectation has accompanied the performance movement both in the United States and internationally (Kettl, 2000; Kelman, 2007; Marr, 2009; National Performance Management Advisory Commission, 2010; Van Dooren, Bouckaert, & Halligan, 2010). Significant accomplishment can be identified (Behn, n.d.; Ingraham, Joyce, & Donahue, 2003; Moynihan, 2008; Sanger, 2008a, 2008b). An abundance of “good government” groups continue to promote, support, and fund best practice efforts for governments at all levels to measure and report their performance, and enormous amounts of resources are devoted to these activities. Most of those we interviewed cited these influences as the compelling reasons that launched their systems. But we have also seen that even cities with the best intentions have failed to fully realize the promise of these methods. Measuring performance and improving it are two different activities, and our research highlights the many questions that remain. Whether we have oversold the performance movement is clearly one question (Van Dooren, Bouckaert, & Halligan, 2010). If we cannot link the performance measurement movement to improved management, and if we cannot demonstrate the performance returns to performance management, we are by necessity weakening the case for continued investment and strengthening the case for improved research on the cost effectiveness and value of these efforts.

We know little about the returns to performance-managed systems and how much they actually work to improve performance.22 Thus, as we see reduced investment in both performance measurement and management with contracting budgets, even in our exemplary cities, we may need to ask ourselves how much it matters and how much effort should continue to be placed in promoting it.23 More and better research will be necessary to answer those questions, but in the face of the evidence to date, our expectations should be modest.

MARY BRYNA SANGER is Deputy Provost and Senior Vice-President of Academic Affairs, Professor of Urban Policy and Management, New School, 66 West 12th Street, New York, NY 10011.

ACKNOWLEDGMENTS

This research was funded by the Einhorn Research Award from the Academy for Governmental Accountability of the Association of Government Accountants. I am grateful to a group of research assistants and interviewers who helped over several years with the data collection effort: Margaret Goodwin, Jackie Moynahan, Kelly Johnstone, Sierra Stoneman-Bell, Andrew French, and Roy Abir.

REFERENCES


22 None of the jurisdiction we interviewed could tell us what the performance returns were from their investments, no less the resource savings. No one was even asking those questions.
23 This is not to say that measuring performance and reporting on it does not have an important value for accountability and transparency. That may be enough to justify the scope and expense of the effort. Here, we are addressing the issue of how much the investment actually results in improved performance.


Behn, R. D. (2011, November). PerformanceStat is a leadership strategy, not a model or a system: Or why MimicStat can't really work. Paper prepared for the Annual Research Conference of the Association for Public Policy Analysis and Management, Washington, D.C.


Kelly, J. M. (2002). If you only knew how well we are performing, you'd be highly satisfied with the quality of our service. National Civic Review, 91, 283–292.


Appendix A: Performance Measurement Cities

1. Abilene, TX
2. Albany, GA
3. Albuquerque, NM
4. Alexandria, VA
5. Alpharetta, GA
6. Anchorage, AK
7. Ankeny, IA
8. Ann Arbor, MI
9. Arlington, TX
10. Asheville, NC
11. Auburn, AL
12. Aurora, CO
13. Austin, TX
14. Baltimore, MD
15. Bellevue, WA
16. Billings, MT
17. Bothell, WA
18. Bowling Green, KY
19. Bridgeport, CT
20. Broken Arrow, OK
21. Bryan, TX
22. Carlsbad, CA
23. Casper, WY
24. Centennial, CO
25. Chandler, AZ
26. Charlotte, NC
27. Charlottesville, VA
28. Chattanooga, TN
29. Chesapeake, VA
30. Cleveland, OH
31. College Station, TX
32. Collinsville, IL
33. Colorado Springs, CO
34. Columbus, OH
35. Concord, NH
36. Coral Springs, FL
37. Corpus Christi, TX
38. Corvallis, OR
39. Dallas, TX
40. Danville, VA
41. Davenport, IA
42. Dayton, OH
43. DeKalb, IL
44. Denver, CO
45. Des Moines, IA
46. Detroit, MI
47. District of Columbia
48. Duluth, GA
49. Duncanville, TX
50. Durham, NC
51. East Providence, RI
52. Elgin, IL
53. Englewood, CO
54. Eugene, OR  
55. Evanston, IL  
56. Farmers Branch, TX  
57. Farmington, NM  
58. Fishers, IN  
59. Fort Collins, CO  
60. Fort Worth, TX  
61. Franklin, TN  
62. Fresno, CA  
63. Fullerton, CA  
64. Gainesville, FL  
65. Gladstone, MO  
66. Goodyear, AZ  
67. Greensboro, NC  
68. Hampton, VA  
69. Harrisonburg, VA  
70. Hartford, CT  
71. Hayward, CA  
72. Henderson, NV  
73. Highland Park, IL  
74. Hillsboro, OR  
75. Houston, TX  
76. Indianapolis, IN  
77. Irving, TX  
78. Jacksonville, FL  
79. Johnson City, TN  
80. Kansas City, MO  
81. Keller, TX  
82. Kennesaw, GA  
83. Kennewick, WA  
84. Kent, WA  
85. Kirkland, WA  
86. Kirkwood, MO  
87. Laredo, TX  
88. Las Cruces, NM  
89. Las Vegas, NV  
90. Lauderhill, FL  
91. Long Beach, CA  
92. Longmont, CO  
93. Longview, TX  
94. Los Angeles, CA  
95. Loveland, CO  
96. Lynchburg, VA  
97. Lynnwood, WA  
98. Marietta, GA  
99. Maryland Heights, MO  
100. McAllen, TX  
101. McHenry, IL  
102. Mesa, AZ  
103. Milwaukee, WI  
104. Minneapolis, MN  
105. Moorhead, MN  
106. Mt. Lebanon, PA  
107. New London, CT  
108. New York City, NY  
109. Newport News, VA  
110. Norfolk, VA
111. North Las Vegas, NV
112. North Richland Hills, TX
113. Northglenn, CO
114. O'Fallon, IL
115. Oakland, CA
116. Oklahoma City, OK
117. Olathe, KS
118. Orlando, FL
119. Overland Park, KS
120. Palm Coast, FL
121. Palo Alto, CA
122. Park Ridge, IL
123. Pasco, WA
124. Peachtree City, GA
125. Peoria, AZ
126. Peoria, IL
127. Philadelphia, PA
128. Phoenix, AZ
129. Plano, TX
130. Plant City, FL
131. Pocatello, ID
132. Portland, OR
133. Portsmouth, VA
134. Raytown, MO
135. Redwood, CA
136. Reno, NV
137. Renton, WA
138. Richland, WA
139. Richmond, VA
140. Rock Hill, SC
141. Rockford, IL
142. Rome, GA
143. Rowlett, TX
144. Sacramento, CA
145. Salem, OR
146. Salisbury, NC
147. Sammamish, WA
148. San Antonio, TX
149. San Diego, CA
150. San Francisco, CA
151. San Jose, CA
152. Sandy Springs, GA
153. Santa Fe, NM
154. Santa Monica, CA
155. Sarasota, FL
156. Savannah, GA
157. Scottsdale, AZ
158. Seattle, WA
159. Shawnee, KA
160. Sherman, TX
161. Shoreline, WA
162. Sioux City, IA
163. Sioux Falls, SD
164. Smyrna, GA
165. Sparks, NV
166. St Charles, IL
167. St. Cloud, MN
168. Sterling Heights, MI
169. Sunnyvale, CA\textsuperscript{a}
170. Tacoma, WA\textsuperscript{a}
171. Tallahassee, FL
172. Tigard, OR
173. Tucson, AZ
174. Tyler, TX
175. University Place, WA
176. Urbandale, IA\textsuperscript{a}
177. Vallejo, CA
178. Vancouver, WA
179. Virginia Beach, VA
180. Vista, CA
181. West Hartford, CT
182. West Jordan, UT
183. Westminster, CO\textsuperscript{a}
184. Winchester, VA
185. Winston-Salem, NC
186. Winter Garden, FL
187. Woodbury, MN
188. Worcester, MA
189. Yakima, WA
190. Yuma, AZ

\textsuperscript{a}Cities that we ranked as exemplary.
\textsuperscript{b}Cities that we ranked as exemplary, but did not interview.