The Wisdom of Crowds: Learning from Administrators' Predictions of Citizen Perceptions
Author(s): Theodore H. Poister and John Clayton Thomas
Source: *Public Administration Review*, Vol. 67, No. 2 (Mar. - Apr., 2007), pp. 279-289
Published by: Wiley on behalf of the American Society for Public Administration
Stable URL: http://www.jstor.org/stable/4624564
Accessed: 24/02/2015 13:43

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Wiley and *American Society for Public Administration* are collaborating with JSTOR to digitize, preserve and extend access to *Public Administration Review*.

http://www.jstor.org

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Theodore H. Poister
John Clayton Thomas
Georgia State University

The Wisdom of Crowds: Learning from Administrators' Predictions of Citizen Perceptions

Essays on Citizen Participation and Governance

Public agencies use surveys to solicit feedback from citizens and targeted customer groups, but many experts question whether the results of these surveys are valuable. This paper explores how a recent innovation in citizen surveys—asking public administrators to predict how citizens will respond to survey questions—may be used to increase that value and, at the same time, provide additional data of interest on its own account. The innovation is explored through two surveys: (1) a public opinion poll of Georgia residents conducted by the authors for the Georgia Department of Transportation in January 2004, and (2) a brief survey of that agency's administrators asking for their predictions of public opinion. The prediction process appeared to increase the agency's interest in the resident survey. The findings document the frequent superiority of groups—what Surowiecki terms "the wisdom of crowds"—over individuals in predicting public opinion.

Agencies are increasingly employing survey research to solicit feedback on public services from both general and targeted samples of citizen-consumers. Yet, in the present ongoing climate of tension between customer feedback and managerial concerns for objective measures of performance (Kelly 2005), there is little agreement on the extent to which such surveys add value to public managers' understanding of program or agency performance. Both scholars and public managers have questioned the worth of these surveys (Kelly and Swindell 2003).

This paper explores how a recent innovation in citizen surveys may be used to increase the value of those surveys while, at the same time, providing additional data of interest on its own account. The innovation, pioneered by one of the current authors and a colleague several years ago (Melkers and Thomas 1998), entails asking public administrators to predict citizen opinions in advance of receiving the survey results. This paper will examine what can be learned from using this innovation in tandem with a survey conducted by the Georgia Department of Transportation (GDOT) in early 2004 to gauge Georgia residents' opinions about transportation issues.

The Value of Customer Feedback

Confidence in the value of customer surveys depends in part on how well they are designed and administered. As more public agencies have committed to gathering systematic customer feedback, methodological approaches for surveys have become increasingly sophisticated (Van Ryzin 2004; Van Ryzin et al. 2004), and other resources have expanded to help agencies and consultants develop and conduct effective surveys (Hatry et al. 1998; Miller and Kobayashi 2000).

But the issue of added value also hinges on the extent to which managers believe that citizen and customer surveys provide them with useful information. Objections on this score originally arose from a substantial stream of research over the past 25 years that has shown only weak or inconsistent correlations between citizens' ratings of or satisfaction with services, on one hand, and more "objective" indicators of program performance used by agencies, on the other hand (Brown and Coulter 1983; Kelly and Swindell 2002; Parks 1984; Stipak 1979; Swindell and Kelly 2000). That objection has lost some of its force in the face of arguments that citizen or customer satisfaction measures provide a different perspective on program performance and service delivery, complementing the more tangible measures of processes, outputs, and outcomes that are typically monitored by administrators (Licari, McLean, and Rice 2005).

> … surveys too often tell administrators what they already know. Many public managers believe they have a good read on customer satisfaction with their organizations based on voluntary compliments, complaints, and other informal contacts.

There remains, though, the objection—familiar to almost anyone who has conducted a survey for a government agency—that surveys too often tell administrators what they already know. Many public

Theodore H. Poister is a professor of public administration in the Andrew Young School of Policy Studies at Georgia State University. His research interests focus on strategic planning and management, performance measurement, and customer service in the public sector, and much of his work addresses these issues in transportation agencies. His book *Measuring Performance in Public and Nonprofit Organizations* was published by Jossey-Bass in 2003.
E-mail: tpoister@gsu.edu.

John Clayton Thomas is a professor in the Department of Public Administration and Urban Studies, Andrew Young School of Policy Studies, Georgia State University. His research focuses on how citizens connect with governments and how those connections can be improved. His research has appeared in such journals as *Public Administration Review, Journal of Public Administration Research and Theory, Administration & Society,* and *Urban Affairs Review.*
E-mail: jcthomas@gsu.edu.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

managers believe they have a good read on customer satisfaction with their organizations based on voluntary compliments, complaints, and other informal contacts. If these managers are correct in assuming that they know how customers perceive their programs, customer surveys may not, in fact, provide them with much additional information. In this regard, when survey findings are reported to agency administrators, it is not unusual to hear one say, "Well, no surprise there." When that happens, the survey findings may not be taken seriously, and the very idea of surveys may lose value in the eyes of agency clients.

The Potential Value of a Survey Innovation

Asking administrators to predict survey results may help address this problem. As Michael Quinn Patton has argued, "Having stakeholders speculate about actual results prior to seeing the real data… helps to increase interest in seeing the actual results" (1986, 250). Increased interest may be the necessary first step toward increased value for survey results.

Administrators who are required to guess citizen opinions in advance may also discover more surprises in the data than they anticipated. Having to predict public opinion can force administrators to come down on one side or another of a question for which there may be two or more seemingly "obvious" yet contrasting answers. Being asked to predict the answer in advance can both alert administrators to the many possible answers *and* force a choice among them, prompting surprise and more respect when the actual answer is revealed.

Obtaining administrators' predictions of survey findings also provides a test of whether public administrators really *do* have a good idea of how customers view their agencies' programs and service-delivery processes. If these managers are correct in believing that they know how customers perceive their programs, customer surveys may not provide them with additional valuable information and perhaps should be deemphasized as an administrative tool. However, as Patton observes,

> Program staff… sometimes argue that they don't need formal evaluations because they know their clients, students, or program participants so well that evaluation findings would just confirm what they already know. I've found that when staff commit their guesses to paper in advance of seeing actual results, the subsequent comparison often calls into question how well some staff do know what is happening in the program. (1986, 251)

When this is the case and managers become aware of the inaccuracy of their perceptions, interest in citizen and customer surveys may grow as managers feel more uncertainty about what those surveys will reveal.

Assuming for the moment that all or most administrators do *not* accurately perceive public opinion on most relevant issues (i.e., services that are relevant to the specific agency), several other questions also arise. For one, do administrators tend to be optimistic, pessimistic, or balanced in their predictions? That is, are their expectations more positive or more negative than actual public opinion? Some evidence suggests that administrators predict overly favorable citizen ratings of public services. For example, in a national survey of 304 city managers, Berman (1997) found that an overwhelming majority of respondents believed that citizens felt that local services were meeting their needs. But other research points in the opposite direction. Earlier research (Melkers and Thomas 1998, 331) found municipal administrators "usually expecting citizens to be less satisfied than they actually are with city services," perhaps because those administrators mostly hear the complaints that citizens report to municipal departments. These results appear consistent with the work of Poister and Henry (1994), who found that, contrary to conventional wisdom, a sample of Georgia residents did not view the quality of public services as inferior to that provided by private institutions.

The data can also speak to the question of which administrators perceive public opinion most accurately. The singular focus here on transportation precludes the kinds of interservice comparisons reported in he earlier research (Melkers and Thomas 1998), but it permits some other comparisons that were not possible there. We can, for example, examine whether central office or local district personnel know public opinion better. Here, too, the likely pattern is not obvious. On one hand, we might expect that regional or local administrators in large decentralized state agencies could better predict citizen feedback because they are "closer to the customer" in the actual delivery of services. On the other hand, top-level managers in central state offices might be expected, by virtue of their broader perspective on the agency's fit with its environment, to perceive customer perspectives more accurately.

We can also examine the purported "wisdom of crowds." Journalist James Surowiecki (2004), drawing on both scholarly research and field examples, has argued that "crowds," meaning small to large groups of people, frequently make better decisions than the most qualified experts. These groups or teams may outperform individuals in making decisions because they bring a wider variety of information, experience, and insight to bear on a problem. Relative to the current research, that pattern could mean that agency administrators collectively understand—and can predict—what citizens think better than any individual manager.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Despite their seeming importance on all of these counts, administrative perceptions of citizen perspectives have only rarely been the subject of systematic research. As best we can tell, no one else has sought to obtain administrative predictions of citizen perceptions in the manner recommended by Melkers and Thomas (1998, 333) since—or before—their work was published (for a limited exception, see Sabatier, Hunter, and McLaughlin 1987).

The current research will explore these issues using new data on state administrators' predictions of citizen perceptions of state transportation services. After explaining the two surveys from which these data are drawn, this paper will discuss the process of presenting customer survey results and administrative predictions to agency managers and, based on that experience, assess the value of using those predictions to increase interest in and the utility of citizen surveys. Next, the article will explore what else can be learned from the data on administrators' predictions. Beyond examining the extent to which managers' predictions are on target, the research will also compare the accuracy of predictions made by managers at different levels of an organization, as well as the accuracy of "group" predictions versus predictions made by individual managers.

The Data: The GDOT Public Opinion Poll and the Administrator Prediction Survey

The customer feedback used as a benchmark for this research was collected as part of an original public opinion poll conducted by the authors for the Georgia Department of Transportation in January 2004. The GDOT had conducted limited surveys of Georgia residents in previous years, most recently in the fall of 2002, but in every case, the earlier surveys consisted of only a small number of questions on a few aspects of GDOT services placed in larger omnibus surveys covering a variety of topics unrelated to transportation. The department received only frequencies and selected cross-tabulations of the answers to those questions, with no accompanying written report, presentation, or conversation about the results with department officials.

The 2004 survey was undertaken with different purposes and a different plan. First, rather than focusing on only a few aspects of the department's work, the survey was intended to cover the full range of GDOT services and performance and to focus exclusively on the department. Second, given the perception that previous surveys had yielded limited value, the development of the survey and the presentation of results were structured to be interactive, with the explicit intent of increasing the department's interest in and perceived value of the survey results.

Development of the Citizen Survey Instrument

With the second purpose in mind, we began developing the survey by interviewing the commissioner, deputy commissioner, chief engineer, and treasurer (the four statutory officers of the department), 10 other members of the executive team, and selected other managers regarding issues to be addressed and questions to be included in the survey. Using a semistructured, open-ended interviewing protocol, we asked each staff member what kind of information he or she would like to obtain from the survey. Team members were asked about a variety of specific question areas, as well as any additional areas that might be of interest.

Based on these interviews, we developed an outline of the areas the survey instrument would likely need to cover. To develop questions for those areas, we examined previous GDOT surveys to maintain longitudinal trend data wherever possible. We also reviewed survey instruments from other state departments of transportation for questions that might fit the GDOT's needs. Where neither source provided satisfactory questions, we formulated original questions. The draft items were then shared with many of the GDOT managers and revised in an iterative process until a consensus developed around an instrument focusing on citizen ratings of Georgia state highways, other transportation services, environmental impacts, communication with the public, and the possibility of an increase in liquid fuel taxes. The principle question formats used throughout the survey instrument included importance ratings on a scale from 1 (not at all important) to 10 (very important) and items asking respondent to grade aspects of the GDOT's performance on an A–F grading scale, in addition to standard five-point Likert items.

The eventual instrument, consisting of some 73 evaluative questions, as well as some items on respondent experiences and characteristics, was administered as a telephone survey in January and February 2004. Random-digit dialing and a survey protocol providing for a minimum of 10 callbacks yielded a total of 806 completed interviews with Georgia residents 18 years of age or older, for an overall response rate of 36 percent. The sample closely mirrored the state's population distribution in terms of gender, age, and the proportion residing in the 13-county Atlanta metropolitan area (as opposed to more rural parts of the state). In terms of racial distribution, however, blacks were underrepresented in the sample by about 8 percentage points, whites were overrepresented by approximately 6 percentage points, and other races were overrepresented by 1 percentage point. All respondents indicated that they were customers of GDOT by virtue of driving on state highways or using public transit systems or other modes of transportation supported by the department.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Obtaining Administrators' Predictions

The initial interviews of the executive team members about questionnaire composition provided one means to increase interest in the survey results. However, many citizen surveys have been developed using the same approach and still produced limited interest in the results. Achieving greater interest, we believed, requires more. For that reason, we decided, even before beginning the development of the citizen survey, to ask the GDOT officials to guess how the public would respond to the survey—parts of it, anyway.

To that end, as the survey went into the field, we turned to developing a "guessing" protocol. The citizen survey was obviously too long—approximately 24 minutes for the average phone interview—to ask GDOT officials to both respond to the survey themselves and to predict the distribution of public opinion on all or most of the survey items. We wanted to be respectful of the many demands on the time of these high-level public officials, but we also wanted to avoid setting up an overly tedious process that might be seen, at the outset at least, as purely academic research with little value to the department.

Because the GDOT officials were to be surveyed by mail, we concluded that a reasonable upper limit on what they could be asked might be the number of questions that could fit on two sides of a page. In this case, that turned out to be 14 questions—a relatively small number of survey questions, but certainly adequate for our predictive purposes. The best questions to include seemed to be those addressing the most general dimensions of the GDOT's work, especially those asking what the GDOT's priorities should be and how well the GDOT was performing. Some of these questions focused on specific functional areas (e.g., "How important is it to you for the Department of Transportation to preserve or improve the environment in planning and building transportation projects?"), whereas others pertained to the GDOT's work overall (e.g., "Thinking about all the services provided by the department, what overall grade on an A–F grading scale would you give the Georgia Department of Transportation?"). For the most part, we avoided including questions that had been asked on earlier surveys so that administrators could not check earlier results before making their predictions.

We debated how much detail to ask GDOT officials to predict on specific questions. In the earlier research (Melkers and Thomas 1998, 330), many questions asked administrators to allocate figures totaling 100 percent across several answer options. In some cases, officials responded with figures that did not total 100 percent; in other cases, we suspect, potential respondents simply chose not to complete the questionnaire rather than attempt to predict percentages that were supposed to total 100 percent on each of a number of questions. To simplify the task this time, we asked GDOT officials to predict only one number on each question, such as the "predicted percentage of A or B grades" for GDOT performance in a particular area. Or, where the public was asked to use a 1–10 point scale to rank the desired priority of an aspect of GDOT services (e.g., "preserving or improving the environment"), we asked GDOT officials to predict only the mean response. The eventual questionnaire asked GDOT officials to predict some aspect of the public's responses on 15 different survey questions. The 15 questions covered all of the major aspects of GDOT services but still fit on two sides of one page.

Our next task was to determine which officials to survey. After consulting with GDOT leaders, we concluded that, because all GDOT managers statewide might be able to use the citizen survey, different groups of managers should be included. We consequently surveyed the full populations of three groups: (1) the executive team, (2) other GDOT central office heads, and (3) district engineers and other office and section heads in the department's seven district offices. At the same time, because the principal survey interest lay with the executive team, only they received follow-up requests to complete the questionnaire, which likely explains why their response rate was higher than the other groups.

The surveys were distributed to all respondents through GDOT interoffice mail, with a request to return the completed forms to a single GDOT mail drop. Each form included a cover letter explaining that all responses would be kept confidential—only summary data would be reported. The survey forms were anonymous, identified only by a code number that was known only to us. We suspect that this anonymity, combined with the fact that the forms were quickly forwarded to us rather than remaining with the GDOT, quieted any concerns that the predictions would become public. In any event, the process resulted in good response rates for all of the target groups, as detailed in table 1.

With responses from 61 of the top managers in the department, we had the necessary data to explore in more detail (1) how these responses might be used to increase administrators' interest in the survey results and (2) what the predictions might tell us about the predictive abilities of the GDOT administrators, both individually and collectively.

Findings

The Value of Administrative Predictions in the Survey Reporting

As our first research question, we wanted to examine whether and to what extent these administrative

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Table 1 Distribution of Survey Respondents

	Population	Respondents	Response Rate
Executive team	14	12	86%
Central office heads	33	20	61%
District engineers and other district staff	50	29	58%
Overall	97	61	63%

predictions might be used to increase the interest in and value of the survey data. Hoping to maximize both, we followed this approach: After analyzing the data from both the citizen and administrator surveys, we prepared a Microsoft PowerPoint presentation for the GDOT executive team on the results of both surveys. In addition, for those questions that asked officials to predict the public's response, we employed the following protocol in reporting findings:

- We presented the data for one question at a time, beginning with the administrators' predictions, including both the mean prediction for all executive team respondents and the range of their predictions. Administrators' predictions were presented first in the hope of heightening interest in the actual public response. Both the range and the average were reported, as individual team members varied widely in their predictions.
- We next presented the actual numbers for the public on the question, at the same time comparing those results to the overall team predictions.
- We concluded by noting which team members had made the best guess on the question, hoping to contribute to a friendly competition among team members.

Did the technique produce the desired results, that is, greater interest in or value of the survey findings? We sought an answer using two approaches. First, we observed how the GDOT administrators responded to our presentation, looking for evidence of interest—or a lack thereof. Second, we asked our principal contact at the GDOT, the director of strategic development, how he felt about the value of the prediction technique, in particular, "Has the prediction process—asking the executive team to predict results in advance—been useful?"

His answer to that question squared with the earlier speculation about the value of asking for predictions. Having made predictions in advance, he argued, "gets folks engaged and orients the thinking of leadership. Their guesses force them to think and risk something." When they see the survey results, he continued, "they don't just react, and anything that forces them beyond just reacting is worthwhile." Moreover, retention of the findings may also increase, as the results "stick more in their minds because their guesses gave them a frame of reference."

Likewise, several reactions during the initial presentation of results to the GDOT executive team suggest that the prediction process increased interest in the results:

- Immediately before the meeting at which the survey results were presented, one executive team member joked about what the best guesses would bring: "The winner gets an early retirement with full benefits, right?" At least that participant was interested in comparing the team members' guesses to reality.
- At many points during the presentation, individual team members commented on how their individual or collective guesses compared to the actual public response, those comments often seeming to reflect a rethinking of their own perceptions. For example, on one question for which the team as a whole predicted a less positive response than the public gave, a team member commented, "We're always too hard on ourselves."
- After the meeting, one of the top GDOT executives raised the possibility of using the same process with the department's governing commission, a governor-appointed body, suggesting that the approach was perceived as helpful. (Such an effort was briefly pursued but did not prove feasible because of the commission's time constraints.)
- The GDOT subsequently asked that we employ the prediction process with surveys of other stakeholder groups (i.e., contractors, consultants), and we have complied.
- This format may enhance the ability to show how public opinion surveys represent only one reading on how the public feels and, at that, a reading that is likely better for tapping representative opinions than intensity of feeling. For example, when several executive team members voiced surprise that the public's assessment of the GDOT's environmental record was not worse, we suggested that their perceptions might be based on hearing more from an intense minority of constituents (e.g., environmental interests) who feel more negatively about the department's performance. Although those intensely felt negative assessments are important and not to be ignored, we thought the department should also recognize that the public as a whole might not be as negative.

On the whole, then, the technique appeared to work as intended, increasing administrators' interest in the survey results.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

It is much more difficult to assess whether the predictive technique led to greater utilization of the findings. The principal problem is that utilization itself is so difficult to measure. As Patton has documented, utilization only rarely involves new findings leading "directly and immediately to the making of major, concrete program decisions" (1986, 34). More typically, utilization occurs when new findings provide "additional pieces of information in the difficult puzzle of program action, permitting some reduction in the uncertainty within which any… decision maker inevitably operates." That is an exceedingly difficult phenomenon to capture.

Still, discussions with the department's director of strategic development did produce evidence of utilization of the survey results, purportedly greater than that of earlier survey reports. We asked him, "Have the survey results been utilized by the department? If so, how?" He responded affirmatively, citing the following:

- The GDOT commissioner used survey results in presentations to the governor to demonstrate that the department's priorities are consistent with the public's priorities, as revealed in the survey.
- The commissioner also used the results to argue for a higher state priority for road maintenance, the highest priority for the public in the public opinion survey.
- The commissioner used the same findings to demonstrate to GDOT employees how the public values their highway maintenance work.
- Two items from the public opinion poll were incorporated as performance targets on the department's strategic planning scorecard.

All of these uses appear to be unique to the current survey and not true of earlier surveys. We cannot say whether this utilization would have occurred had we not employed the predictive process, but we suspect that process heightened interest in a manner that contributed to this utilization.

The Accuracy of Administrators' Predictions

The second set of research questions concerned the accuracy of the GDOT administrators' predictions of citizen perceptions and opinions. Comparing the predictions made by the GDOT administrators with the feedback received from the sample of citizen-customers in the public opinion poll yielded some interesting findings.

Table 2 displays the results of the administrators' survey for the 14 items in question, including the mean response from the public opinion poll and the 99 percent confidence interval estimates for each of these parameters. For example, item 2 in the public opinion poll asked respondents to grade the quality of state highways in terms of condition and ride quality on an A–F scale. As table 2 indicates, 72.9 percent of the survey respondents gave the department an A or B grade in this regard. Given a sample size of 806 respondents (a simple random sample) from the population of all Georgia residents, we can be 99 percent certain that between 68.8 percent and 76.9 percent of the full population of adult residents of Georgia would assign the same grades.

We decided to consider a prediction by any group of GDOT top executives, office heads, or district managers, also shown in table 2, accurate or "on target" if the prediction fell within this 99 percent confidence interval. We chose the 99 percent confidence level rather than the 95 percent or 90 percent confidence level because we wanted to be as certain as possible that the resulting interval estimates would, in fact, contain the true value of the population parameters in question as an accurate basis for determining whether the administrators' predictions were on target. Given a sample size of 806 respondents, the precision of these intervals was quite tight—less than ±10 percent for all but three of the 10 test items—and thus provided a solid basis for discerning the extent to which the administrators' predictions were on target, below, or above the presumed true responses in the population of Georgia residents. Based on that standard, as summarized in table 3, the GDOT groups accurately predicted citizens' responses only 36 percent of the time (15 of 42 group predictions).

The picture changed dramatically, however, when the focus was narrowed to only the highest-priority GDOT service—highways. On the five questions about highways, the GDOT administrators were on target on 11 of 15 predictions, or 73 percent of the time. That accuracy cannot be explained by any repetition of questions from earlier surveys, as none of these five questions had been included in any earlier survey. Like the other half of the study, on questions about lower-ranked GDOT priorities (preserving the environment, providing a variety of transportation options, and communicating well)—lower in the eyes of both the department and the public—GDOT administrators were on target on only 1 of 18 predictions, or only 5 percent of the time.

On the survey items that were not predicted accurately, 20 of the 26 (77 percent) across the three groups were underpredicted, and only six (23 percent) were overpredicted. In other words, when these administrators collectively erred, they usually expected the public to be more negative than it actually was, a pattern that is consistent with earlier findings

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Table 2 Accuracy of GDOT Team Predictions of Public Opinion

Questions	Public Opinion Poll Response, Actual Survey (99% Confidence Interval)	Exec. Team Mean	Office Heads Mean	District Staff Mean	Exec. Team	Office Heads	District Staff
Condition of state highways (Grade A and B, percent)	72.9 *(68.84–76.94)*	74.4	73.8	75.9	On target	On target	On target
Congestion on state highways (Grade A and B, percent)	45.5 *(40.99–50.10)*	43.3	35.9	49.7	On target	Below	On target
Safety of state highways (Grade A and B, percent)	66.5 *(62.15–70.77)*	66.4	63.0	67.0	On target	On target	On target
GA highways are better than other states (percent)	54.3 *(49.29–59.28)*	75.7	68.5	71.0	Above	Above	Above
GA highways are worse than other states (percent)	8.9 *(6.02–11.72)*	8.2	9.1	7.9	On target	On target	On target
Importance of providing a variety of transportation options (mean)	8.4 *(8.25–8.60)*	7.2	6.7	6.1	Below	Below	Below
Performance in providing a variety of options (Grade A and B, percent)	54.2 *(49.61–58.81)*	34.6	36.2	37.4	Below	Below	Below
Importance of preserving environment (mean)	8.7 *(8.49–8.82)*	8.3	8.3	7.5	Below	Below	Below
Performance in preserving environment (Grade A and B, percent)	57.5 *(52.84–62.21)*	44.9	41.2	42.8	Below	Below	Below
Importance of communicating (mean)	8.8 *(8.62–8.91)*	8.0	8.9	8.3	Below	On target	Below
Contacted GDOT (percent)	11.2 *(8.31–14.05)*	15.9	19.7	17.2	Above	Above	Above
Contactors reporting GDOT helpful (percent)	69.5 *(57.39–81.61)*	63.5	63.8	65.5	On target	On target	On target
Performance in communicating (Grade A and B, percent)	62.4 *(57.93–66.85)*	53.6	48.6	55.4	Below	Below	Below
Overall services (Grade A and B, percent)	69.7 *(65.55- 73.93)*	54.8	48.2	56.5	Below	Below	Below

Note: The shaded boxes represent the best prediction for each survey item.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Table 3 Summary Accuracy of GDOT Team Predictions

	Exec. Team	Office Heads	District Staff	Total	Percent
All items					
Above	2	2	2	6	14
On target	6	5	5	16	38
Below	6	7	7	20	48
Totals	14	14	14	42	100
Highways items only[a]					
Above	1	1	1	3	20
On target	4	3	4	11	73
Below	0	1	0	1	7
Total	5	5	5	15	100
Items on lower priorities[b]					
Above	0	0	0	0	0
On target	0	1	0	1	6
Below	6	5	6	17	94
Total	6	6	6	18	100

[a]Includes ratings of five aspects of highways: condition, congestion, highways better or worse than in other states.

[b]Includes ratings of six areas other than highways, including priorities given to and performance on providing variety of transportation options, preserving the environment, and communicating well.

(Melkers and Thomas 1998, 330–31). However, that general pattern actually conceals two contrasting patterns. On questions about highways, the GDOT administrators leaned slightly toward optimism, more often expecting the public to be more positive than it was. On questions about the three other priority areas, by contrast, the GDOT predictions were almost uniformly pessimistic, with 17 of the 18 predictions (94 percent) below the actual public response.

Executives versus Office Heads and District Managers

The data in tables 1 and 2 also speak to comparative predictive abilities of the three groups of GDOT administrators: (1) members of the executive team, (2) office heads and senior-level managers in the department's central office, and (3) district engineers and other professional office staff in GDOT's seven districts around Georgia. Comparing each group's average predictions against those of the respondents to the public opinion poll can tell us about the relative accuracy of the three groups in predicting the public's grading of the department's performance. Though members of the executive team certainly had the most cumulative experience in the department and might well have been expected to have the greatest sensitivity to prevailing attitudes and regional variations in perspectives on transportation around the state, the district managers could be expected to have a better reading of the pulse of the public by virtue of operating at the point of service delivery and being "closer to the customer."

As it turned out, the executive team performed the best in predicting the public's response to the 14 survey items, making the most accurate predictions (the shaded cells in table 2) on six items, compared to five items for the office heads and four for the district managers. (The number of "winners" totaled 15 rather than 14 because the executives and office heads tied on one item.) As table 3 shows, the executive team also was on target in one more of its collective predictions than either the office heads or the district staff. Thus, no one group of GDOT administrators dominated the others in predicting the public's responses to these survey items, but the executive team proved slightly more perceptive than either of the other two groups of managers.

However, the differences between the three groups were less pronounced than their similarities. As a review of table 2 shows, the three groups agreed in the direction of their predictions—on target, above, or below—on 12 of the 14 items. Even on the three other items on which they did not fully agree, the differences were not large: In each case, two of the three groups offered similar predictions, and the third was only slightly different (i.e., in no case did one group predict below and another above on an item). These similarities are all the more striking because they appear among artificial groups, that is, groupings for which numerical averages were calculated, even though the groups had never actually discussed public opinion *as* groups.

> Administrative groupings at different levels of the GDOT—top executive team, central office heads, and district office professional staff—on average perceived public opinion very similarly, but they were mostly correct only on high-priority highway questions, whereas they were mostly incorrect and pessimistic on questions about other departmental priorities.

These findings present a curious contrast. Administrative groupings at different levels of the GDOT—top executive team, central office heads, and district

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

office professional staff—on average perceived public opinion very similarly, but they were mostly correct only on high-priority highway questions, whereas they were mostly incorrect and pessimistic on questions about other departmental priorities.

This pessimism may be traced to the difficult politics that state departments of transportation have faced recently around these other priorities, including environmental sensitivity, attention to alternative forms of transportation, and communication with stakeholders. As Swope has reported on the recent history of state transportation agencies, "Over and over again, DOTs found themselves fighting with local governments, historic preservationists and community activists" (2005, 28) about these very priorities. We suspect this fighting may have left GDOT staffers with an exaggerated sense of how negatively the public as a whole feels about the department's efforts in these areas. By contrast, because the conflict did not focus on highway quality—that is, condition, safety, and congestion—GDOT staff perceptions of public opinion about that quality remained relatively positive.

Individual versus Group Predictions

What, finally, of the "wisdom of crowds"—that is, the superiority of groups of GDOT administrators over individual administrators in predicting public opinion? Table 4 speaks to this question solely for the GDOT executives, reporting for each item the mean response on the public opinion poll, the executive team's mean prediction of that response, the range of individual predictions among the 12 executives, the number of individual predictions that were more accurate than the mean team prediction, and, conversely, the number of times that the team average was more accurate than *any* individual's prediction.

As shown at the bottom of table 4, of the 168 individual predictions (12 GDOT executives predicting on 14 items), only 45 (27 percent) were closer to the mean public response than the average prediction of the executive team. In other words, the executive team's average predictions were closer to the public's response than 123 of the 168 individual predictions. In addition, the average team prediction proved superior to the prediction of *any* of the individual team members on 5 of the 14 items. Notably, in light of the earlier discussion, four of those five items focused on the highest-priority area of highways. (The sixth question asked about the importance of preserving the environment in GDOT programming.) None of these questions had been asked on any previous survey. In short, the executive team "crowd" proved wiser than most of its individual members on all but one of the questions and wiser than *any* of its individual members on most of the highest-priority GDOT issues.

Table 4 GDOT Executive Team Predictions

Questions	Public Opinion Survey Response	Exec. Team Mean	Exec. Team Range	Best Exec. Team Prediction	Individual Predictions Better Than or Equal to Team Mean	Team Predictions Better Than Individual Predictions
Condition of state highways (Grade A and B, percent)	72.9	74.4	50–90	75	0	12
Congestion of state highways (Grade A and B, percent)	45.5	43.3	15–80	40 & 50	0	12
Safety of state highways (Grade A and B, percent)	66.5	66.4	50–90	65	0	12
GA highways better than other states (percent)	54.3	75.7	45–100	50	5	7
GA highways worse than other states (percent)	8.9	8.2	0–20	10	0	12
Importance of providing a variety of transportation options (mean)	8.4	7.2	4–9	8 & 9	5	7
Performance in providing variety of options (Grade A and B, percent)	54.2	34.6	2–70	55	5	7
Importance of preserving environment (mean)	8.7	8.3	6–10	8 & 9	0	12
Performance in preserving environment (Grade A and B, percent)	57.5	44.9	4–90	60	4	8
Importance of communicating well (mean)	8.8	8.0	3–10	9	8	4
Contacted GDOT (percent)	11.2	15.9	5–40	10	5	7
Contactors reporting GDOT helpful (percent)	69.5	63.5	5–100	70	5	7
Performance in communicating (Grade A and B, percent)	62.4	53.6	10–90	60 & 65	3	9
Overall services (Grade A and B, percent)	69.7	54.8	30–80	70	5	7
Total					45	123

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

Similar results were obtained for both office directors and district managers (not shown). For the 20 office directors, only 74 of the 280 (26.4 percent) individual predictions were closer to the public's responses than the average across the directors. Similarly, only 90 of 406 (22.2 percent) individual predictions by GDOT district managers were closer than the group average. In short, for all three groups of GDOT administrators, these imputed group predictions—the wisdom of the GDOT administrative crowd—proved far superior to the individual predictions of the public's responses to survey items.

It is important to underscore, however, that these are only imputed groups. Of all the groups, only the executive team actually met as a group, although it had not attempted to function as a group in predicting public opinion. Thus, the GDOT had not availed itself of any of the potential advantages of group predictions over individual predictions.

Conclusions

Working with municipal administrators and the city of Atlanta, Melkers and Thomas (1998) pioneered the technique of asking public administrators to predict citizen opinions in advance of receiving survey results. The research reported in this essay involved a more systematic use of the technique for professional staff of the Georgia Department of Transportation in conjunction with an early 2004 GDOT survey of Georgia residents about state transportation issues. This paper has explored whether this predictive technique might both increase the value of citizen surveys for clients and yield other valuable insights about public administrators themselves. The findings suggest affirmative answers on both counts.

On the first count, asking for advance predictions did appear to enhance the interest of GDOT administrators in the eventual survey results while, at the same time, making more evident to those administrators where they were right or wrong in their expectations of public priorities and perceptions. Both outcomes, we believe, likely increased the value of the survey results to the department by giving administrators a frame of reference for retaining the findings.

Analysis of the administrative predictions produced interesting findings on the accuracy of those predictions. To summarize,

- Groups of GDOT administrators were on target in their predictions (i.e., their predictions fell within a 99 percent confidence interval around the public's mean response) only slightly more than a third of the time, but strikingly, those on-target predictions were concentrated in the GDOT's highest-priority area of highways.
- When inaccurate, the predictions tilted toward pessimism, expecting public sentiment to be more negative than it actually was. Again, however, the pattern varied: GDOT administrators tilted slightly toward optimism on higher-priority highway questions but were almost uniformly pessimistic on questions about lower-priority issues. We suspect that recent political conflicts around these lesser priorities may explain this contrast.
- Although the top-level executives proved slightly more accurate in their predictions than the other groups of GDOT administrators, the real story lay in the striking similarities between the predictions of the different administrator groups—the three groups were of a similar mind on almost all of the questions. Those similarities are all the more striking given that all of the groups were artificial, their predictions reflecting only a numerical average rather than any collective group product.
- Collectively, the public administrators proved more perceptive—often substantially more perceptive—than they were individually. This "wisdom of crowds" was especially evident on questions about high-priority aspects of highways, for which the GDOT executives as a group were almost uniformly more accurate in their predictions than any of the individual members.

That wisdom, as Surowiecki (2004) argues, may depend on the members of the "crowd" being able to bring diversity and independence, as well as intelligence, to their efforts. The GDOT groups in the current research met all of these criteria: (1) intelligence in their knowledge of and experience with the GDOT, (2) diversity across the different offices and levels of the department, and (3) independence in that each administrator was asked to make predictions separate from the full group. In combination, those criteria may have led to more accurate group averages because, as Surowiecki writes, they "represent not what any one person in the group thinks but rather, in some sense, what they all think" (xix).

Potential for Broader Use

These results add to our conviction about the value of asking for administrators' predictions of citizen opinions. The use of this technique in conjunction with citizen and customer surveys may yield substantial gains at a relatively small cost.

As for the cost side, the principal costs in the GDOT case involved time, which was minimal: a few hours to construct the administrator questionnaire, negligible time to mail the surveys, a few hours for data entry and analysis, and an hour or two more to decide how to incorporate the comparisons in the presentation to the GDOT officials.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions

The technique could also have broader use than we achieved. Had circumstances permitted, we could have used the technique with the GDOT's governing board. Had the department desired, we also could have presented results and predictions to other GDOT work groups, in addition to the executive team. Such a feedback process could facilitate a "cascading" of survey results down through the organization, a common goal as many organizations (including the GDOT) attempt to make performance measurement a pervasive organizational concern.

Administrators might also benefit from hearing about the kinds of patterns in their predictions, as we have reported here. For example, the GDOT administrators might have gained confidence from hearing about the accuracy of their predictions of public sentiment and perhaps found encouragement from hearing about the pessimism in their guesses about public sentiment on other departmental priorities. They might also have learned from having the opportunity to discuss the meaning of both.

One of the few limitations we see in the use of this technique lies in the marginal value of repeating an administrative prediction process. An identical repetition of the exercise makes little sense because officials might remember—or check their records for—the results of the earlier exercise. Some potential for repetition remains, however, as administrators could be asked to guess (1) opinions on new questions that have been added to ongoing surveys (as in our case) or (2) the direction of opinion change on repeated questions, though in the latter case, perhaps only every several years, not annually.

Even with that limitation, we encourage more survey experts to consider the use of a technique that may yield increased practitioner interest in survey results, as well as other insights, both benefits coming at minimal cost. To that end, we are committed to that use in our own future research.

References

Berman, Evan M. 1997. Dealing with Cynical Citizens. *Public Administration Review* 57(2): 105–12.

Brown, Karin, and Philip Coulter. 1983. Subjective and Objective Measures of Police Service Delivery. *Public Administration Review* 43(1): 50–58.

Hatry, Harry P., John E. Marcotte, Therese van Houten, and Carol H. Weiss. 1998. *Customer Surveys for Agency Managers: What Managers Need to Know*. Washington, DC: Urban Institute Press.

Kelly, Janet M. 2005. The Dilemma of the Unsatisfied Customer in a Market Model of Public Administration. *Public Administration Review* 65(1): 76–84.

Kelly, Janet M., and David Swindell. 2002. A Multiple-Indicator Approach to Municipal Service Evaluation: Correlating Performance Measurement and Citizen Satisfaction across Jurisdictions. *Public Administration Review* 62(5): 610–20.

———. 2003. The Case for the Inexperienced User: Rethinking Filter Questions in Citizen Evaluation Surveys. *American Review of Public Administration* 33(1): 91–108.

Licari, Michael J., William McLean, and Tom W. Rice. 2005. The Condition of Community Streets and Parks: A Comparison of Resident and Nonresident Evaluations. *Public Administration Review* 65(3): 360–68.

Melkers, Julia, and John Clayton Thomas. 1998. What Do Administrators Think Citizens Think? Administrator Predictions as an Adjunct to Citizen Surveys. *Public Administration Review* 58(4): 327–34.

Miller, Thomas I., and Michelle M. Kobayashi. 2000. *Citizen Surveys: How to Do Them, How to Use Them, What They Mean*. Washington, DC: International City/County Management Association.

Parks, Roger B. 1984. Linking Objective and Subjective Measures of Performance. *Public Administration Review* 44(2): 118–27.

Patton, Michael Quinn. 1986. *Utilization-Focused Evaluation*. 2nd ed. Beverly Hills, CA: Sage Publications.

Poister, Theodore H., and Gary T. Henry. 1994. Citizen Ratings of Public and Private Service Quality: A Comparative Perspective. *Public Administration Review* 54(2): 155–60.

Sabatier, Paul, Susan Hunter, and Susan McLaughlin. 1987. The Devil Shift: Perceptions and Misperceptions of Opponents. *Western Political Quarterly* 40(3): 449–76.

Stipak, Brian. 1979. Citizen Satisfaction with Urban Services: Potential Misuse as a Performance Indicator. *Public Administration Review* 39(1): 46–52.

Surowiecki, James. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York: Doubleday.

Swindell, David, and Janet M. Kelly. 2000. Linking Citizen Satisfaction Data to Performance Measures: A Preliminary Evaluation. *Public Productivity and Management Review* 24(1): 30–52.

Swope, Christopher. 2005. Rethinking the Urban Speedway. *Governing*, October, 26–28, 30, 32, 34.

Van Ryzin, Gregg G. 2004. The Measurement of Overall Citizen Satisfaction. *Public Performance and Management Review* 27(3): 9–28.

Van Ryzin, Gregg G., Douglas Muzzio, Stephen Immerwahr, Lisa Gulick, and Eve Martinez. 2004. Drivers and Consequences of Citizen Satisfaction: An Application of the American Customer Satisfaction Index Model to New York City. *Public Administration Review* 64(3): 331–41.

This content downloaded from 103.27.8.44 on Tue, 24 Feb 2015 13:43:57 PM
All use subject to JSTOR Terms and Conditions