An Experimental Test of the Expectancy-Disconfirmation Theory of Citizen Satisfaction

Gregg G. Van Ryzin

Abstract

A number of prior studies have found evidence for the expectancy-disconfirmation theory of citizen satisfaction with public services, which holds that citizens judge public services not only on experienced service quality but also on an implicit comparison of service quality with prior expectations. But the evidence to date has been based on surveys (observational studies) and on subjective measures of expectations and performance, which are likely endogenous. Thus, the present study aimed to test the expectancy-disconfirmation theory of citizen satisfaction with public services using an experimental method. Participants in an Internet panel (N = 964) were randomly assigned to receive either low- or high-expectations statements from a hypothetical government official and to view either low- or high-performance street cleanliness photographs, in an online survey experiment. The findings are in line with previous research and generally confirm the core relationships in the theory, although the effect of expectations varied by age and political ideology. Because this study is a true randomized experiment, it provides better evidence than previous studies regarding the true causal nature of these relationships. © 2013 by the Association for Public Policy Analysis and Management.

INTRODUCTION

In public management worldwide, emphasis on the measurement of government performance is growing, with a particular push to measure outcomes (Kettl, 2005; Pollitt & Bouckaert, 2011; Van Dooren, Bouckaert, & Halligan, 2010). At the same time, it is widely recognized that the outcomes of many government services are elusive and difficult to measure objectively (Hatry, 2007; Holzer & Yang, 2004; Poister, 2003). For this reason, attention has focused on the use of citizen surveys as a way to try to gauge public performance in the form of satisfaction or quality ratings (Bouckaert & van de Walle, 2003; Folz, 1996; Hatry et al., 2006; Miller & Miller Kobayashi, 2000; Web & Hatry, 1973). Indeed, governments around the world now conduct citizen surveys and report the results as a measure of the quality of government services at the local, regional, national, and even international levels (Bouckaert, van de Walle, & Kampen, 2005; Miller, Miller Kobayashi, & Hayden, 2008; Stipak, 1980).

However, skepticism exists in both practitioner and academic circles about the validity of citizen ratings or satisfaction as a gauge of true government
performance. Some scholars suggest that the results of citizen surveys more closely reflect the characteristics and attitudes of the respondents than the actual quality of government services (Bouckaert & van de Walle, 2003; DeHoog, Lowery, & Lyons, 1990; Stipak, 1979a). Studies have found that respondent characteristics and political attitudes influence performance perceptions (Brown & Coulter, 1983; Hero & Durand, 1985; Lyons, Lowery, & DeHoog, 1992; Stipak, 1979b; Van Ryzin et al., 2004), and some empirical tests have found little correlation between subjective and objective measures of public services (Brown & Coulter, 1983; Kelly, 2003; Kelly & Swindell, 2002a). However, other studies have suggested that citizen perceptions may indeed reflect real government performance, particularly if the linkage between objective and subjective measures is better specified (Brudney & England 1982; Licari, McLean, & Rice, 2005; Parks, 1984; Percy, 1986; Van Ryzin, Immerwahr, & Altman, 2008).

Expectation-disconfirmation theory has been advocated as a way to explain the difference between perceived performance and more objectively measured performance, as well as to understand citizens’ judgments of government performance more generally. The theory holds that citizens judge public services based on an implicit comparison of service quality with their prior expectations, and there is growing evidence in support of the theory in public management research (James, 2009; Morgeson, 2012; Roch & Poister, 2006; Van Ryzin, 2004, 2006). But this evidence comes from surveys (observational studies) in which expectations and performance perceptions are likely to be endogenous with respect to satisfaction judgments. Thus, the aim of the present study is to test the expectancy-disconfirmation theory using a randomized experiment to provide firmer evidence of the causal nature of the key relationships in the theory.

The next section reviews expectancy-disconfirmation theory and prior empirical tests of the theory in the public management literature. This is followed by a description of the experimental design of the present study, in which online participants were randomly assigned to receive either low- or high-expectations statements from a hypothetical government official and to view either low- or high-performance street cleanliness photographs. Results are then presented in the form of path analysis (the method used in prior tests of the theory) and more traditional regression analysis of main and interaction effects. The paper concludes with a discussion of the results and their implications for expectancy-disconfirmation theory as well as for the practice of public management.

EXPECTANCY-DISCONFIRMATION THEORY

Originating in the study of consumer behavior, expectancy-disconfirmation theory holds that consumers develop different expectations of service quality from personal experience, word of mouth, advertising, or in other ways, and thus consumers apply different standards in forming their subjective ratings or satisfaction judgments (Cardozo, 1965; Erevelles & Leavitt, 1992; Oliver, 1997). Expectations thus constitute an adaptation level of sorts against which consumers contrast their experiences of service quality, in the same way that people’s perceptions of hot or cold are based in part on the temperatures to which their bodies have already adjusted (Helson, 1947; Oliver, 1980). And just as people stepping out of a dark theater experience the bright afternoon sunshine differently from passersby already accustomed to the daylight, varying expectations may well explain how consumers differentially perceive, and in turn judge, the quality of the same goods or services. The resulting difference or gap between expectations and performance is referred to as disconfirmation, which can be either positive (with performance exceeding expectations) or negative (with performance falling short of expectations). Expectancy
disconfirmation has been found to be a determinant of customer satisfaction in studies of a wide variety of private-sector products and services (Anderson & Sullivan, 1993; Bearden & Teel, 1983; Cardozo, 1965; Churchill & Suprenant, 1982; Oliver, 1980; Oliver & DeSarbo, 1988; Oliver & Swan, 1989; Tse & Wilton, 1988).

Expectancy-disconfirmation theory has relevance for research about public services as well as public management strategy and practice. As mentioned, the theory can shed light on the validity of subjective measures of government performance as well as explain how these subjective evaluations, derived from citizen surveys, differ from more objectively measured performance. More fundamentally, the theory can help illuminate the decisionmaking process ordinary citizens use to make evaluations of government. As such, it can help provide a theoretical foundation for a body of empirical work on the determinants of citizen satisfaction with various types of government services (Hero & Durand, 1985; Kelly & Swindell, 2002b; Lyons, Lowery, & DeHoog, 1992; Stipak, 1977; Stipak, 1979a; Van Ryzin et al., 2004). On a more practical or political level, expectancy-disconfirmation theory can potentially provide insights on how governments should direct or control their communications and public relations strategies (Liu, Horsley, & Yang, 2012). For example, when public managers and political leaders are faced with declining resources or service cutbacks, is it better to lower expectations or bolster them? The theory may also help make sense of media effects on public opinion about government performance (Bouckaert, van de Walle, & Kampen, 2005; James, 2011; Van de Walle & Bouckaert, 2007).

Figure 1 (adapted from Oliver, 1980 and Van Ryzin, 2004) shows the expectancy-disconfirmation model and its key hypothesized relationships. Link A represents the hypothesis that high expectations produce more negative disconfirmation, with all else equal, while link B represents the hypothesis that high performance leads to more positive disconfirmation, again with all else equal. Disconfirmation, in turn, is

![Disconfirmation Model Diagram](image)

positively related to satisfaction (link C). That is, positive disconfirmation (exceeding expectations) produces higher satisfaction, while negative disconfirmation (falling short of expectations) produces lower satisfaction. These three links in the model (A, B, and C) represent the basic expectancy-disconfirmation process (Oliver, 1980, 1997). Expectations and performance are often assumed to be positively correlated (link D), especially in observational studies, although the causal direction of this relationship is typically not specified.

Link E represents the direct effect of performance on satisfaction, and link F represents the direct effect of expectations on satisfaction. That good performance can directly and positively influence satisfaction (link E), separate and apart from its effect on disconfirmation, is intuitive and consistent with prior empirical studies (DeHoog, Lowery, & Lyons, 1990; Lyons, Lowery, & DeHoog, 1992; Van Ryzin et al., 2004). In other words, Link E implies that performance matters in some absolute sense, above and beyond its relative standing with respect to varying individual expectations. Link F, however, is less intuitive and refers to the direct effect of expectations on satisfaction, apart from disconfirmation. Some evidence suggests a negative direct effect (Poister & Thomas, 2011), consistent with the basic cognitive process of expectancy-disconfirmation as the mechanism of performance perception. But other studies find a positive direct effect of expectations (Oliver & DeSarbo, 1988; Van Ryzin, 2004), which may seem counterintuitive at first glance. But expectations could have a positive net effect on satisfaction to the extent that they create an anchoring or focusing effect (Kahneman, 2011). In other words, people may employ their prior expectations as a baseline or starting point when forming a satisfaction judgment, and thus their satisfaction ratings may tend to anchor on this baseline, especially when the evaluative task is ambiguous or difficult (Cardozo, 1965).

Expectancy-disconfirmation theory has been adopted and applied with increasing frequency in the public management literature. Using data from a citizen survey about New York City government services, Van Ryzin (2004) found that expectations were positively related to disconfirmation, and that perceived performance, an index of quality ratings of various urban services, was negatively correlated with disconfirmation, as the theory predicts. Disconfirmation, in turn, was a strong positive predictor of overall citizen satisfaction. Performance had a positive direct effect on satisfaction, beyond its indirect effect through disconfirmation, as did expectations. In a later replication using a U.S. online panel (Van Ryzin, 2006), similar results were obtained using subtractive disconfirmation, measured with a quality rating minus expectations, but not using perceived disconfirmation, measured with a single question about quality being above or below expectations. In particular, expectations did not have the predicted negative effect on perceived disconfirmation, although perceived disconfirmation still had a large positive effect on satisfaction.

Using data from a statewide survey in Georgia, Roch and Poister (2006) found that citizen satisfaction with trash collection, policing, and schools was significantly higher when perceived performance exceeded expectations, and significantly lower when it did not, even when controlling for the performance rating (measured as a letter grade) given to the service. The direct effect of performance itself (the letter grade) was positive in all analyses. These authors, however, were not able to measure the direct effect of expectations on satisfaction. Additional support for the role of expectancy-disconfirmation comes from a study by James (2009) using data from an Internet panel in the United Kingdom. James found that “the predicted probability of being satisfied rises as expectations of service quality are perceived to be increasingly met and exceeded” (James, 2009, p. 118). Interestingly, he observes an even larger negative effect on the probability of being dissatisfied. James also finds a positive direct effect of expectations on satisfaction, as well as a direct effect of performance on satisfaction, although these relationships were not estimated.
An Experimental Test of Expectancy-Disconfirmation

simultaneously. A study by Poister and Thomas (2011), using a statewide survey in Georgia on motorists’ satisfaction with highways, again finds that expectancy-disconfirmation (as measured primarily by perceived disconfirmation) has a large effect on satisfaction with road conditions, traffic flow, and safety. In all their analyses, expectations were negatively related to disconfirmation, and performance positively related, as the theory holds. These authors also found a strong positive direct effect of performance on satisfaction, as well as a weak negative direct effect of expectations on satisfaction. Finally, a recent study by Morgeson (2012), who tested the expectancy-disconfirmation theory in the context of satisfaction with the U.S. federal government using data from the American Customer Satisfaction Index, found results that were substantially similar to Van Ryzin’s (2006) study of local government satisfaction.

In sum, the core relationships forming the expectancy-disconfirmation model in Figure 1—links A, B, and especially C—has been supported by a number of previous investigations. These studies have also found fairly consistent evidence of a direct effect of performance on satisfaction (link E), although the magnitude of this direct effect varies a great deal across studies. There are, however, conflicting findings with respect to the expectations variable in particular, which has not always been shown to be related to disconfirmation (link A), as hypothesized, and has been found to have direct effects on satisfaction (link F) that are negative in some studies but positive in others. These conflicting findings for the expectations variable likely have much to do with the difficulty of measuring expectations, which often must be asked retrospectively, as well as differences in how the models are estimated in different studies. Another complication, however, is suggested by results of an interesting experimental study by James (2011) that showed important differences between how citizens form positive expectations (regarding what performance will be) versus normative expectations (regarding what performance should be). Specifically, his results showed that positive expectations can be influenced by exposure to credible government performance information, but that normative expectations appear more ingrained and less modifiable. Not enough attention has been paid to this distinction, with some studies operationalizing the construct as positive expectations and others as normative expectations. Thus, the role and relationship of expectations in the formation of satisfaction judgments remains unclear.

Tests of the expectancy-disconfirmation theory in the public management literature to date rely predominantly on surveys in which expectations, performance, and satisfaction are measured as part of the same cross-sectional, observational study. As a result, the presumed causal variables—expectations, performance, and, in turn, disconfirmation (whether subtractive or perceived)—are likely to be endogenous with respect to satisfaction. That is, causation could work the other way around (e.g., satisfaction could drive reported expectations) or there could be a lurking variable (like a general predisposition toward government or even toward responding to survey questions in a patterned way) that produces a spurious relationship between the variables. Thus, in discussing his somewhat conflicting findings, Van Ryzin (2006) suggests that, to get at the true causal structure of these complex relationships, “experimental research designs that independently manipulate expectations and performance are clearly necessary” (p. 609). With the exception of the work of James (2011), experiments along these lines have been absent from the public management literature, even though experimental methods have been used for some time in the field of consumer behavior to test various aspects of the theory (Cardozo, 1965; Oliver & DeSarbo, 1988; Tse & Wilton, 1988). In addition, it has been suggested that public management research should make more use of experiments to test hypotheses and develop theory (Brewer & Brewer, Jr., 2011; Margetts, 2011).

Thus, the main objective of this study was to experimentally test the effects of citizen expectations, perceived performance, and, in turn, disconfirmation, on
citizen satisfaction with government services. More specifically, the study aimed to estimate the direction and magnitude of the presumed causal relationships in the full expectancy-disconfirmation model, shown in Figure 1, using experimentally manipulated expectations and performance. Such experimental evidence can help confirm as well as clarify findings from prior observational studies that, although more realistic in some ways, are much more limited in terms of clearly demonstrating causation. It should be pointed out, however, that even randomized experiments remain limited in terms of their ability to fully identify mediation effects, a methodological issue that has been raised recently in the social sciences but that does not have an easy solution (Bullock, Green, & Ha, 2010; Imai et al., 2011). Thus, in addition to estimating the kind of mediation (path) model used in previous studies and shown in Figure 1, the data for this experiment was also analyzed using a more traditional regression framework.

EXPERIMENTAL DESIGN AND PARTICIPANTS

The study was designed as a $2 \times 2$ factorial experiment in which participants were randomized to one of four arms, as shown in Figure 2. Each arm of the experiment was a different version of a Web-based questionnaire, with the initial email list fully randomized into four segments linked to each questionnaire. The participants ($N = 964$) were respondents to an emailed study invitation sent to individuals in the CivicPanel project, a university-affiliated Internet research panel (Welcome to CivicPanel, 2011). CivicPanel recruits on an ongoing basis using regular Web directory listings, social media postings, and Craigslist (http://www.craigslist.org) ads, and includes panelists of various ages and income levels from across the United States and other countries (primarily Canada and Europe). Data were collected online from September 13 to 24, 2010, in response to an emailed invitation to 12,211 panelists, 1,489 of whom opened the email message (representing a contact rate of 12 percent, a cooperation rate of 65 percent, and an overall panel response rate of 8 percent).

The participants were predominantly white (81 percent) and a majority were female (64 percent); the median age was 44. Twenty-one percent of the participants had an annual household income of less than $25,000, while 31 percent had over $75,000 in annual household income. In terms of education level, 24 percent had only a high school education or less, while 28 percent had a four-year college degree.
or more. Eighty-three percent of participants were from the United States, with the remainder from various other countries. Of the U.S. participants, 33 percent were from the Northeast, 29 percent from the South, 19 percent from the Midwest, and 19 percent from the West. Thirty-four percent of all participants lived in big cities, 27 percent in suburbs, 28 percent in small cities or towns, and 13 percent in rural areas. In terms of their political views, 33 percent identified themselves as liberals, 41 percent as moderates, and 26 percent as conservatives. In sum, although the participants in the experiment were a voluntary, nonprobability sample that cannot be statistically projected to the U.S. population, they were nevertheless a geographically diverse group of adults of various ages, incomes, education levels, and political views. As would be expected from random assignment, statistical tests (ANOVA and Pearson chi square) revealed no significant differences on any of these variables across the four arms (or treatment groups) in the experiment.

The online experiment started with a section that asked a short series of closed-ended questions, including the quality of the participant’s local community as a place to live, their perceived change in the community over the last few years, and the participant’s interest in politics and national affairs, trust of the national government, trust of the local government, and political ideology. These questions were designed to prime participants to think about their own local community and its quality of life, as well as to get baseline data on political interest, trust, and ideology. Participants were then given these instructions:

“In this next section, you will be asked to read descriptions, look at photographs, and answer questions about a hypothetical city named 'Hometown.' Please consider this information carefully and answer the questions as if you were a resident of Hometown.”

Participants were then exposed to the first randomized factor, which involved either a high- or low-expectations statement from a hypothetical public official. In the low-expectations statement, the official mentions the current recession and a decline in tax revenues (the survey was conducted, it should be noted, in the midst of the Great Recession) and then cautions citizens to expect a decline in the quality of public services. In the high-expectations statement, the official also mentions the recession and the decline in tax revenues but then promises citizens that the government will maintain high-quality public services. The complete statements as presented to participants are shown in Table 1. A manipulation check after presentation of the statement was done by asking participants this follow-up question: “Based on what you have read about Hometown, how would you rate your expectations for the city government's performance? (1 = very low expectations to 7 = very high expectations).” As Figure 3 shows, the mean for participants randomly assigned to the high-expectations statement was significantly higher than the mean for those randomly assigned to the low-expectations statement—evidence that the manipulation did alter participants’ expectations.

Next, participants were exposed to the second randomized factor: either a high- or low-performance picture of street cleanliness, as shown in Figure 4. The images were taken from the photographic standards used in New York City's street cleanliness scorecard (City of New York, 2010), a well-developed and established performance measurement system. It should be noted that the two photos used in the experiment are not far apart on the seven-category scale used by New York City to rate the level of street litter: the high-performance photograph is the second cleanest, just clean enough to be rated “acceptable,” and the low-performance photograph is the fourth of a total of seven levels (and thus the middle of the range). This was done so that the variation in performance would lie within a fairly normal range and not appear too extreme (e.g., a perfectly spotless street compared to a filthy street covered in litter), which might potentially swamp any expectations effect. After presenting the photograph, participants were asked the following question: “How would you rate
Table 1. Expectations manipulation.

**Introductory language (all participants)**

Hometown is a medium-size U.S. city with a growing population. Recently, Hometown’s city administrator made the following public statement about the current economic situation and the city’s budget difficulties:

**Low-expectations version (n = 508)**

(AUGUST 2010): My fellow citizens, Hometown is facing challenges as a result of the ongoing recession and a continuing decline in local tax revenues. The city has been forced all this year to make cuts across the board in order to balance our budget.

So I need to be honest with you: These cuts will result in noticeable declines in the quality of many public services. The streets may not be as clean as you expect, potholes may not be fixed as fast as they should be, and many other city services will no longer meet our usual high standards. I regret that we cannot maintain the quality of services that you expect of city government, but we all must make sacrifices during these difficult times and hope, together, for better days ahead.

**High-expectations version (n = 456)**

(AUGUST 2010): My fellow citizens, Hometown is facing challenges as a result of the ongoing recession and a continuing decline in local tax revenues. The city has been forced all this year to make cuts across the board in order to balance our budget.

But I want to assure you of one thing: The necessary cuts we have made will not reduce the quality of public services. Indeed, the city will find a way to make sure that the streets stay as clean as you expect them to be, that potholes continue to be fixed as fast as they should be, and that all city services will be maintained at our usual high standards. I promise you that we will continue to deliver the high-quality public services that you expect of city government, and that we will find a way, together, to do more with less during these difficult times.

Based on what you have read about Hometown, how would you rate your expectations for the city government’s performance? (1 = very low expectations to 7 = very high expectations).

Note: Test of difference in means: $t = 8.15$, $df = 913$, $p < .001$.

**Figure 3.** Mean Expectations Ratings by Group.
An Experimental Test of Expectancy-Disconfirmation

Figure 4. Performance Manipulation.

The picture below shows the cleanliness of a typical city street in Hometown. Please look at the picture, then answer the questions below.

<table>
<thead>
<tr>
<th>Low-performance photograph (n = 477)</th>
<th>High-performance photograph (n = 487)</th>
</tr>
</thead>
</table>

How would rate the cleanliness of city streets in Hometown? (1 = poor to 7 = excellent). Note: Test of difference in means: t = 12.70, df = 906, p < .001.

Figure 5. Mean Performance Ratings by Group.

the cleanliness of city streets in Hometown? (1 = poor to 7 = excellent).” As Figure 5 shows, the performance manipulation also had its expected effect: participants randomly assigned to the low-performance photograph rated the streets as much less clean than did those randomly assigned to the high-performance photograph.

Finally, participants were asked this question about their satisfaction with government: “Based on the cleanliness of Hometown, how satisfied would you be with the city government’s performance? (1 = very dissatisfied to 7 = very satisfied).”
This question serves as the dependent variable in the test of the full expectancy-disconfirmation model.

**MODELING RESULTS**

In order to estimate the full expectancy-disconfirmation model (Figure 1), a measure of disconfirmation was constructed by subtracting the self-reported expectations rating (on a 1 to 7 scale) from the performance (cleanliness) rating (also on a 1 to 7 scale). This new measure thus ranges from \(-6\) to \(+6\) and has a mean of \(-0.72\), a standard deviation of 2.09; it is operationally similar to the subtractive measures of disconfirmation used in several prior studies (James, 2009; Van Ryzin, 2004, 2006). With this measure of disconfirmation, all of the variables are in place to estimate the full model in Figure 1. These variables include:

- Expectations (1 = high-expectations statement, 0 = low-expectations statement).
- Performance (1 = high-performance photo, 0 = low-performance photo).
- Disconfirmation (\(-6\) to \(+6\) subtractive measure, see paragraph above for operational definition).
- Satisfaction (on a 1 to 7 scale).

Regression-based path analysis, which is equivalent to structural equation modeling when all variables are observed rather than latent (as they are in this case), was used to estimate the full model. Although not the usual approach to the analysis of experiments, path analysis remains the form in which the expectancy-disconfirmation theory has been presented and tested in nearly all prior studies, so it is used here initially for purposes of comparison. Figure 6 presents the results, including both unstandardized as well as standardized path coefficients, along with significance tests. It should be noted that no correlation was estimated between expectations...
An Experimental Test of Expectancy-Disconfirmation

and performance (link D in Figure 1) because these variables were randomly assigned treatments and thus independent of each other by design (the correlation between these variables was $r = -0.0005$).

Overall, the empirical model in Figure 6 explains 26 percent of variance in satisfaction. More importantly, all of the core relationships were in the expected direction and were both substantively strong and highly significant statistically ($p < .001$ for all path coefficients). High expectations produced nearly a 1 point decline ($-0.9$) in disconfirmation, and high performance produced an even larger 1.4 point increase in disconfirmation (consistent with the differences in means shown in Figures 3 and 5). In turn, each point increase in disconfirmation led to over a third of a point (0.35) increase in satisfaction. In standardized terms, a 1 standard deviation increase in disconfirmation produced a nearly one half (0.44) standard deviation increase in satisfaction.

The direct effects of performance on satisfaction were moderately large, over half a point increase (0.58) in satisfaction from being exposed to the high-performance photograph and holding constant disconfirmation. Putting this direct effect together with the positive indirect effect of performance on satisfaction through disconfirmation ($1.445 \times 0.354 = 0.512$), the result was a 1-point (1.095) total effect of performance on satisfaction. In standardized units, the total effect of performance on satisfaction was nearly a third (0.32) of a standard deviation (i.e., the standardized total effect $= 0.171 + (0.345 \times 0.435)$).

The direct effect of expectations on satisfaction was positive and moderately large, representing nearly a third of a point increase (0.31) in satisfaction from being exposed to the high-expectations statement and again holding disconfirmation constant. But this positive direct effect was completely offset by the negative indirect effect of expectations on satisfaction through disconfirmation (which was $-0.903 \times 0.354 = -0.320$). As a result, the total effect of expectations on satisfaction was practically zero. Of course, it follows that the standardized total effect of expectations on satisfaction was also practically zero.

As mentioned, this kind of path analysis has been the main approach used in the public management literature to test the expectancy-disconfirmation theory when using survey data. In this study, however, expectations and performance were experimentally manipulated rather than simply observed. Still, the role of disconfirmation in the theory remains somewhat ambiguous because of the limitations, even in a randomized experiment, of examining mediation (Bullock, Green, & Ha, 2010; Imai et al., 2011). Specifically, disconfirmation as a mediator may still be influenced by pre- or posttreatment omitted variables that also influence satisfaction judgments. Thus, it is important to look at these experimental results in a more traditional regression (ANOVA) framework, ignoring the mediating role of disconfirmation in the path model, in order to estimate the main effects of the two experimentally manipulated factors (expectations and performance) along with the possibility of an interaction effect. To make the regression results directly comparable to ANOVA, both the expectations factor and the performance factor were effect coded (i.e., +1 for high expectations/performance and −1 for low expectations/performance), which resulted in coefficients that are directly interpretable as experimental main effects (Montgomery, 2008). The regression results for the experiment as a whole (all participants) appear in Table 2 (top row), and Figure 7 illustrates these overall results graphically. Similar to the path analysis, these results clearly show that performance had a large and statistically significant main effect on satisfaction. The expectations factor had only a very small negative main effect, which was not at all significant statistically. The interaction effect was positive, meaning that high expectations seemed to amplify the effect of performance on satisfaction judgments. Although not quite significant statistically, this somewhat amplified performance

Journal of Policy Analysis and Management  DOI: 10.1002/pam
Published on behalf of the Association for Public Policy Analysis and Management
Table 2. Regression analysis, including subgroup analysis.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Expectations</th>
<th>Performance</th>
<th>Interaction</th>
<th>Constant</th>
<th>R-sqr</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants</td>
<td>908</td>
<td>0.55***</td>
<td>0.08</td>
<td>3.49***</td>
<td>0.10</td>
</tr>
<tr>
<td>Females</td>
<td>556</td>
<td>0.55***</td>
<td>0.08</td>
<td>3.42***</td>
<td>0.10</td>
</tr>
<tr>
<td>Males</td>
<td>317</td>
<td>0.13</td>
<td>0.03</td>
<td>3.60***</td>
<td>0.12</td>
</tr>
<tr>
<td>Younger (less than 45)</td>
<td>449</td>
<td>0.04</td>
<td>-0.02</td>
<td>3.67***</td>
<td>0.09</td>
</tr>
<tr>
<td>Older (45 or more)</td>
<td>459</td>
<td>-0.03</td>
<td>0.18**</td>
<td>3.32***</td>
<td>0.13</td>
</tr>
<tr>
<td>Liberal</td>
<td>304</td>
<td>0.14*</td>
<td>0.13</td>
<td>3.41***</td>
<td>0.12</td>
</tr>
<tr>
<td>Moderate</td>
<td>367</td>
<td>0.14*</td>
<td>0.13</td>
<td>3.41***</td>
<td>0.12</td>
</tr>
<tr>
<td>Conservative</td>
<td>237</td>
<td>-0.22***</td>
<td>-0.01</td>
<td>3.30***</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Notes: Unstandardized coefficients shown; *p < .10, **p < .05, ***p < .01 (two-tailed tests).

Based on the cleanliness of Hometown, how satisfied would you be with the city government’s performance? (1 = very dissatisfied to 7 = very satisfied).

Note: The performance main effect is significant (p < .01) but the expectations main effect and the interaction effect are not (see top row of Table 2 for the statistics).

**Figure 7.** Graph of Satisfaction Ratings by Experimental Arm.

effect when expectations are high can be seen in the graph in Figure 7 and is a suggestive finding that will be discussed in the next section.

Table 2 also shows the same regression analysis of main and interaction effects for select subgroups of participants. Because participants were randomized, the experimental design holds within each of these subgroups considered separately (although the sample size is of course reduced). For all subgroups, the main effect of performance was consistently strong and highly significant statistically. Interestingly, the main effect of expectations was negative and clearly significant statistically for politically conservative participants, meaning they expressed less satisfaction with government performance when initially prompted to have high expectations.
In contrast, the main effect of expectations was slightly positive for political moderates, although this finding is only of borderline significance statistically. Another interesting finding from the subgroup analysis was the significant and fairly large positive interaction effort for older participants. This positive interaction indicates, as mentioned earlier, that expectations intensify or amplify the effect of performance on the satisfaction judgments of older people in the experiment.

DISCUSSION AND IMPLICATIONS

The results of this online experiment involving street cleanliness generally provide support for the expectancy-disconfirmation theory of citizen satisfaction with government performance. More importantly, because this study used an experimental design, these results supply some initial evidence in the field of public management research that at least some of the key relationships in the theory are truly causal. Disconfirmation was shown to be negatively determined by expectations and positively determined by performance, as hypothesized. Disconfirmation, in turn, had the expected positive effect on satisfaction, although this relationship is less clearly causal because disconfirmation is a mediator and thus not directly manipulated experimentally. Performance was found to have a positive direct effect on satisfaction as well. All of these relationships are consistent with expectancy-disconfirmation theory as well as previous studies.

But interestingly, the results also suggest little or no net total effect of expectations on satisfaction, at least in the aggregate, with the hypothesized negative indirect effect on satisfaction largely offset by a positive direct effect. This lack of a net effect of expectations is confirmed also by the more traditional regression analysis of experimental main and interaction effects for the full sample. These findings have potential practical implications, as prior studies have suggested that government officials could perhaps manage citizens' expectations, through public statements or social marketing, in an effort to maximize satisfaction (James, 2009; Van Ryzin, 2004). The statements from the city administrator varied in this experiment indeed simulate two distinct communication strategies for managing public expectations in a context of economic recession, revenue shortfalls, and service cutbacks. But the overall results suggest that neither strategy would work especially well—because manipulating expectations did not have a net causal impact on satisfaction in the experiment as a whole.

The subgroup regression results, however, did show that political conservatives express less contentment overall with government performance when they approach the situation with high expectations. It is unclear why this ideological subgroup in particular seems susceptible to this expectations effect, but perhaps it is because political conservatives tend to have a generally more negative or critical view of government to begin with (Jost, Federico, & Napier, 2009). Certainly the role of political ideology and its relation to expectations is an interesting question to consider in future research. Another interesting subgroup finding was that high expectations seem to amplify the salience of performance in the minds of older citizens, but this positive interaction effect was also evident (although not quite significant statistically) in the full sample and some other subgroups (females as well as political liberals and moderates). Although not a strong finding, this pattern at least hints at the possibility that high expectations among citizens or clients may increase both the potential rewards (if performance is good) as well as the risks (if performance is poor) for public service-delivery organizations. Certainly, additional empirical exploration of this expectations-performance interaction effect, involving different public services and in varying contexts (and with diverse populations), should be considered, as it may help explain previously conflicting findings.
An important implication of the results of this experiment is that performance alone—the photograph of street cleanliness—clearly appears to be the main causal driver of satisfaction in both the path analysis and the regressions, including all of the subgroup regressions. This finding is especially noteworthy given the relatively subtle difference in street cleanliness between the low- and high-performance photographs that were shown to participants (Figure 4). At least with respect to this one commonplace public service, people appear quite capable of perceiving fairly fine distinctions in performance and making their evaluative judgments of government accordingly. This finding has potential implications for the use of citizen surveys as a measure of public performance, which has long been criticized for being inescapably biased and fundamentally invalid (see especially Stipak, 1979a, 1979b). These experimental results correspond closely with a New York City study by Van Ryzin, Immerwahr, and Altman (2008), using a large citizen survey, which found that perceptions of street cleanliness were highly correlated with the city government’s street cleanliness scorecard (the source of the photographs in Figure 4). Thus, satisfaction ratings may reflect public performance more accurately than widely assumed, at least with this kind of public service.

Limitations to this experiment must be taken into account when drawing conclusions from these results. To begin with, street cleanliness is a clearly visible service that citizens experience on a daily basis and that they can presumably judge quite confidently. As a result, there may be less potential for street cleanliness judgments (based on photographs) to be influenced by prior expectations and other sources of bias. It would be useful to replicate this kind of experiment in an area where quality or performance judgments are based on more complex or abstract performance information. For example, people may be more ambiguous or uncertain about judging disease rates, air quality measures, crime rates, or emergency response times, as well even more abstract performance metrics such as standardized test scores for public school students, or broad quality of life indexes. Indeed, it would be interesting in future experiments to examine whether expectations play a stronger role in the formation of satisfaction judgments as the performance information becomes more complex or difficult for people to interpret.

Another limitation of the study is that the expectations manipulation, although confirmed by a manipulation check, may not reflect how expectations operate in the real world. The hypothetical statement by a public official does have the advantage of being similar to the kind of communication strategy that has been mentioned as a practical application of expectancy-disconfirmation theory in the public sector (James, 2009; Van Ryzin, 2004). But it can also be argued that such statements provoke a rather shallow or short-term form of expectations that are perhaps fairly readily disregarded when forming quality or satisfaction judgments. Deeper or more established expectations, formed by years of experience with government, could perhaps have more of an effect. One way to test this might be to experiment with priming techniques to elicit more ingrained positive or negative images of government. As suggested by James (2011), another important distinction that may matter in this context is the difference between positive and normative expectations. This study manipulated essentially positive expectations only, and it may be that manipulation of normative expectations would result in different findings. However, as suggested by James (2011), it may be more difficult to influence and thus manipulate normative expectations. Finally, as one reviewer of this study pointed out, the expectations statement in this experiment also included information about the economic recession and declining tax revenues, perhaps suggesting a scapegoat for worse-than-expected performance (in the minds of participants) and thus diluting the theorized expectations effect. A less ambiguously worded official statement might have prompted more firmly held or low high expectations, and
this issue should be considered in the design of future experiments along these lines.

In conclusion, this study provides experimental evidence that the expectancy-disconfirmation model may indeed describe a causal process through which citizens form their judgments about the quality of public services and in turn the performance of government. The results suggest that performance is the key determinant of citizen satisfaction and that expectations, although influencing disconfirmation, may have less of a net main effect on how citizens form their satisfaction judgments than previously thought. However, the effect of expectations seems to vary somewhat by age and ideological subgroups and, importantly, may act to amplify the impact of government performance in the minds of citizens. Additional studies that look at other definitions of expectations, especially normative expectations, as well as other areas of public service delivery, which are numerous, need to be conducted before such conclusions can be generalized. Finally, it is worth noting that randomized experiments would seem to offer important methodological advantages for understanding the true causal nature of the kind of relationships hypothesized in the expectancy-disconfirmation model. At the very least, the field of public management research should put more emphasis on the use of experiments to test the causal structure of models that have previously been developed using only surveys or other observational (nonexperimental) studies.

GREGG G. VAN RYZIN is Associate Professor, School of Public Affairs and Administration, Rutgers University—Newark, 111 Washington Street, Newark, NJ 07102.

REFERENCES


Stipak, B. (1979a). Are there sensible ways to analyze and use subjective indicators of urban service quality? Social Indicators Research, 6, 421–438.


An Experimental Test of Expectancy-Disconfirmation

