



Københavns Universitet



The Appropriate Response to Performance Information by Citizens: A Large-scale Experiment with Local Politicians

Olsen, Asmus Leth

Publication date:
2015

Document Version
Tidlig version også kaldet pre-print

Citation for published version (APA):
Olsen, A. L. (2015). The Appropriate Response to Performance Information by Citizens: A Large-scale Experiment with Local Politicians.

The Appropriate Response to Performance Information by Citizens: A Large-scale Experiment with Local Politicians

Asmus Leth Olsen*
Department of Political Science
University of Copenhagen
ajlo@ifs.ku.dk
www.asmusolsen.com

First draft. December 2014.

Abstract

How do politicians view the external response from citizens to performance information? We answer this question by testing how politicians expect citizens to respond to relative performance, social and historical performance comparisons, good and bad performance, and the assignment of responsibility for performance. This constitutes the first study of how politicians view the external role of performance information. We draw on two unique survey vignette experiments conducted with a large sample of Danish local councilors (n=814). The results show that councilors expect citizens to care a lot about both relative social and historical performance. Interestingly, the results also show very divergent asymmetries in councilors expectations to how citizens respond to relatively good and relatively bad performance. Specifically, councilors expect a negativity bias if blame is assigned to others and a positivity bias if praise is assigned to the council.

KEYWORDS: performance information · blame avoidance · citizen satisfaction

*This research was made possible with grant 0602-02577B from the Danish Research Council given to Asmus Leth Olsen. A previous version of this paper was presented at the workshop: *Citizen Attitudes and Behaviors from an Experimental Perspective: Towards an Experimental Public Administration?* Speyer, Germany, September 09-10, 2014. The author would like to thank Oliver James, Sebastian Jilke, Martin Baekgaard, Lene Holm Pedersen, Gregg van Ryzin, and Kurt Houlberg for comments on the experiments or earlier drafts of the manuscript.

Introduction

Recently there has been an increasing interest in how managers and politicians perceive and use performance information (Askim 2007; Askim et al. 2008; Askim 2009; Moynihan and Pandey 2010; Kroll 2013; 2014a; 2014b). Still, many have recognized that our empirical knowledge about how politicians perceive performance information is limited (Poister and Streib, 1999; Moynihan, 2005b; Askim 2007, Nielsen and Baekgaard 2013). Specifically, most studies focusing on how politicians or managers view performance information focus on the *internal* organizational role of the performance information (Askim 2009). However, performance information also serves the purpose of helping citizens evaluate the value that government creates for them (Osborne and Plastrik 2000, 247). Yet, as James (2011b) has argued, public management research has neglected how politicians and managers view the *external* accountability role of performance information. This stands in stark contrast to the many recent efforts to experimentally understand the external role of performance information from the perspective of citizens. This research has highlighted how relative evaluations, references points, and a negativity bias shape how citizens use and perceive performance information (Charbonneau & Van Ryzin, 2013; James 2011a; 2011b; Olsen, 2013a; 2013b; 2013c; Hansen et al., 2014). This raises the question if politicians *normative* expectations to the external role of performance information differs from or matches what we observe among citizens: How do politicians expect citizens to respond to performance information about an organization? Do politicians expect citizens to engage in relative performance evaluations and be more responsive to negative rather than positive information as citizen-level research has shown? And to what extent are politicians' expectations dependent on who responsibility for good and bad performance assigned to?

This article turns the focus to how politicians view the external effect of performance information among citizens by responding to the above posed questions. The argument here

is that we can learn something about politicians use of performance information by studying their expectations to how citizens respond. The underlying assumption is that politicians perceptions and actions are motivated by their beliefs about the publics response to performance information. Specifically, four aspects are derived from the existing experimental literature in order to develop a set of questions guiding the experimental study of politicians expectations to citizens response to performance information. The four aspects are: How politicians view (1) the importance of relative performance evaluations compared with absolute evaluations (James 2011b; 2011b), (2) the relative strength of social and historical comparisons in relative performance evaluations (Charbonneau and Van Ryzin 2013; Olsen 2013d; Hansen et al. 2014), (3) the presence of a negativity bias in citizens evaluations (John and James 2007; Boyne et al. 2009; Olsen 2013c), and finally (4) how the assignment of responsibility affects politicians' view of citizens' use of performance information (Moynihan 2005a; 2005b; 2008; Kroll 2014a).

In order to test these four aspects of politicians' view of citizens' response to performance information, we conduct two vignette experiments in a large sample of Danish local councilors (n=814). Danish local councilors constitute a highly relevant population for the study of performance information as they are politically responsible for the majority of expenses in the large Danish welfare state. Accordingly, they are regularly exposed to all sorts of performance information from a diverse set of public sector services.

The broader contribution of the article is twofold: First, while there is a growing use of experimental research to study citizens, only few studies of performance information has been conducted on a larger scale among politicians (Nielsen and Baekgaard 2013). In the experiments, we asked councilors how satisfied citizens should be with various services. We experimentally manipulated the organizational performance and the relative performance to various reference points – including past performance and the performance of others (Simon 1939). We hereby offer the first experimental study of how politicians expect citizens to

respond to various forms of performance.

Second, we offer a comparative perspective on the response to performance information by contrasting experimental evidence found among citizens with that of politicians. By contrasting the two we are able to identify differences and similarities in citizens response to performance information and how politicians expect them to respond. This provides a closer integration of the research into how various actors respond to performance information. Understanding similarities and differences may help understand when and why performance information has intended or unintended consequences (Kelman and Friedman, 2009).

How do Politicians Expect Citizens to Respond to Performance Information?

The existing research on citizens response to performance information poses a number of questions to how politicians expect citizens to react. Performance information is here defined broadly as any regular, systematic, and quantitative data accessible to the public which capture the inputs, actions, outputs, or outcomes of public organizations or political entities (James 2011b). In the experimental study of citizens response to performance information focus has mainly been on three aspects: 1) The importance of relative performance evaluation compared with absolute evaluations (James 2011a; 2011b), 2) the relative strength of social and historical comparisons in relative performance evaluations (Charbonneau and Van Ryzin 2013; Olsen 2013d; Hansen et al. 2014), 3) the presence of a negativity bias in performance evaluations (John and James 2007; Boyne et al. 2009; Olsen 2013c). In addition, from the perspective of politicians, there has been focus on a fourth aspect namely how the assignment of responsibility affect politicians view of citizens use of performance information (Weaver 1986; Moynihan 2005a; 2005b). Each of the four aspects of the existing body of research leads to four distinct empirical questions about politicians' expectation to citizens' response to performance information. These will be outlined in further detail below.

It has long been recognized that reference points can serve as goals or aspirational levels for performance for organizations or individuals. In the early work of Herbert Simon it was noted how, from the perspective of the administrator, "the only sound basis for decisions about numbers is numerical factual information about past experiences or the experiences of others nothing more nor less than comparative statistics." (Simon 1939, 106). Reference dependence implies that performance is coded in terms of loss or gain, success or failure, negative or positive, or good or bad depending on how the performance of the organizations in question is doing relative to some reference point (Festinger 1954; Simon 1955; March and Simon 1958; Tversky and Kahneman 1974; Tversky and Kahneman 1981; Heath 1999). The reference point is therefore crucial for forming expectations and aspirations to match with performance feedback. Hereby, comparing performance to some reference point provides a measure of attainment discrepancy and places performance in a domain of either falling short or making the cut. The existing studies at the citizen level point to that citizens care a lot about relative performance when evaluating performance information (James 2011a; 2011b; Olsen 2013d; Hansen et al. 2014).

This leads us to the first expectation how politicians' think citizens should respond to performance information: *To what extent do politicians expect citizens to rely on reference points for performance evaluations and thereby engage in a relative mode of evaluation?*

A natural extension of the reference point discussion is to focus on variation in the impact of various reference points. Specifically, it has been an integrated part of organizational research to analyze how organizations engage in comparisons with other organizations, i.e. *social comparisons*, and with their own past performance, i.e. *historical comparisons* (March and Simon 1958; Cyert and March 1963; Heath 1999). In social psychology, there has long been a similar distinction between (historical) comparisons in time and (social) comparisons with others (Albert 1977; Festinger 1954). This distinction has also gotten some attention in recent experimental research on the effect of performance information on citizens. A striking

finding has been that social reference points tend to affect attitudes more than historical ones (Charbonneau and Van Ryzin, 2013; Olsen, 2013d, Hansen et al., 2014). That is, citizens are more affected by information about how other organizations are doing relative to the one they are evaluating than simply judging how the organization in question is performing over time. For instance, Olsen (2013d) that social reference points (e.g. other organizations) have a stronger impact on citizen evaluations than historical comparisons (e.g. past years performance). Hansen et al. (2014) also finds that voters prospective preferences are also highly affected by cross-national concerns: How are we doing relative to other neighboring countries?

This leads to a natural second question to politicians' expectations to citizens' responses to performance information: *To what extent do politicians expect citizens to rely more strongly on social reference points than historical ones?*

A related matter in the experimental study of citizens' reaction to performance information has been on the negativity bias. The negativity bias is defined by the fact that "negative events are more salient, potent, dominant in combinations, and generally efficacious than positive events" (Rozin and Royzman 2001, 297). For the case of performance information it implies that positive and negative performance of the same magnitude have asymmetrical effects on citizens with the negative information showing a stronger impact. The experimental evidence of this effect has been mixed. For instance, James and Moseley (2014) find a stronger negative effect of absolute and relative performance on citizens' satisfaction with public services. On the other hand, James (2011b) finds no clear support for a negativity bias in a laboratory experiment where participants were exposed to "good" and "bad" performance information. Observational studies find more clear support for a negativity bias in citizens response to performance information (Soroka 2006; Boyne et al. 2009).

This leads to our third question to politicians' expectations to citizens' responses to performance information: *To what extent do politicians expect citizens to respond with a nega-*

tivity bias to performance information, that is, by giving more weight to relatively bad performance compared with relatively good performance?

Finally, the negativity bias has often been linked with blame avoidance behavior among politicians and managers (Weaver 1986; Boyne et al. 2009; Johnsen 2012). Pollitt (2006a: 41) has argued that politicians "do not take much interest—unless and until disasters, scandals, or breakdowns come along...". Not only citizens but also the media coverage of performance information are overly focused on performance problems (James 2004; Soroka 2006; Dixon et al. 2013). This faces politicians and managers with some equally asymmetrical incentives to avoid bad performance rather than seek good performance. Some have noted that a motivation underlying politicians use of performance information can be *opportunistic*, that is, they use the available information strategically (Moynihan 2005; Askim 2007). The implication is that politicians may have different wishes to citizens' responses depending on how the performance information relates to a service for which the politicians themselves are to be either blamed or praised.

This leads to our fourth and final question to politicians' expectations to citizens' responses to performance information: *To what extent do politicians' expectations to citizens' response to performance information depend on the degree to which the performance information assigns accountability to politicians?*

An Experimental Research Design

In order to answer the four questions a large-scale experiment was conducted among real politicians. Specifically, a survey experiment was administered to a large sample of Danish local councilors. Denmark has 98 municipalities with around 2,450 local councilors filling the 20 to 30 council seats in each municipality. The mayor is the only full time paid councilor while most other councilors use between 15 to 25 hours a week on council related work and ob-

tain a payment of between 20.000 to 30.000 USDs a year. Danish municipalities are seen as the main provider of public services to citizens and account for more the half of all public spending. Danish local councilors are therefore important politicians to analyze as they are routinely confronted with performance information on schools, child care, administrative service satisfaction, and much more.

During the spring of 2013 all Danish councilors were emailed a survey containing questions about their views on local matters and their role as councilors. In addition to the initial invitation, three reminder emails were sent out. This resulted in a large sample of which some drop out during the survey which results in a slightly smaller sample for the results presented here (n=814)

Towards the end of the survey councilors were randomly assigned to a number of treatments in two separate vignette experiments. In both vignettes councilors were asked to evaluate organizations given a combination of absolute and relative performance information. The experiments draw on recent experimental work done among citizens (Charbonneau and Van Ryzin 2013; James 2011b; Olsen 2013d). This allows us to more precisely compare politicians' beliefs about citizens' response to performance information and the existing experimental work on citizens' responses.

The first experimental vignette dealt with municipal unemployment rates and the other presented councilors with grade averages from a hypothetical local school. Both experiments contained three treatments each. The two service areas are chosen because performance information for both are frequently reported to citizens. It is therefore likely that councilors have expectations to how citizens respond to both areas of performance. Furthermore, the two areas allow for variation in the attribution of responsibility for the observed performance. For the case of unemployment, attribution of responsibility will often be assigned at the municipal level as unemployment statistics cover each municipality as a single unit. However, for the case of school performance, the unit in question is often individual schools within a

municipality.

In the case of unemployment, all three experimental conditions asked councilors to evaluate a hypothetical scenario outlining the unemployment situation in their own municipality. Specifically, councilors were asked to rate how satisfied they think that citizens *should* be with various local unemployment rates. This was done on a 11-point scale ranging from *0=not at all satisfied* to *10=very satisfied*.

The three treatments varied the reference point for performance comparison. The exact texts of each treatment are outlined in table 1. In short, the conditions contained: 1) an absolute value of local unemployment, 2) an absolute value of local unemployment and a social comparison value, and 3) an absolute value of local unemployment and a historical comparison value. The three treatments allow us to test the first and second question concerning the importance of relative performance and whether social or historical comparisons have a stronger impact on evaluations.

In the first treatment councilors are only provided with a hypothetical unemployment rate of their municipality. In the second treatment the very same information is provided along with a hypothetical average rate for all other municipalities in the country which is intended to serve as a social reference point. In the third treatment councilors are provided their own hypothetical unemployment rate along with another hypothetical rate representing last year's unemployment.

The experiment also contained a second stage of randomized treatment in which the numerical value of the unemployment rates are randomized (Olsen 2013d). Randomizing the numerical performance information has a number of advantages. First, it overcomes the potential problems by choosing to experimentally expose councilors with just one or two arbitrary values. Performance numbers can themselves contain unwarranted information and symbolic meaning which can be difficult to determine a priori (Wildavsky 1964; Olsen 2014b). Second, by randomizing the numbers we can also test the importance of distance

between target and reference point. Some councilors have been assigned to cases where the difference was very narrow while others have experienced large differences. Third, it allows for comparing potential different effects of moving away from the reference point in both worse and better directions. This allows us to measure a potential negativity bias as stated in our third question to politicians' expectations.

The randomized numerical values are drawn from a predefined normal distribution with an average of 6.3% and a standard deviation of 1.0%. These parameters correspond roughly to national unemployment rate during the survey period. This distribution is used both for the specific organization in question and the reference point and both values are drawn independently of each other.¹ The two levels of randomization implies that we have both a discrete level of treatment (absolute, no reference point / historical reference point / social reference point) and a continuous level of treatment

[Table 1 about here.]

The second experimental vignette deals with school performance. The basic setup is similar to that of the unemployment experiment: a between-subjects-design with three treatment groups containing 1) only a single grade for one school, 2) a school grade average along with the combined average of schools in the same municipality, and 3) a school grade average along with the previous year's grade average for that school. The exact texts of each treatment are outlined in table 2.

As in the employment experiment the school performance scenarios also contained a second stage of randomization of the actual grades presented to councilors. Again the grades were drawn from a normal distribution with a mean of 6.5 and a standard deviation of 1.0.²

¹An example of the distribution of randomized numerical values in the experiments can be viewed in the appendix.

²The Danish grade scale is a 7-point scale with values from worst to best being: -3, 0, 2, 4, 7, 10, and 12 which corresponds to American grades of scale as: F, Fx, E, D, C, B, and A.

[Table 2 about here.]

Apartment for the substantial differences between school and unemployment performance, the two experimental vignettes are also systematically different in how responsibility for performance is assigned. In the case of schools the councilors are asked to make judgment about the satisfaction of "*parents with children at this school*" given the provided information. That is, responsibility in case of relatively good or bad performance is directed towards the individual school. In the case of unemployment, councilors instead make judgment about the satisfaction of "*citizens in your municipality*" on the basis of the randomized performance information. That is, responsibility is directed towards the council. This is important to investigate the fourth question we posted to councilors expectations which aimed at variation in assignment of responsibility to politicians.

Empirical Findings

In figure 1 the mean evaluations across the treatment groups in both experiments are shown with 95% confidence intervals. First, we can note that the councilors think that citizens should be much less satisfied in the unemployment case ($M=4.3$, $SD=2.2$) than in the school case ($M=5.2$, $SD=2.2$). The difference in mean responses is highly significant ($d = 0.9$, $p < 0.01$).

Across both experiments we can note how the expected satisfaction is higher under the historical reference point. We also see greater variations in average evaluations across the unemployment experiment than in the school experiment. For school performance only the social and historical reference point evaluations differ ($d = 0.4$, $p < 0.05$). For the case of unemployment, the historical reference point is significantly higher than both the social reference point ($d = 0.4$, $p < 0.05$) and the no reference point condition ($d = 1.1$, $p < 0.001$). Here the social reference point also provides significantly higher evaluations than the no reference point condition ($d = 0.7$, $p < 0.01$).

[Figure 1 about here.]

Next we turn to answering the four questions we posed. First, we look at how important absolute and relative performance was for councilors' expectation to citizens' evaluations. The main results are reported in table 3. The first set of models use the randomized absolute value for each school/municipality and the randomized reference point (either social or historical). The second line of models looks at reference dependence and symmetries which we will return to later.

First, we look at the school experiment. As expected the model with no reference point shows that councilors expect citizens to be more satisfied with better performing schools. The positive coefficient in model I (no comparison) indicate that a one grade point shift changes the expected citizen satisfaction by about one unit on the 11-point scale. This amounts to about half a standard deviation change for a one standard deviation change in the assigned grade average. This indicates that councilors attended to the information in the experiment and responded in accordance with the intended manipulation of performance information.

Next we turn to model I with a social reference point. Here both randomized numbers in the form of the absolute grade and the reference point are introduced in one model. The school grade average itself has about the same effect while the reference point has about half the effect. The negative sign on the reference point coefficient indicate that councilors expects parents to evaluate the school 0.5 points worse as the reference point increase by one grade point average. This is a sizable effect which denote a strong relative component in evaluations. A similar but numerically slightly stronger effect is found for the historical reference point. In summary, relative performance has a profound effect on councilors expectations to citizens evaluations of schools performance.

For the unemployment experiment the absolute unemployment rate has a negative effect on evaluations and reduce them by 0.7 for each percentage point increase in the unemploy-

ment rate. This effect is smaller than the absolute effect in the school performance vignette. We can also see that the model fit is substantially better in the school experiments than in the unemployment experiments. This tells us that councilors generally expected performance information to affect parents' satisfaction with schools more than performance information should affect citizens' satisfaction with the municipality.

The effects of the social and historical reference points are about the same as in the school case namely around half a point. The positive sign indicates that councilors expect citizens satisfaction to increase as the unemployment rate rises for the reference point. Again we can note that councilors expect relative performance to matter a lot to citizens.

Findings from both the school and unemployment case are therefore both supportive and conflicting with experimental evidence among citizens. It supports the finding in previous studies that relative performance matters a lot to citizens. Local councilors expect citizens rely on reference points. However, we do not see evidence of a greater potency of social reference points if compared with historical ones as has been the case for citizen focused experiments.

[Table 3 about here.]

Next we turn to the interaction models in order to identify a negativity bias in the response to relatively good and relatively bad performance. In table 3 these results are provided in the second model for both experimental vignettes under the social and historical treatments. In both treatments across both cases have a variable indicating relative performance been coded. It measures how relatively good the organization in questions is doing by subtracting the organizations performance from the reference point performance. In the unemployment case the subtraction is done in reverse in order to also let higher values reflect relatively better performance. In addition, a dummy is coded if performance was relatively bad, that is, if performance was below the reference point (and zero if performance was equal to or better

than the reference point). By interacting the two we obtain an estimate of how sensitive councilors expectations to citizens' satisfactions response is to relatively bad vs. relatively good performance.

In figure 2 the asymmetries between relatively good and bad performance are highlighted for the school experiment and reflect the results of table 3. We can clearly see the difference in steepness between relatively good and relatively bad performance. For negative grades (i.e. when school performance is below the reference point) councilors reacts with a stronger negative expectation to citizen satisfaction than if the school performs above the reference point. Looking at the estimated effects in table 3 we can see that the interaction is only significant for the historical comparison case. The effect is sizable: the effect of changes in performance on councilors expectation to citizens satisfaction is around twice as large for relatively negative performance compared with relatively positive performance. This finding is in line with studies on loss aversion and a negativity bias in responses to performance information among citizens (James 2011a; 2011b; Soroka 2006; Olsen 2013c).

[Figure 2 about here.]

In figure 3 the asymmetries between relatively good and bad performance are highlighted for the municipal unemployment experiment. Here the exact opposite effect than in the school experiment is visible: Councilors respond with a stronger positive expectation to citizens' evaluations when unemployment is above the reference point than if the municipal unemployment is worse than the reference point (i.e., negative values). This corresponds to a positivity bias in the sense that councilors expect performance above the reference point to be evaluated better than the negative response to performance below the reference point of a similar magnitude. If we look at the models in table 3 we see negative signs on the coefficients for the interactions which corresponds to this effect. This asymmetry is again only significant for the historical comparison and again the difference in coefficients for relatively good and bad

performance is of a sizable magnitude.

[Figure 3 about here.]

In sum, we find evidence of an striking difference in councilors asymmetrical expectations to citizens response. For the school case where attribution of responsibility is directed towards the school, councilors expect citizens to show a negativity bias: relatively bad performance is given more weigh than relatively good performance of the same magnitude. This expectation become reversed in the unemployment case, where attribution of responsibility to a greater extent is directed at the council.

Conclusion

The article has offered the first experimental study of how politicians view the external role of performance information, i.e., how they expect citizens to respond when faced with various forms of performance information. The argument has been that we can learn something about how politicians respond to performance information by studying how they expect citizens to respond to performance information. Reelection seeking politicians will have an edge if they skillfully are able hold valid beliefs and wishes about how citizens respond to performance information.

In the article we have approached the question of how politicians think citizens should respond to relatively good and bad performance with an experimental design. This was done by embedding a set of experiments in a survey among a large sample of Danish local councilors. The results provide the first possibility of comparing how politicians expect citizens to respond to performance information with the existing experimental knowledge of how citizens actually respond. In sum, the findings pointed to than when asked about how citizens *should* respond to various types of performance information we obtain both valid beliefs about citizens responses and elements of opportunistic or biased beliefs.

First, we found that politicians expect citizens to care a lot about relative performance in terms of either other organizations performance or own past performance. Politicians beliefs about the importance of relative performance reflects experimental studies on how citizens actually respond to performance information (Charbonneau and Van Ryzin, 2013; James 2011a; 2011b; Olsen, 2013a; 2013b; 2013d; Hansen et al. 2014). However, they did not expect citizens to be more affected by social reference points than historical reference points as it has been found among citizens.

Second, we found compelling opposite asymmetries in how councilors think citizens should respond to relatively bad and good performance for the two experimental cases. In the case of schools, the councilors thought that citizens should respond more negatively to relatively bad performance than positively to relatively good performance. This effect was consistent with a negativity bias. However, for the case of unemployment we found evidence of a positivity bias.

This divergence may be explained by how responsibility attribution was assigned for the two experiments. For the case of schools, councilors were asked about parents' satisfaction at the school in question. Here the question of accountability was assigned at the school level. However, for the case of unemployment, the councilors were asked about the satisfaction of all citizens in the municipality. Here the question of accountability was assigned more directly to the level of the council itself. This finding thus may reflect that councilors hope citizens will appreciate success on policies where responsibility is assigned to the council (unemployment), while they care less about this for policies where responsibility is assigned to others (schools). In other words: councilors have high hopes for blame avoidance when they are on the line, but expect harsh reactions to relatively negative performance when others have the main responsibility.

It is important to note that this difference cannot be explained by the fact that councilors see unemployment performance as generally more difficult to handle or less mutable than

school performance. If this was the case, we should expect councilors to think that citizens should be less affected in case of both relatively bad and relatively good performance—not just in the latter case. Furthermore, we would have expected that the average expectation of citizens satisfaction would have been higher across all conditions for the unemployment case than the school case. The data showed the opposite.

This article has only offered a first look at similarities and differences between how politicians and citizens view the external role of performance information. The results point to a number of implications for future studies: First, the fact that politicians recognize the importance of relative comparison to historical and social reference points calls for further systematic investigation of how such comparisons are sought out in a real world setting. Most existing experimental studies have provided respondents with a fixed reference point, however, in a more naturalistic setting multiple reference points will be available. Here is a opportunity for taking the experimental results into an observational setting.

Second, with politicians' recognition of the importance of reference points we should turn more focus to how organizations aim to affect how, when, and if they are compared with historical and social reference points. From a research perspective we can approach the use of reference points from the perspective of measurement gaming and manipulation. Reference points may be easier to manipulate or strategically change than actual performance. Our efforts to understand unintended behavior induced by performance information should therefore also be directed towards the reference points which are used and the reference points which are not.

References

- Askim, J., Johnsen, Aa., & Christophersen, K. A. 2008. Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory* 18(2), 297–320.
- Askim, J. 2007. How do politicians use performance information? An analysis of the Norwegian local government experience. *International Review of Administrative Sciences* 73(3), 453–472.
- Askim, J. 2009. The demand side of performance measurement: explaining councillors' utilization of performance information in policymaking. *International Public Management Journal* 12(1), 24–47.
- Boyne, George A., Oliver James, Peter John, and Nicolai Petrovsky. 2009. Democracy and Government Performance: Holding Incumbents Accountable in English Local Governments. *Journal of Politics* 71(4): 1273–1284.
- Charbonneau, É., & Van Ryzin, G. G. 2013. Benchmarks and Citizen Judgments of Local Government Performance: Findings from a survey experiment. *Public Management Review*.
- Festinger, Leon. 1954. A theory of social comparison processes. *Human Relations* 7(2): 117–140.
- Hansen, K. M., A. L. Olsen, & M. Bech. 2013. Cross-national Yardstick Comparisons: A Choice Experiment on a Forgotten Voter Heuristic. *Political Behavior*.
- Heath, C. 1999. Goals as Reference Points. *Cognitive Psychology* 38(1): 79–109.
- Heinrich, C. J. 2002. Outcomes-based performance management in the public sector: implications for government accountability and effectiveness. *Public Administration Review* 62(6), 712–725.
- James, Oliver. 2011a. Managing Citizens' Expectations of Public Service Performance: Evidence from Observation and Experimentation in Local Government. *Public Administration* 89(4): 1419–1435.
- James, Oliver. 2011b. Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments. *Journal of Public Administration Research and Theory* 21(3): 399–418.
- James, Oliver, & Peter John. 2007. Public Management at the Ballot Box: Performance Information and Electoral Support for Incumbent English Local Governments. *Journal of Public Administration Research and Theory* 17(4): 567–580.
- Kelman, Steven, & John Friedman. 2009. Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service. *Journal of Public Administration Research and Theory* 19(4): 917–46.

- March, J. G., & H. A. Simon. 1958. *Organizations*. New York: Wiley
- March, James G. 1988. Introduction: A Chronicle of speculations about organizational decision-making. Oxford, UK.
- Moynihan, D. P. 2005a. Why and How do State Governments Adopt and Implement 'Managing for Results' Reforms? *Journal of Public Administration Research and Theory* 15(2): 219–243.
- Moynihan, D. P. 2005b. Goal-based Learning and the Future of Performance Management. *Public Administration Review* 65(2): 203–216.
- Moynihan, D. P. 2008. *The Dynamics of Performance Management: Constructing Information and Reform*. Washington, D.C: Georgetown University Press
- Moynihan, D. P., & Pandey, S. K. 2010. The big question for performance management: why do managers use performance information? *Journal of public administration research and theory*.
- Kroll, A. 2013. Explaining the Use of Performance Information by Public Managers: A Planned-Behavior Approach. *American Review of Public Administration*
- Kroll, A. 2014a. Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. Forthcoming in *Public Performance & Management Review*.
- Kroll, A. 2014b. Why Performance Information Use Varies Among Public Managers: Testing Manager-Related Explanations. *International Public Management Journal* 17(2), 174–201.
- Lau, Richard R. 1982. Negativity in Political Perception. *Political Behavior* 4(4): 353–377.
- Nielsen, P. A., & Baekgaard, M. 2013. Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment. *Journal of Public Administration Research and Theory*.
- Olsen, A. L. 2013a. The politics of digits: evidence of odd taxation. *Public Choice*, 154 (1-2), 59–73.
- Olsen, A. L. 2013b. Leftmost-digit-bias in an enumerated public sector? An experiment on citizens judgment of performance information. *Judgment and Decision Making*, 8(3): 365–371.
- Olsen, A. L. 2013d. Compared to What? Experimental Evidence on Social and Historical Reference Points in Performance Evaluation. Paper presented at the 11th Public Management Research Conference Madison, Wisconsin, June 20–23, 2013.
- Olsen, A. L. 2015a. Citizen (Dis)satisfaction: An Experimental Equivalence Framing Study. *Public Administration Review*.
- Olsen, A. L. 2015b. "Simon Said," We Didn't Jump. *Public Administration Review*.

- Olsen, A. L. 2015c. The Numerical Psychology of Performance Information – Implications for Citizens, Managers, and Policy Makers. *Public Performance & Management Review*
- Osborne, D. and P. Plastrik. 2000. *The reinventor's Fieldbook: Tools for transforming your government*. Jossey-Bass San Francisco, CA.
- Pollitt, C. (2006a). Performance Management in Practice: A Comparative Study of Executive Agencies. *Journal of Public Administration Research and Theory* 16(1): 25–44.
- Pollitt, C. (2006b). Performance Information for Democracy—The Missing Link? Evaluation: *The International Journal of Theory, Research and Practice*, 1(2): 38–55.
- Poister, Theodore H. and Streib, Gregory (1999) Performance Measurement in Municipal Government: Assessing the State of the Practice. *Public Administration Review* 59(4): 325–35.
- Simon, H. A., & C. E. Ridley (1938). Trends in Municipal Reporting. *The Public Opinion Quarterly*, 2(3), 465-468.
- Simon, H.A. (1938). Inter-city contets. *Municipal Year Book*, 210-216.
- Simon, H.A. (1941). Measurement techniques in administrative research. *Civic Affairs*, 8(9), 1-2, 5-8.
- Soroka, Stuart N. 2006. Good News and Bad News: Asymmetric Responses to Economic Information. *Journal of Politics* 68(2): 372–385.
- Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (4481): 453–458.
- Tversky, A., and D. Kahneman. 1991. Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics* 106(4): 1039-1061.
- Van Ryzin, G. G., Muzzio, D., Immerwahr, S., Gulick, L., & Martinez, E. 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), 331-341.
- Weaver, R. Kent. 1986. The Politics of Blame Avoidance. *Journal of Public Policy* 6(4): 371–398.
- Wildavsky, Aaron B. 1964. *Politics of the budgetary process*. US: Little, Brown, and Company.

Table 1: Experimental design: Unemployment rates and reference points

Frame		Question wording	Numerical treatment
<i>Absolute</i> (n=278)	<i>level only</i>	Imagine that your municipality has an unemployment rate of $x\%$. How satisfied do you think that the citizens in your municipality should be with this unemployment rate?	$x \in N(\mu = 6.3, \sigma = 1.0)$
<i>Social</i> (n=275)	<i>reference</i>	Imagine that your municipality has an unemployment rate of $x\%$. <i>Imagine as well that the unemployment average for all other municipalities in the country is $x_2\%$.</i> How satisfied do you think that the citizens in your municipality should be with this unemployment rate?	$x_1 \in N(\mu = 6.3, \sigma = 1.0)$ $x_2 \in N(\mu = 6.3, \sigma = 1.0)$ $x_1 \perp x_2$
<i>Historical</i> (n=252)	<i>reference</i>	Imagine that your municipality has an unemployment rate of $x\%$. <i>Imagine as well that the unemployment rate in the same municipality last year was $x_2\%$.</i> How satisfied do you think that the citizens in your municipality should be with this unemployment rate?	$x_1 \in N(\mu = 6.3, \sigma = 1.0)$ $x_2 \in N(\mu = 6.3, \sigma = 1.0)$ $x_1 \perp x_2$

Table 2: Experimental design: School performance and reference points

Frame		Question wording	Numerical treatment
<i>Absolute</i> (n=258)	<i>level only</i>	Every year the Ministry of Education releases a grade average for the country's schools. Imagine that a regular public school in your municipality obtains an average grade of X on the new grade scale for students taking their final exam in 2013. How satisfied do you think that parents with children at this school should be given this information?	$x \in N(\mu = 6.5, \sigma = 1.0)$
<i>Social</i> (n=270)	<i>reference</i>	Every year the Ministry of Education releases a grade average for the country's schools. Imagine that a regular public school in your municipality obtains an average grade of X on the new grade scale for students taking their final exam in 2013. <i>Imagine as well that the grade average for all other public schools in your municipality is X.</i> How satisfied do you think that parents with children at this school should be given this information?	$x_1 \in N(\mu = 6.5, \sigma = 1.0)$ $x_2 \in N(\mu = 6.5, \sigma = 1.0)$ $x_1 \perp x_2$
<i>Historical</i> (n=286)	<i>reference</i>	Every year the Ministry of Education releases a grade average for the country's schools. Imagine that a regular public school in your municipality obtains an average grade of X on the new grade scale for students taking their final exam in 2013. <i>Imagine as well that the grade average for the same school last year was X.</i> How satisfied do you think that parents with children at this school should be given this information?	$x_1 \in N(\mu = 6.5, \sigma = 1.0)$ $x_2 \in N(\mu = 6.5, \sigma = 1.0)$ $x_1 \perp x_2$

Table 3: Experiments: Social vs. Historical Performance Information

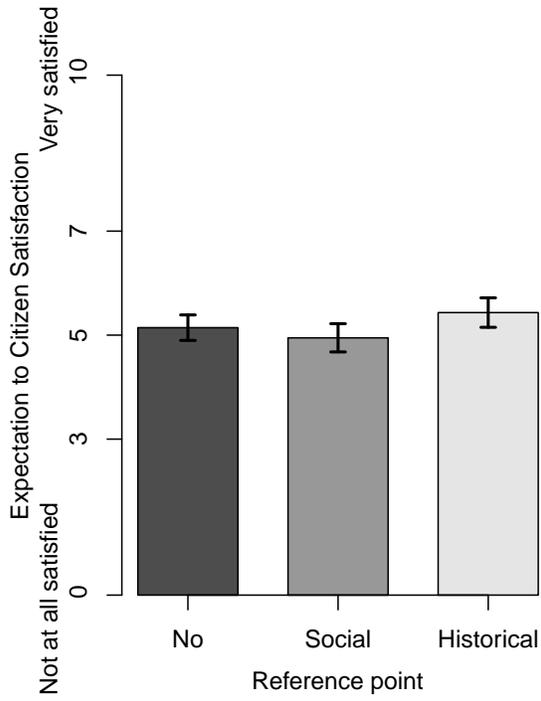
	<i>No comparison</i>	<i>Social comparison</i>		<i>Historical comparison</i>	
	I	I	II	I	II
<i>Exp. 1: Grades</i>					
School average	0.98** (0.12)	0.88** (0.13)		0.96** (0.12)	
Reference point average		-0.48** (0.12)		-0.64** (0.13)	
Relative performance			0.23 (0.20)		0.47* (0.19)
Relative bad performance			-0.68 (0.41)		-0.26 (0.40)
Interaction			0.51 (0.29)		0.57* (0.28)
Intercept	-1.28 (0.75)	2.43* (1.18)	5.65** (0.28)	3.20** (1.14)	5.80** (0.27)
<i>DF</i>	256	267	266	283	282
adj. <i>R</i> ²	0.22	0.18	0.18	0.25	0.26
<i>Exp. 2: Unemployment</i>					
Municipal rate	-0.67** (0.15)	-0.80** (0.14)		-0.93** (0.11)	
Reference point rate		0.52** (0.13)		0.47** (0.13)	
Relative performance			0.53* (0.23)		0.92** (0.17)
Relative bad performance			-0.90* (0.40)		-0.31 (0.38)
Interaction			-0.32 (0.33)		-0.60* (0.28)
Intercept	7.89** (0.92)	6.27** (1.14)	4.82** (0.29)	7.63** (1.07)	4.59** (0.25)
<i>DF</i>	276	272	271	249	248
adj. <i>R</i> ²	0.07	0.14	0.15	0.24	0.23

Coefficients obtained by OLS estimation. Standard errors in parentheses.

Significance levels correspond to: * $p < .05$ and ** $p < .01$ (all two-tailed tests).

Figure 1: Main effects of the experiments

(a) School performance



(b) Municipal unemployment

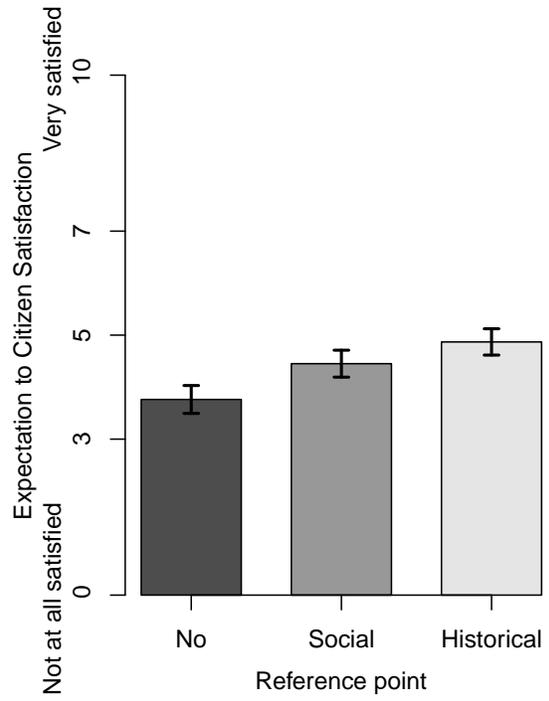
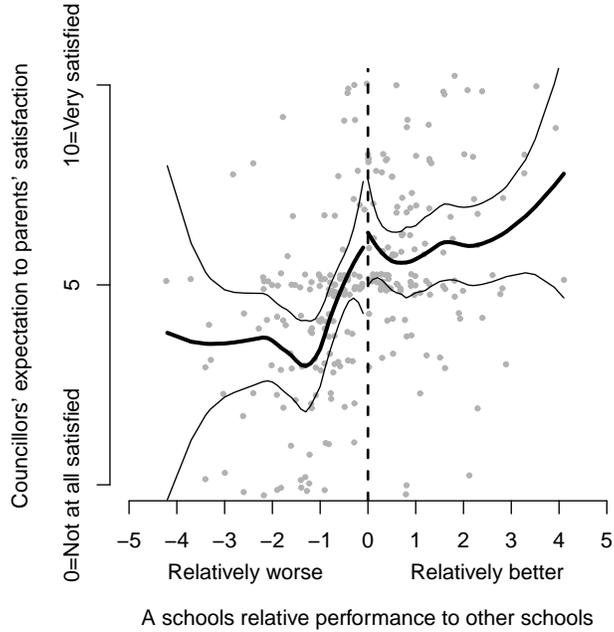


Figure 2: School performance experiment

(a) **Social reference point (n=270)**. Councilors' expectation to parents' evaluations plotted against the randomized relative performance (*school performance - other schools in same municipality*).



(b) **Historical reference point (n=286)**. Councilors' expectation to parents' evaluations plotted against the randomized relative performance (*school performance - same municipality last year*).

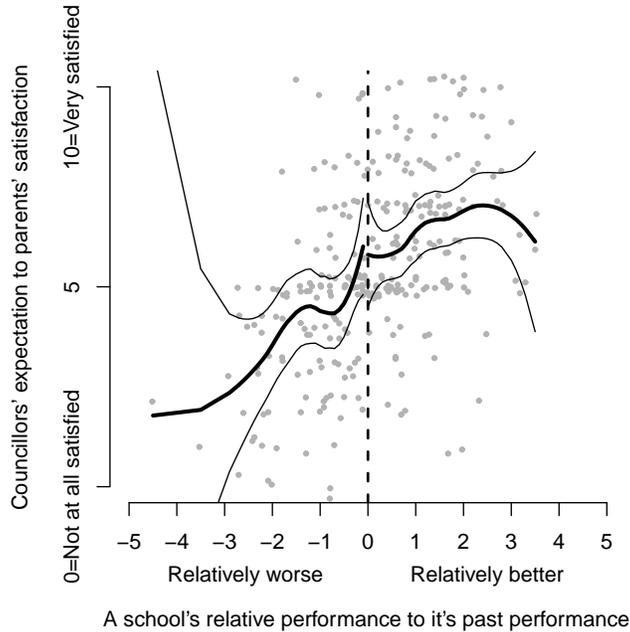
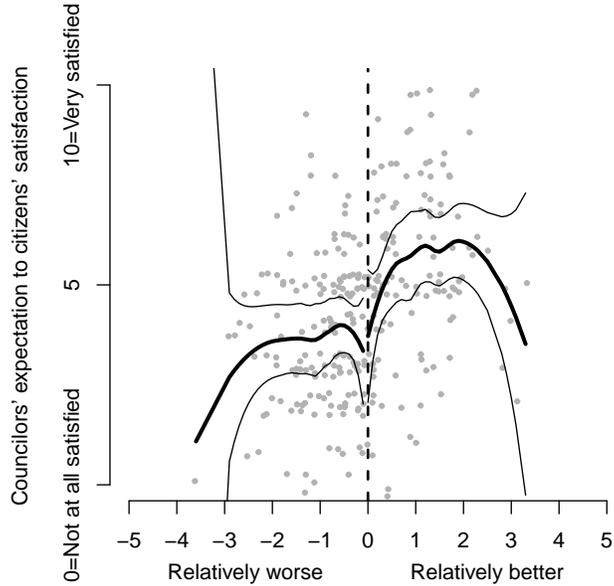


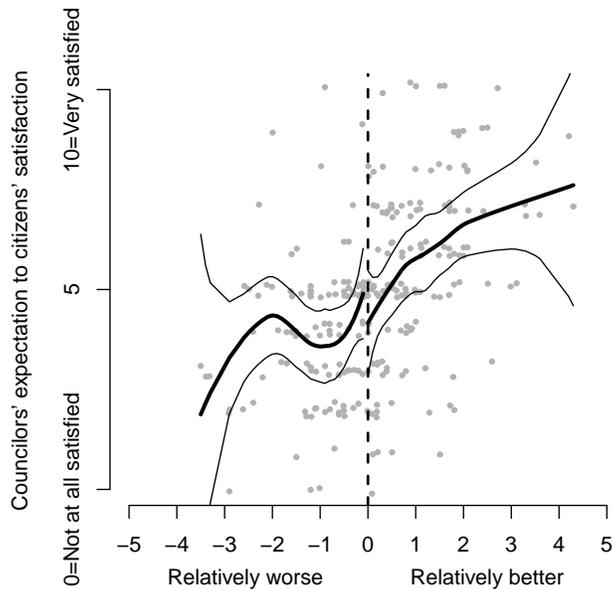
Figure 3: Municipal unemployment experiment

(a) **Social reference point (n=275)**. Councilors' expectation to citizens' evaluations plotted against the randomized relative performance (*the national unemployment rate - a municipality's unemployment rate*).



A municipality's unemployment relative to the national rate

(b) **Historical reference point (n=252)**. Councilors' expectation to citizens' evaluations plotted against the randomized relative performance (*A municipality's unemployment rate last year - the same municipality's unemployment rate today*).



A municipality's unemployment relative to the year before