Journal of Behavioral
Public Administration
Vol 2(1), pp. 1-12
DOI: 10.30636/jbpa.21.32

Citizen Assessments of Local Government Sustainability Performance: A Bayesian Approach

Aaron Deslatte*

Abstract: Citizens are increasingly critical information-processors, and the use of government performance information has become ubiquitous to the tenants of democratic accountability. Yet human perceptions of governmental policies and outcomes are increasingly partisan and resistant to updating. Partisan motivated reasoning can lead to inaccurate or biased assessments of both the merit of specific policies and governmental performance. Combining previous findings with a new experimental design, this study examines whether provision of performance information on local government implementation of federally initiated sustainability efforts ameliorates the partisan motivated reasoning of citizens. Employing Bayesian methods, the study finds evidence of attitude-strengthening in the face of disconfirming performance, as well as suggesting partisan cues may dampen this effect. A case is made for the use of Bayesian inference for experimental work on information-processing.

Keywords: Performance information, Bayesian inference, local government sustainability, motivated reasoning, affective intelligence
Supplements: Open data, Open materials

Public organizations are increasingly under pressure to demonstrate performance gains and preserve accountability to their citizens (Clarke & Margetts, 2014; Lavertu, 2016; Yang, 2016). Government performance information has therefore become ubiquitous to democratic responsiveness (Baekgaard, 2015; James, 2011; Moynihan & Pandey, 2005; Van Dooren & Van de Walle, 2016).

Communicating government performance to a diverse and politically tribalized citizenry remains one of the "big questions" of public administration (Moynihan, 2018) and poses potential tradeoffs. On the one hand, transparency may contribute to trust in government (Grimmelikhuijsen & Meijer, 2012) increased participation (Porumbescu, 2017) and legitimacy (De Fine Licht et al., 2014) depending upon the strategy for how information is presented (Piotrowski & Van Ryzin, 2007; Piotrowski, Grimmelikhuijsen, & Deat, 2017). On the other hand, human perceptions of governmental policies and outcomes are increasingly partisan and resistant to updating (Lodge and Taber 2013). Cognitive limitations and partisan motivated reasoning can lead to inaccurate or biased assessments of both the merit of specific policies (Bolsen, Druckman, & Cook, 2014) and the performance of government (Baekgaard & Serritzlew, 2016; Marvel, 2016).

The goal of this article is to demonstrate the advantages of a Bayesian inferential strategy for building cumulative knowledge to address such questions. Combining previous findings with a new experimental design, this study examines whether provision of performance information on local government implementation of federally initiated sustainability efforts ameliorates the motivated reasoning of citizens. The study focuses on the performance of local governments in achieving energy savings through the federal Energy Efficiency and Conservation Block Grant Program, established under the American Recovery and Reinvestment Act of 2009. This allows for a novel assessment of the impact of performance information on citizen

* Department of Public Administration, Northern Illinois University
Address correspondence to Aaron Deslatte at (mailto:adeslatte@niu.edu)

evaluation of inter-governmental performance outcomes. The evidence suggests attitude-strengthening occurs in the face of disconfirming performance information, although this strengthening is dampened by partisan cues.

Local Sustainability, Ideology and Performance Gaps

Local sustainability is a burgeoning area of government activity (Fiorino, 2010; Wang, Hawkins, & Lebredo, 2012; Opp & Saunders, 2013; Sharp, Daley, & Lynch, 2010). "Sustainability" as an organizing set of objectives is typically defined in accordance with the United Nations' Brundtland Commission report which articulated inter-generational goals of preserving or minimizing harm to the environment, the economy and social equity (Brundtland & Khalid, 1987). In recent years, hundreds of cities in the United States have adopted green "scorecards" and "climate action plans," while international organizations such as ICLEI-Local Governments for Sustainability boast of networks of more than 1,500 localities committed to reduce their carbon-footprints (Krause, 2011; Kousky & Schneider, 2003). While local sustainability policy commitments have been extensively explored, the question of whether local government activities square with prevailing citizen beliefs has drawn less attention (Krause, 2011; Yi, Krause, & Feiock, 2017).

Citizen ideology and partisan identification play a role in government performance assessments. This study defines ideology narrowly as the degree to which citizens believe government should play a more or less active role in fostering sustainability (Ellis & Stimson, 2012). This is consistent with its application in both political science studies on citizens' policy preferences (Bolsen, Druckman, & Cook, 2014) and public management research on the role that ideological beliefs play in assessing public performance (Baekgaard & Serritzlew, 2016).

The links between ideological beliefs, party identification and support for government policy have been well-established in political psychology literature. In their seminal work on the rationalization processes of voters, Milton Lodge and Charles Taber (2013) have demonstrated the dual-processes of affective and cognitive evaluation voters undergo when they receive new information on candidates or issues. When presented with new information, voters unconsciously activate affective tags associated with political objects in memory (Lodge & Taber, 2005). The positive or negative valence of these affective associations prompts "hot cognition," which colors the updating of their summary, online tally of beliefs and the processing of new information (Lodge & Taber, 2013).

Political science research has pinpointed the near automaticity of affective responses citizens display toward political candidates, groups and issues, where the positive or negative valence tagged to an object in memory is aroused prior to the cognitive evaluation (Lodge and Taber 2005; Taber and Lodge 2006). Motivated reasoning can occur when citizens evaluate policies differently depending on the strength of the frame used to describe them (Chong & Druckman, 2007; Laurian, Walker, & Crawford, 2017). Elite polarization surrounding environmental issues can activate citizen motivated reasoning when partisan endorsements overpower substantive evaluation of the issues (Druckman, Peterson, & Slothuus, 2013).

A missing link in these studies is whether the evaluation of government performance on divisive issues is similarly colored by activation of affective tags. When evaluating issues in which parties take opposing positions, it is possible citizens may become aware of the pros and cons and may not be able to immediately retrieve an affective tag, particularly when they are more ambivalent about an issue (Lodge & Taber, 2013; Zaller & Feldman, 1992; Zaller, 1992). While Lodge & Taber (2013) find "hot cognition" extends to evaluations of issues themselves, tests of affective activation in which citizens must evaluate support for issues based on varying levels of government performance are considerably rarer.

Here, activation of ideological beliefs and assessment of performance evaluation are interconnected cognitive processes. Policy support is presumed to be a based on prior attitudes, summarized by citizens' online tally, which introduces motivated bias into the consideration of new information (Gerber & Green, 1999; Lodge & Taber, 2013). While political science has focused on the manipulation of this information, public administration scholars have focused on how the positive or negative valence of even "straightforward" performance information produces asymmetrical responses. Indeed, perceptions of performance can influence how public managers and political elites

make decisions (Nielsen & Moynihan, 2017). Public administration researchers have also advanced behavioral models highlighting the special role that performance gaps play in triggering searches for innovative solutions (Rutherford & Meier, 2015; Salge, 2011).

Public administration researchers have found evidence of a negativity bias toward both public organizations in general (Hvidman & Andersen, 2016) and public performance at both the federal and local levels (James, 2011; Marvel, 2016). Performance information may be systematically misinterpreted based on prior beliefs and the affective evaluations of new information (Baekgaard & Serritzlew, 2016; Redlawsk, 2002; Redlawsk, Civettini, & Emmerson, 2010). This research has tended to give less attention to the role that both beliefs and partisan cues play in asymmetrical citizen responses to performance gaps. In the next section, this article provides a fuller summary of this literature to justify a Bayesian inferential approach.

Incorporating Prior Knowledge of Citizen Performance Evaluation

Further complicating the chore of assessing the influence of performance information is the difficulty in making cumulative knowledge claims when studies fail to replicate across different political and environmental contexts. For nearly two decades, some scholars have argued a Bayesian approach was ideal for public administration research because of reliance on population data rather than random pulls which can be repeatedly re-sampled (Gill & Witko, 2013; Gill & Meier, 2000; Meier, Favero, & Zhu, 2015). Bayesian methods have begun appearing in public administration research to overcome certain data limitations (e.g., Sinclair & Whitford, 2012; Zhu, Robinson, & Torenvlied 2015). Typically, these studies have used diffuse (or uninformed) priors (Deslatte & Swann, 2017; Deslatte, Swann, & Feiock, 2017), which essentially fit a likelihood model estimating the posterior median, mean or other quantiles of interest based solely on new data (Gill & Witko, 2013). While often appropriate, the use of uninformed priors nullifies one of the key advantages of Bayesian estimation: leveraging existing knowledge to quantitatively update our beliefs about the phenomena under investigation.

The Bayesian paradigm provides advantages in dealing with differing study contexts (Boyne et al., 2005). Bayesian inference differs from the frequentist assumption that phenomena of interest have fixed but unknowable values. Bayesian estimation assumes the opposite that these parameters come from a random probability distribution and can by summarized more intuitively via probability statements. The approach involves estimating posterior parameters for quantities of interest by combining new data via a likelihood function with a prior distribution derived from existing knowledge. Given the prevalence of replication failures and "desk shelf" effects in experimental studies, a Bayesian evaluation of prior research could provide new insights on the complexity and dynamism of context through evidence which may be otherwise disregarded for not reaching an arbitrarily set level of statistical significance. Instead, the Bayesian approach allows us to quantify our uncertainty through intuitive statements about the probability of observing an effect.

Informed prior distributions typically come from knowledge about the size and direction of relationships in previous studies. For instance, motivated reasoning has been shown to influence how citizens process performance information for contentious programs such as the U.S. Affordable Care Act (James & Van Ryzin, 2017). The relationship between provision of performance information and citizen assessments is also highly susceptible to positive and negative framing effects (Olsen, 2015), and biases in interpreting numerical performance information (Olsen, 2018). In one study of relevance to this research, James (2011) found evidence that credible performance information on English local governments can be used to manage citizens' positive expectations, although normative expectations were more resilient to this type of approach, evidence of a negativity bias (James, 2011). Recent work has found that the types of messaging strategies government officials take -- straight provision of information versus stronger framing of performance within -- disproportionately impacts engaged and disengaged citizens (Piotrowski, Grimmelikhuijsen, & Deat, 2017). Engaged citizens responded more favorably to straight information provision while less-engaged citizens require a "transformational" communications strategy.

To some degree, it may be possible that actions of any type are rewarded by the public, regardless of outcome (Olsen, 2017c). However, it is rea-

sonable to assume that citizens asked to assess episodic details of performance information on a controversial policy will display systemic partisan motivated reasoning (Bolsen, Druckman, & Cook, 2014), and that this information will differentially impact engaged versus passive individuals (Piotrowski, Grimmelikhuijsen, & Deat, 2017).

This information is used to construct appropriate, informed priors in an online experiment exploring citizen assessments of high and low performance within a federal program designed to encourage local government energy savings and conservation. The U.S. Department of Energy (DOE) Energy Efficiency and Conservation Block Grant (EECBG) Program was passed by Congress in 2009 to quickly create jobs and generate energy savings through grants to local governments (Terman & Feiock, 2015; Terman, 2015). Grants could be used for a wide array of projects, from installing HVAC systems in affordable housing units to energy retrofitting government buildings, buying fuel-efficient vehicle fleets and solar panels, to curbing greenhouse gas emissions from landfills (DOE, 2011; GAO, 2011, 2012). Based on a systematic survey of media coverage of EECBG projects, citizens in communities across the U.S. could have encountered a variety of episodic and numeric-influenced narratives of local government activities and outcomes under the program. Thus, the EECBG represents an ideal testbed for constructing realistic vignettes of inter-governmental performance and augmenting the findings from prior literature for how citizens are likely to process such information.

Experimental Data and Design

Citizen motivated reasoning was tested in an online survey experiment. Subjects for the survey were recruited via Amazon's Mechanical Turk (MTurk), the online labor market in which individuals are paid small sums for human intelligence tasks such as participating in market research or academic surveys.

Performance-management systems in local governments are dominated by quantitative information (Julnes & Holzer, 2001; Kelly & Swindell, 2002), which scholars have long posited as essential to communicating unambiguous organizational activities, outputs and outcomes to the public (Moynihan, 2008; Yang & Holzer, 2006). Recent evidence suggests citizens are influenced more by episodic than statistical data (Olsen, 2017b), and may not discern between more- or less-precise information (Olsen, 2018). Because quantitative information-processing can be systematically biased (James & Olsen, 2017; Olsen, 2017a), and government performance activities, outputs and outcomes in the sustainability arena can be inherently ambiguous (Deslatte & Swann, 2016), this study seeks to explore citizen evaluations of local performance under conditions of ambiguity on one objective: energy efficiency.

This experiment employed a 2 x 3 between-subjects design in which participants were assigned to one of six groups and presented with hypothetical vignettes of episodic performance based on actual local government experiences with the EECBG program (GAO, 2011). The six groups included: a) a *control group* in which the EECBG program is described but no partisan cues or performance information are provided; b) a *baseline partisan cue* group in which participants were informed that "[t]he Congressional vote to authorize the program was largely along party lines, with Democrats in the House and Senate overwhelmingly voting in favor and Republicans mostly voting against it"; c) a *high-performance, no-partisan-cue* group, in which participants were told to imagine their local government had used grants to install LED streetlights, energy efficiency upgrades to city buildings and solar panels on the roof of the City Hall for electric-vehicle charging stations, resulting in the city saving costs on energy; d) a *low-performance, no-partisan-cue* group, in which participants were told the same activities had resulting in no savings for the City; e) a *high-performance, partisan-cue* group in which both high performance information and the partisan cue was provided; and f) a *low-performance, partisan-cue* group. Participants (N=1,001) were paid $0.70 for completing the surveys. Table 1 describes the demographics of the full sample of respondents by reported gender, ethnicity/race, income group, education, and party affiliation, and group subsamples demonstrated no statistically significant differences.

To measure existing beliefs regarding the role of government in sustainability issues, respondents were asked on a four-point scale what level of responsibility (1 = "Not at all responsible," 4 = "Completely responsible") the federal government, state governments, local governments, nonprofits, corporations and individuals had for "taking actions to protect the environment." Factor

Table 1
Sample Description (N=1,001)

	Female (%)	White (%)	Income (modal category)	Education (modal category)	Dem. (%)	Ind. (%)
Full Sample	50.6	79.9	$45,000-$59,999	2-year degree	41.1	27.9

analysis showed the federal, state and local government items loaded onto the same factor (factor scores each > .6) with a Cronbach's alpha of .83. The combined *environmental ideology* index was then re-scaled to run from 0-100 with higher values reflecting a greater belief in governmental responsibility for environmental stewardship. For the two outcomes, respondents were asked whether the federal government should continue to appropriate funds for the EECBG Program, with responses along a 7-point scale from "completely disagree" to "completely agree." A second question asked whether their local government should continue to fund the program after the federal grants had been exhausted.

Prior distributions are statements about the probability of a particular parameter, β, independent of new information (Gill & Witko, 2013). We say that $p(\beta) = k, a < \beta < b$ when we want to specify a uniform prior, meaning that $p(\beta)$ is constant within the domain $[a, b]$. Given the recent behavioral public administration findings on performance information use, the analysis estimates models with both informed and uninformed priors for the relationship between *environmental ideology* and citizen assessments of local government sustainability performance. Based on previous studies, I expect that citizens' assessments of performance information will display systemic partisan motivated reasoning (Bolsen, Druckman, & Cook, 2014), and that this information will differentially impact engaged versus passive individuals (Piotrowski, Grimmelikhuijsen, & Deat, 2017). Using an informed prior to express this belief simply requires specifying normally distributed priors, $\beta \sim N(\mu, \sigma^2)$, for the *environmental ideology* mean and variance, with $\mu = 1$ to indicate prior belief in a positive direct association between *ideology* and *federal* and *local* support, and $\mu = -.5$ for the interaction term of the treatment and *ideology*. Belief that high and low performance information can attenuate this effect can also be expressed through a similar positive, negative or "skeptical" prior of $\mu = 0$, with a large σ^2 used when there is less certainty of the relationship. All other variables in the models were given uninformed priors.

The models were estimated using Stata 14. Bayesian estimation involves sampling from a simulated probability distribution, which was done using a Markov chain Monte Carlo (MCMC) method utilizing a Metropolis-Hastings sampling algorithm. To improve convergence, 240,000 iterations were run with a 40,000 iteration "burn-in" period. Diagnostic plots indicated strong evidence of model convergence. The models with informed and uninformed prior distributions were then compared via a form of sensitivity testing for model selection called a Bayes factor.

Results

The mean scores of support for federal and local government program continuation show some descriptive group differences.[1] In the seven-point scale, M ≥ 5 signifies support for both programs. Table 2 shows that mean scores for the partisan baseline group dip slightly for both federal support (M=4.99, SD=1.73) and local support (M= 4.94, SD=1.76) compared to the control group. An analysis of variance (ANOVA) shows that high performance information without a partisan cue increased mean support for local policy continuation relative to the control group, $F(1,331) = 6.68$; $p < .05$.

Similarly, low performance information without a partisan cue lowered mean support for the local policy, $F(1,331) = 6.38$; $p < .05$. When a partisan cue is included, we see a strengthening of the statistical significance for high-performance information, $F(1,332) = 10.26$; $p < .01$, and a weakening significance for low-performance, $F(1,332) =$

Table 2
Mean Support for Federal and Local Energy Efficiency Program

	Federal Program (scale: 1-7)	Local Program (scale:1-7)
C1: Control Group (N=166)	5.26 (1.58)	5.08 (1.55)
C2: Partisan Baseline (N=167)	4.99 (1.73)	4.94 (1.76)
T1: High Performance, No Partisan Cue (N=167)	5.46 (1.6)	5.51** (1.48)
T2: Low Performance, No Partisan Cue (N=167)	4.69*** (1.88)	4.61** (1.85)
T3: High Performance, Partisan Cue (N=167)	5.45** (1.58)	5.5*** (1.39)
T4: Low Performance, Partisan Cue (N=167)	4.59** (1.93)	4.57* (1.8)

Notes: Standard Deviations appear in the parentheses.
** p < 0.1, ** p < 0.05, *** p < 0.01 (two-tailed test).*

3.62, $p < .1$. Federal program mean support displays similar results, with the exception of the high-performance, no partisan cue group. High and low performance appear to impact mean support for program continuation, with or without a partisan cue.

Results for the Bayesian ordered probit models for local support without partisan cues are reported in Table 3. Bayes factors (BF) are the appropriate method for selecting a model among a set of candidates, because unlike other information criteria approaches (BIC, AIC, DIC) they account for the use of informed priors. Bayes factors are the ratios of the marginal likelihoods of two comparison models. Across all the groups, M1 represents the base model with uninformed priors, M2 is the informed prior model, and $\mathrm{BF}_{M1,M2} > 20$, which is strong evidence that the models using informative priors were superior. We interpret results with Bayesian interval hypothesis tests reported in Table 5, which allow for making probabilistic statements about whether a parameter falls along a specified interval (for instance, the probability of observing $\beta > 0$ or $\beta < 0$). True to convention, prior support for government involvement in sustainability appears to positively influence support for local sustainability activities in all models, while we observe a .71 probability that high-performance has a positive effect.

The assessment of low performance is key to motivated reasoning arguments. In the low-performance, no-partisan-cue model, there is a .13 probability of observing the expected negative effect of lower-performance information on policy support. In other words, the evidence suggests that provision of a concise, episodic negative assessment of government performance has an 87% chance of having a positive direct influence on support. The negative interaction effect of ideology and low-performance also conforms with previous findings which suggest more engaged citizens are more responsive to informational appeals and require less emotionally appealing or transformative packaging of performance information (Piotrowski, Grimmelikhuijsen, & Deat, 2017). Here, the interaction suggests that support for the local policy among those with stronger prior pro-government beliefs is "correctly" attenuated downward by low-performance information.

The attitude-strengthening interaction of affect associated with prior beliefs with the affective response to moderately disconfirming information could explain the attitude-strengthening among citizens predisposed to support sustainability efforts (Redlawsk, 2002; Redlawsk, Civettini, & Emmerson, 2010). This polarization effect leads citizens with existing positive attitudes about a government role in sustainability to become more positive in the face of some level of negative performance information (Lord, Ross, & Lepper, 1979),

Table 3
Bayesian Ordered Probit Regressions: Local Support, No Partisan Cue

Performance Cue:	High (T1/C1)			Low (T2/C1)		
	Mean	MCSE	95% C.I.	Mean	MCSE	95% C.I.
Environmental ideology	.023	.0001	.017; .029	.022	.00009	.015; .028
Treatment dummy	.176	.008	-.40; .799	.304	.007	-.233; .831
Treatment * Ideology	.003	.0001	-.006; .013	-.011	.0001	-.019; -.002
Male	-.108	.0009	-.339; .126	.068	.0009	-.159; .295
White	.216	.002	-.069; .503	-.016	.002	-.304; .268
Democrat	.726	.002	.446; 1.01	.84	.002	.555; 1.13
Independent	.121	.001	-.172; .414	.199	.001	-.091; .488
Education	.01	.0008	-.088; .108	.062	.0007	-.032; .153
Income	.017	.0004	-.047; .082	.046	.0003	-.016; .109
Age	.038	.0005	-.047; .124	-.014	.0006	-.072; .099
MCMC			200,000			200,000
Acceptance rate			.429			.414
Efficiency			.027			.027
N			333			333

although some studies have found this affective response may be curvilinear and correct itself as negative information proliferates (Redlawsk, Civettini, & Emmerson, 2010). To test this explanation, we need to examine whether partisan cues strengthen this "perverse" reaction (favoring a cognitive dissonance explanation) or weaken it (evidence of an affective tipping point).

Turning to the partisan cue models reported in Table 4, the evidence suggests partisan cues increase support for the policy among Democrats in both high- and low-performance groups nearly identically. Partisan motivated reasoning would explain this result. But the Bayesian hypothesis tests find a 99% chance that high-performance information provision also positively impacts support and a 77% chance that low-performance produces the opposite, negative effect. Given the lack of such "accuracy" evidence in the models with no partisan cue, these findings would seem to support an affective intelligence argument that partisan cues -- in this case, the knowledge that the program was adopted amid partisan conflict -- alert respondents to the potential contentiousness of the issue. Once alerted, respondents shift to active processing in anticipation of negative affective stimuli. In this

Table 4
Bayesian Ordered Probit Regressions: Local Support, Partisan Cue

Performance Cue:	High (T3/C2)			Low (T4/C2)		
	Mean	MCSE	95% C.I.	Mean	MCSE	95% C.I.
Environmental ideology	.022	.00009	.016; .029	.023	.00008	.016; .03
Treatment dummy	.662	.009	.061; 1.27	-.217	.008	-.799; .365
Treatment * Ideology	-.007	.0002	-.016; .002	.002	.0001	-.011; .008
Male	.049	.0009	-.183; .282	.077	.0008	-.147; .302
White	-.161	.002	-.437; .112	-.039	.002	-.31; .229
Democrat	.978	.002	.684; 1.27	.971	.001	.688; 1.25
Independent	.402	.001	.095; .707	.394	.001	.105; .683
Education	.044	.0008	-.05; .139	.073	.0008	-.018; .164
Income	.019	.0003	-.044; .083	-.016	.0003	-.077; .044
Age	.055	.0006	-.029; .139	.02	.0006	-.06; .102
MCMC			200,000			200,000
Acceptance rate			.438			.431
Efficiency			.021			.024
N			334			334

Table 5
Bayesian Interval Hypothesis Tests (w/ prior expected probabilities)

	No Partisan Cue		Partisan Cue	
Performance Cue:	High	Low	High	Low
Ideology	.999 ($p > 0$)	.999 ($p > 0$)	.999 ($p > 0$)	.999 ($p > 0$)
Treatment	.716 ($p > 0$)	.132 ($p < 0$)	.999 ($p > 0$)	.766 ($p < 0$)
T * Ideology	.249 ($p < 0$)	.999 ($p < 0$)	.928 ($p < 0$)	.641 ($p < 0$)

sense, the minor anxiety associated with consideration of potentially disconfirming information leads to an information-processing outcome more akin to Bayesian updating, controlling for the motivated

reasoning which also influences the evaluation of this new information. While this process is not directly tested in this experiment, it is the most plausible explanation given the evidence, and subsequent studies should both attempt to replicate such findings and test the limitations of such a Bayesian-like updating process under stronger and weaker performance cues.

Conclusion

Organizational performance remains a paramount concern in public administration. Yet, human beings adjudicating government performance usually fall far short of being rational Bayesian updaters. In order to more effectively utilize performance information to elicit public support for programs and policies, public administrators need a more holistic understanding of how citizens process performance information in an increasingly tribalized political environment.

The performance management literature suggests negative performance information leads to negativity bias in citizen assessments (Baekgaard & Serritzlew, 2016; James, 2011; Marvel, 2016). However, theories of systems justification and cognitive dissonance might also lead one to expect that when presented with negative information for a policy citizens favor, they display differential degrees of tolerance of bad performance in order to ease discomfort with the disconfirming evidence (Jost, Banaji, & Nosek, 2004). Yet another explanation comes from behavioral research on affective "tipping points" (Redlawsk, 2002; Redlawsk, Civettini, & Emmerson, 2010). Voters maintain a summary online tally (a summary evaluative feeling) for both political candidates and policies. Motivated reasoning studies suggest the affective evaluation of new information is conditioned upon the prior online tally, which can strengthen their views when confronted with incongruous stimuli (Taber & Lodge, 2006). Studies of affective intelligence suggest that at some tipping point, new stimuli incongruent with existing expectations shift information-processing from a passive, subconscious state to an active one where information is more carefully examined (Marcus & MacKuen, 1993).

Filling a gap in extant behavioral research, this study suggests citizens may reach an affective tipping-point in evaluating a partisan-charged policy via one type of performance information: simple frames of episodic "success" or "failure." However, establishing such a claim will require additional work to unpack the causal mechanisms and contextual caveats. Given research showing biased assessments of numeric information (Olsen, 2018), episodic framing of performance may have the potential to be more accurately assessed by citizens. Episodic and numeric performance comparisons are necessary to further explore the mechanisms of affective intelligence and the magnitude of disconfirming information required to flip citizens to more accurate updating of beliefs. Beyond the message, the type of delivery mechanism likely matters. How does the strength of a positive or negative frame influence citizen assessment? How do multiple or conflicting accounts of performance moderate or mediate this effect? Lastly, eroding trust in public officials and institutions is a systemic governance problem. Do local public administrators benefit from a "messenger effect," or do citizens fail to distinguish between political frames and neutral performance appraisals from unelected officials?

Evaluating performance is easier said than done in the realm of sustainability. Success can mean dollars saved on energy bills, health improvements, or climate change adaptations delivering intergenerational benefits often discounted by present citizens. No experimental design can completely eliminate bias, and it is also necessary to replicate and extend this study to minimize the possibility of design or instrument error. More studies using different types of outcomes of government sustainability are also required to better establish the validity and replicability of these findings.

Lastly, the work on motivated reasoning clearly demonstrates that context matters. This is the principal argument by Gelman (2014) that a Bayesian paradigm is ideal for dealing with replication failures and "desk shelf" publication bias. This study is a salient example, because the evidence on partisan cues would have been disregarded under a frequentist approach for not reaching an arbitrarily set level of statistical significance. Instead, the Bayesian approach provides a method for quantifying our uncertainty. Much like a weather prediction of rain, readers and researchers can then judge for themselves whether the probability merits closer consideration.

Acknowledgement

This study was supported by a National Research Foundation of Korea Grant from the Korean Government [NRF-2017S1A3A2065838].

Notes

1. I interpret these differences using a frequentist, analysis of variance (ANOVA), because the initial evaluation is only whether statistically significant differences across groups are present. As such, this portion of the analysis does not need to bring prior information into the estimation.

References

Baekgaard, M. (2015). Performance information and citizen service attitudes: Do cost information and service use affect the relationship? *International Public Management Journal*, *18*(2), 228–245.

Baekgaard, M., & Serritzlew, S. (2016). Interpreting performance information: Motivated reasoning or unbiased comprehension. *Public Administration Review*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/puar.12406/full

Bolsen, T., Druckman, J. N., & Cook, F. L. (2014). The influence of partisan motivated reasoning on public opinion. *Political Behavior*, *36*(2), 235–262.

Boyne, G. A., Meier, K. J., O'Toole, L. J., & Walker, R. M. (2005). Where next? Research directions on performance in public organizations. *Journal of Public Administration Research and Theory*, *15*(4), 633–639.

Brundtland, G., & Khalid, M. (1987). UN Brundtland commission report. *Our Common Future*.

Chong, D., & Druckman, J. N. (2007). Framing public opinion in competitive democracies. *The American Political Science Review*, *101*(4), 637–655.

Clarke, A., & Margetts, H. (2014). Governments and citizens getting to know each other? Open, closed, and big data in public management reform. *Policy & Internet*, *6*(4), 393–417.

De Fine Licht, J., Naurin, D., Esaiasson, P., & Gilljam, M. (2014). When does transparency generate legitimacy? Experimenting on a context-bound relationship. *Governance*, *27*(1), 111–134.

Deslatte, A., & Swann, W. L. (2016). Is the price right? Gauging the marketplace for local sustainable policy tools. *Journal of Urban Affairs*, *38*(4), 581–596.

Deslatte, A., & Swann, W. L. (2017). Context matters: A Bayesian analysis of how organizational environments shape the strategic management of sustainable development. *Public Administration*, *95*(3), 807–824.

Deslatte, A., Swann, W. L., & Feiock, R. C. (2017). Three sides of the same coin? A Bayesian analysis of strategic management, comprehensive planning, and inclusionary values in land use. *Journal of Public Administration Research and Theory*, *27*(3), 415–432.

DOE. (2011). *EECBG Program Notice 11-002*. Washington, DC: Department of Energy.

Ellis, C., & Stimson, J. A. (2012). *Ideology in America*. Cambridge University Press.

Fiorino, D. J. (2010). Sustainability as a conceptual focus for public administration. *Public Administration Review*, 70, s78–s88.

GAO. (2011). *EECBG under the Recovery Act, 11–279*. Washington, DC: Government Printing Office.

———. (2012). *Grants to state and local governments. 12-1016*. Washington, DC: Government Printing Office.

Gelman, A. (2014). The connection between varying treatment effects and the crisis of unreplicable research: A Bayesian perspective. *Journal of Management*, *41*(2), 632–643.

Gerber, A., & Green, D. (1999). Misperceptions about perceptual bias. *Annual Review of Political Science*, *2*(1), 189–210.

Gill, J., & Meier, K. J. (2000). Public administration research and practice: A methodological manifesto. *Journal of Public Administration Research and Theory*, *10*(1), 157–199.

Gill, J., & Witko, C. (2013). Bayesian analytical methods: A methodological prescription for public administration. *Journal of Public Administration Research and Theory*, *23*(2), 457–494.

Grimmelikhuijsen, S. G., & Meijer, A. J. (2012). Effects of transparency on the perceived trustworthiness of a government organization: Evidence from an online experiment. *Journal of Public Administration Research and Theory*, *24*(1), 137–157.

James, O. (2011). Managing citizens' expectations of public service performance: Evidence from observation and experimentation in local government. *Public Administration*, 89(4), 1419-1435.

James, O. (2011). 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, O., & Van Ryzin, G. G. (2017). Motivated reasoning about public performance: An experimental study of how citizens judge the Affordable Care Act. *Journal of Public Administration Research and Theory*, *27*(1), 197–209.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 881–919.

Julnes, P. de L., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, *61*(6), 693–708.

Kelly, J. M., & Swindell, D. (2002). A multiple-indicator approach to municipal service evaluation: Correlating performance measurement and citizen satisfaction across jurisdictions. *Public Administration Review*, *62*(5), 610–621.

Kousky, C., & Schneider, S. H. (2003). Global climate policy: will cities lead the way? *Climate Policy*, *3*(4), 359–372.

Krause, R. M. (2011). Symbolic or substantive policy? Measuring the extent of local commitment to climate protection. *Environment and Planning. C, Government & Policy*, *29*(1), 46–62.

Laurian, L., Walker, M., & Crawford, J. (2017). Implementing environmental sustainability in local government: The impacts of framing, agency culture, and structure in US cities and counties. *International Journal of Public Administration*, *40*(3), 270–283.

Lavertu, S. (2016). We all need help: "Big data" and the mismeasure of public administration. *Public Administration Review*, *76*(6), 864–872.

Lodge, M., & Taber, C. S. (2005). The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. *Political Psychology*, *26*(3), 455–482.

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge University Press.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098.

Marcus, G. E., & MacKuen, M. B. (1993). Anxiety, enthusiasm, and the vote: The emotional underpinnings of learning and involvement during presidential campaigns. *The American Political Science Review*, *87*(3), 672–685.

Marvel, J. D. (2016). Unconscious bias in citizens' evaluations of public sector performance. *Journal of Public Administration Research and Theory*, *26*(1), 143–158.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *The American Psychologist*, *70*(6), 487.

Meier, K. J., Favero, N., & Zhu, L. (2015). Performance gaps and managerial decisions: A Bayesian decision theory of managerial action. *Journal of Public Administration Research and Theory*, *25*(4), 1221–1246.

Moynihan, D. (2018). A great schism approaching? Towards a micro and macro public administration. *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.15

Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. Georgetown University Press.

Moynihan, D. P., & Pandey, S. K. (2005). Testing how management matters in an era of government by performance management. *Journal of Public Administration Research and Theory*, *15*(3), 421–439.

Nielsen, P. A., & Moynihan, D. P. (2017). How do politicians attribute bureaucratic responsibility for performance? Negativity bias and interest group advocacy. *Journal of Public Administration Research and Theory*, *27*(2), 269–283.

Olsen, A. L. (2015). Citizen (dis)satisfaction: An experimental equivalence framing study. *Public Administration Review*, *75*(3), 469–478.

Olsen, A. L. (2017a). Citizens and public performance measures: Making sense of performance information. In *Experiments in public management research: Challenges and contributions* (pp. 270–290). Cambridge University Press.

Olsen, A. L. (2017b). Human interest or hard numbers? Experiments on citizens' selection, exposure, and recall of performance information. *Public Administration Review*, *77*(3), 408–420.

Olsen, A. L. (2017c). Responding to problems: actions are rewarded, regardless of the outcome. *Public Management Review*, *19*(9), 1352–1364.

Olsen, A. L. (2018). Precise performance: Do citizens rely on numerical precision as a cue of confidence? *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.19

Opp, S. M., & Saunders, K. L. (2013). Pillar talk: local sustainability initiatives and policies in the United States—finding evidence of the "three E's": economic development, environmental protection, and social equity. *Urban Affairs Review*, 49(5), 678-717.

Piotrowski, S., Grimmelikhuijsen, S., & Deat, F. (2017). Numbers over narratives? How government message strategies affect citizens' attitudes. *Public Performance & Management Review*, 1–24.

Piotrowski, S. J., & Van Ryzin, G. G. (2007). Citizen attitudes toward transparency in local government. *American Review of Public Administration*, *37*(3), 306–323.

Porumbescu, G. (2017). Linking transparency to trust in government and voice. *American Review of Public Administration*, *47*(5), 520–537.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *The Journal of Politics*, *64*(4), 1021–1044.

Redlawsk, D. P., Civettini, A. J. W., & Emmerson, K. M. (2010). The affective tipping point: Do motivated reasoners ever "get it"? *Political Psychology*, *31*(4), 563–593.

Rutherford, A., & Meier, K. J. (2015). Managerial goals in a performance-driven system: Theory and empirical tests in higher education. *Public Administration*, *93*(1), 17–33.

Salge, T. O. (2011). A behavioral model of innovative search: Evidence from public hospital services. *Journal of Public Administration Research and Theory*, *21*(1), 181–210.

Sharp, E. B., Daley, D. M., & Lynch, M. S. (2010). Understanding local adoption and implementation of climate change mitigation policy. *Urban Affairs Review*, *47*(3), 433–457.

Sinclair, A. H. & Whitford, A. B. (2012). Separation and integration in public health: Evidence from organizational structure in the States. *Journal of Public Administration Research and Theory*, 23 (1): 55-77.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769.

Terman, J. N. (2015). Performance goal achievement in fiscal federalism: The influence of state partisan environments and regulatory regimes. *Policy Studies Journal: The Journal of the Policy Studies Organization*, *43*(3), 333–354.

Terman, J. N., & Feiock, R. C. (2015). Improving outcomes in fiscal federalism: Local political leadership and administrative capacity. *Journal of Public Administration Research and Theory*, *25*(4), 1059–1080.

Van Dooren, W., & Van de Walle, S. (2016). *Performance information in the public sector: How it is used*. Springer.

Wang, X. H., Hawkins, C. V., & Lebredo, N. (2012). Capacity to sustain sustainability: A study of US cities. *Public Administration*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6210.2012.02566.x/full

Yang, K. (2016). Creating public value and institutional innovations across boundaries: An integrative process of participation, legitimation, and implementation. *Public Administration Review*, *76*(6), 873–885.

Yang, K., & Holzer, M. (2006). The Performance-trust link: Implications for performance measurement. *Public Administration Review*, *66*(1), 114–126.

Yi, H., Krause, R. M., & Feiock, R. C. (2017). Back-pedaling or continuing quietly? Assessing the impact of ICLEI membership termination on cities' sustainability actions. *Environmental Politics*. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09644016.2016.1244968

Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American Journal of Political Science*, 579–616.

Zaller, J. R., & R, Z. J. (1992). *The nature and origins of mass opinion*. Cambridge University Press.

Zhu, L., Robinson, S. E. & Torenvlied, R. (2015). A Bayesian approach to measurement bias in networking studies. *The American Review of Public Administration*, 45 (5): 542-564.