Vogel & Xu (JBPA 2021 - Everything hacked? What is the evidential value of the experimental public administration literature?):

Supplement

S1: *p*-curve disclosure table

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Alon-Barkat, 2020		Symbolic elements in communications are likely to have a greater positive effect on citizens' attitudes about government organizations and policies the lower the perceived personal relevance of the communication.	one factor between subject	t-test	The real symbols condition (in the left panel) has a positive and significant main effect on trust in that policy plan compared with the no symbols condition	t(232)=3.197	No	Effect calculated from text (p. 558)
An et al., 2019		H1: The effect of transformational leadership training on managers' transformational leadership behavior will be larger in public organizations compared to private organizations.	one factor between subject	OLS regression coefficient	The public × transformational coefficient (+7.937) indicates the difference in treatment effect between the public sector and the private sector, a statistically significant difference.	r(2437)=0.051	No	Effect calculated from Table 1, Model 3
Andersen & Guul, 2019	1	In sum, evidence suggests that minorities are being discriminated in many different settings in the United States and in some settings outside the United States as well.	2x2 between subject	OLS regression coefficient	Model 1 in the table shows that the minority name cue affects teachers' willingness to include the student in their own classroom negatively and significantly (p < .002).	r(842)=-0.110	No	Effect calculated from Table 6, Column 1
Andersen & Guul, 2019	4	In sum, evidence suggests that minorities are being discriminated in many different settings in the United States and in some settings outside the United States as well.	2x2 between subject	OLS regression coefficient	Model 1 confirms the finding from study 1 (and previous research) that the minority cue has an overall negative effect on the willingness to include a student with special needs in the classroom (p < .004)	r(157)=-0.233	No	Effect calculated from Table 13, Column 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Andersen & Hjortskov, 2016	1	To test whether perceived performance can also be manipulated in areas where people have personal experience, we designed a field experiment (study 1) with manipulated budget information	2x2 between subject	Odds ratios	the cutback treatment affects the perception of the experiences with the more abstract performance in a negative direction.	r(676)=0.144	No	Effect calculated from Table 4, Model 1
Andersen & Jakobsen, 2018	2	The more pressure from political principals for adoption of an organizational innovation, the more likely an organization is to adopt the innovation.	2x3 between subject	OLS regression coefficient	The effect on the tactics question suggests that this effect on investment in expertise comes exactly because they evaluate that the net benefit of implementing the strategy would exceed the (political) costs of not doing it. When the three outcomes are joined in one adoption index, the effect of the political pressure is significant.	r(738)=0.076	No	Effect calculated from Table 3, Model 4
Andersen & Moynihan, 2018		Hypothesis 1: The introduction of a newcomer will result in greater team reflection processes.	field experiment	regression coefficient	Collectively, the results show that the introduction of newcomers, whether they share the dominant educational background of existing staff or not, is associated with more team reflection processes	r(495)=0.144	No	Effect calculated from Table 1, Model 2
Andersen, 2017		Bureaucrats with minority background will all else being equal be more positive toward a program that promotes the interest of minority groups than white bureaucrats will.	one factor between subject	OLS regression coefficient	Figure 1 confirms that principals with minority background in the control group are indeed significantly more positive toward the program than white bureaucrats.	r(399)=0.201	No	Effect calculated from Table A4, Model 1
Anderson & Stritch, 2016		Increases in goal clarity are associated with increases in performance.	2x3 between subject	ANOVA	A two-way ANOVA indicates the treatment effect of goal clarity is statistically significant for quantity transcribed (F (2, 212) = 5.89, p = 0.00) and quantity transcribed accurately (F (2, 212) = 4.15, p = 0.02) but there is no statistically significant effect of task significance or the interaction of clarity and significance	F(2,212)=5.89	No	

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Avellaneda & Olvera, 2018		H1: Chief executives are more likely to accept an influx of immigrants when the cause of immigration is known.	one factor between subject	OLS regression	Consequently, H1 receives empirical support through both tests.	r(58)=0.362	No	Effect calculated from Table 4, Model 1
Baekgaard & George, 2018		Politicians assess ethnic minority applicants less favorably than applicants from the ethnic majority.	2x2x2 between subject	OLS regression coefficient	Consistent with the hypothesis, there is a substantial negative effect of having an Arab name on being perceived as a representative of the public at large.	r(1683)=0.941	No	Effect calculated from Table 5, Model 1
Baekgaard & Kjaergaard, 2016		An increase in grants has a stronger stimulatory impact on public spending than an increase of equivalent size in citizens' income.	one factor between subject	regression	This hypothesis finds support in case the politicians are significantly more inclined to favour spending in Vignette 1 than in Vignette 2.	r(646)=-0.080	No	Effect calculated from Table 5, Model 1
Baekgaard & Serritzlew, 2016	1	when confronted with new information that is inconsistent with prior beliefs, actors will tend to selectively use heuristics to interpret the information in a way that is consistent with prior beliefs	2x2 between subject	logistic regression	It appears that subjects' prior beliefs have a negative impact on their ability to interpret performance information correctly if the private hospital is actually performing best (treatment 1, presented in model 1).	chi2(1)=28.43	No	Effect calculated from Table 3, Model 1
Baekgaard & Serritzlew, 2016	2	when confronted with new information that is inconsistent with prior beliefs, actors will tend to selectively use heuristics to interpret the information in a way that is consistent with prior beliefs	2x2 between subject	logistic regression	Specifically, prior beliefs have a significant negative impact on the ability to interpret performance correctly in the two treatments in which the private school performs best according to the performance information.	r(1405)=0.0604	No	Effect size calculated from Table 5
Baekgaard & Serritzlew, 2020	1	Performance information has less impact on opinion for individuals with higher numeracy	one factor between subject	Ordered logistic regression coefficient	The results of the survey experiment are consistent with the persuasion hypothesis: the opinions of highly numerate respondents are less affected by exposure to performance information.	r(415)=-0.1155	No	Effect calculated from Table 6, Model 1
Baekgaard et al., 2019		In line with these results, we expect that the framing of infor-	4x3x2 between subject	OLS regression	We find that respondents prefer more funding when performance information is framed negatively, possibly because negatively framed information	r(1398)=0.095	No	Effect calculated from Table 2, Model 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		mation matters to how the information is interpreted by political decision-makers.			generates a greater desire to spend more to avoid blame. The average preferred change in funding under the negative frame condition is 4.5 compared to 3.8 and 3.6 in the neutral and positive versions respectively. Thus, the difference between the positive and negative frames amounts to about one fourth of a standard deviation of the dependent variable. The finding that only negatively framed information has an impact is much in line with the literature on negativity bias.			
Baekgaard et al., 2020	1	We expect a relationship for partisan identity and assessments of health plans labeled as being part of the CA, with Democrats tending to be more favorable to the plan labeled as being part of the ACA and Republicans being more favorable to the performance of the plan labeled as being employer-provided (and thus not part of the ACA).	3x2 between subject	OLS regression coefficient	labeling of the plans as ACA/employer is associated with a significant reduction of about 12 percentage points in the rate of correct interpretations by Democrats, which is statistically significant (Table 1, Model 1).	r(584)=-0.118	No	Effect calculated from Table 1, Column 1
Baekgaard et al., 2020	2	We expect a relationship for partisan identity and assessments of health plans labeled as being part of the CA, with Democrats tending to be more favorable to the plan labeled as being part of the ACA and Republicans being more favorable to the performance of the plan labeled as being employer-provided (and thus not part of the ACA).	3x2 between subject	OLS regression coefficient	Beginning with Figure 5, among people who favor the public sector, the labeling of information as public/private hospital was associated with lower accuracy of interpretation, contrasted with the more neutral hospital A/B comparison. This main effect of labeling is statistically significant (Table 2, Model 1).	r(1453)=-0.096	No	Effect calculated from Table 2, Column 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Baekgaard, 2015	1	Information on high and low performance has a positive impact on citizen attitudes to service expansion while information on average performance has a negative impact on citizen attitudes.	2x3 between subject	Ordered logistic regression coefficient	The analysis shows that information on high performance has a positive impact on citizen support for more services, corresponding to a seven percentage point difference in the number of respondents who report that they "totally agree" or "agree" with the proposal about expanding services, while information on average and low performance has no impact. Hence, H1 is only partly supported.	r(1861)=0.057	No	Effect size calculated from Table 4, Model 1
Baekgaard, 2017	1	On average, citizens prefer certain to risky reform outcomes.	one factor between subject	t-test	With 58.9 per cent of the respondents preferring the certain to the risky reform, the experimental findings lend support to the hypothesis (t = 6.773; p < 0.01 in a two-tailed test).	t(1395)=6.773	No	
Baekgaard, 2017	2	On average, citizens prefer certain to risky reform outcomes.	one factor between subject	t-test	The results strongly support the risk aversion hypothesis since 64.9 per cent in experiment 2 (t = 12.774; p < 0.01; two-tailed t-test) and 68.1 per cent (t = 15.848; p < 0.01; two-tailed t-test) in experiment 3 prefer the certain to the risky option	t(1678)=12.774	No	
Baekgaard, 2017	3	On average, citizens prefer certain to risky reform outcomes.	one factor between subject	t-test	The results strongly support the risk aversion hypothesis since 64.9 per cent in experiment 2 (t = 12.774; p < 0.01; two-tailed t-test) and 68.1 per cent (t = 15.848; p < 0.01; two-tailed t-test) in experiment 3 prefer the certain to the risky option	t(1678)=15.848	No	
Barnes et al., 2018		Individuals will perceive female police officers as more successful at combating corruption when stereotypes about women's honesty are activated	between subject	t-test		t(516)=2.508	No	Effect calculated from Table 2

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Barrows et al., 2016		We therefore expect that providing relative performance information will influence citizens' evaluations of their local schools.	one factor between subject	difference-of proportions tests	Informing respondents of local school performance relative to the state, nation, or internationally depresses the probability of assigning local schools A or B by approximately 7, 10, and 11 percentage points, respectively.	r(2129)=0.07259	Yes	Effect size calculated from p-value in Table 1 Row 1
Belardinelli et al., 2018	1	Hypothesis 1: Framing effects will be stronger under ex post uses of performance information than under ex ante uses.	one factor between subject	t-test	Subjects in the negative framing condition gave a lower rating to the director of the sports facilities (M = 64.56, SD = 26.11) compared with respondents in the positive framing condition (M = 77.19, SD = 12.63), p < .001.	t(243.486)=-4.921	No	Effect calculated from Table 2, Row 1
Belardinelli et al., 2018	2	Hypothesis 1: Framing effects will be stronger under ex post uses of performance information than under ex ante uses.	one factor between subject	t-test	Subjects in the negative framing condition gave again a lower rating to the director of the sports facilities (M = 56.43, SD = 22.89) relative to respondents in the positive framing condition (M = 73.68, SD = 16.80), p < .001	t(51.597)=-3.457	No	Effect calculated from Table 3, Row 1
Belardinelli et al., 2018	3	Hypothesis 1: Framing effects will be stronger under ex post uses of performance information than under ex ante uses.	one factor between subject	t-test	In experiment Gamma, public managers in the negative framing condition gave a lower rating to the director of the sports facilities (M=64.00, SD=25.43) relative to subjects in the positive framing condition (M=75.70, SD=11.70), p<.05	t(80.341)=-2.418	No	Effect calculated from Table 4, Row 1
Belardinelli et al., 2018	4	Hypothesis 1: Framing effects will be stronger under ex post uses of performance information than under ex ante uses.	one factor between subject	t-test	The average rating of the sports facilities director was again lower in the negative framing condition (M=64.07, SD=24.27) relative to the average rating in the positive framing condition (M=75.83, SD=11.76), p<.01	t(67.274)=-2.662	No	Effect calculated from Table 7, Row 1
Belardinelli et al., 2018	5	Hypothesis 1: Framing effects will be stronger under ex post uses of performance information than under ex ante uses.	one factor between subject	t-test	When the customer (dis)satisfaction rate was negatively framed, participants gave a lower rating to the director of the sports facilities (M=60.11, SD=25.06) relative to their colleagues	t(616.106)=-4.535	No	Effect calculated from Table 10, Row 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
					in the positive framing condition (M=68.84, SD=22.83), p<.001			
Bellé & Cantarelli, 2018	1	Public employees attach a positive value to the opportunity to do good for others and the society through their job.	discrete choice experiment	Multilevel mixed effects linear model	Our data for Experiment 1 show that respondents preferred job positions that had an impact on many rather than few citizens, that were interesting rather than boring, and that required working under a leader with higher rather than lower ethical standards. Thus, these results provide support for hypotheses 1, 3, and 5, respectively.	Z=10.46	No	Effect calculated from text (p. 200)
Bellé & Cantarelli, 2018	2	Public employees attach a positive value to the opportunity to do good for others and the society through their job.	discrete choice experiment	Multilevel mixed effects linear model	More precisely, other things being equal, the odds that respondents would choose a job position rose by 5.0 times (p < .001) when the job provided the opportunity to serve many rather than few citizens	Z=9.47	No	Effect calculated from Table 2
Bellé & Cantarelli, 2019	1	Hypothesis 1: Leaders' ethical example reduces followers' unethical behavior.	between subject	Chi2	The percentage of participants who reported a 5 and, therefore, maximized their payoff was significantly higher than the expected 1/6 among students who did not listen to the ethical leadership speech (p = .012) but not among those who heard the ethical leadership messages (p = .104). This finding is in line with the prediction of Hypothesis 1.	t(118)=10.4734	No	Effect calculated from Table 2
Bellé et al., 2017	1	Whether anchoring effect exists in performance appraisal	one factor between subject	t-test	On average, raters' assessment of the ratee's performance was higher in the high anchor group (M = 88.47, SD = 15.63) than in the low anchor group (M = 71.07, SD = 13.01), p < .001. This is in line with our theoretical expectations about the presence of the anchoring effect in performance appraisal in the public sector.	t(589.198)=14.854	No	Effect calculated from text (p. 284)

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Bellé et al., 2017	2	Whether anchoring effect exists in performance appraisal	one factor between subject	t-test	On average, raters in the low accuracy group scored ratee's interpersonal skills lower (M = 60.30, SD = 22.40) than raters in the high accuracy group did (M = 67.74, SD = 19.46), p < .001	t(569.014)=-4.326	No	Effect calculated from text (p. 286)
Bellé et al., 2018	1	We expected participants in our experiments to behave in a systematically different manner depending on whether they were exposed to positively or negatively framed information.	factorial experiment	difference-of proportions tests	The proportion of risk-averse civil servants was higher under a positive frame (.82) than under a negative frame (.33), p < .001. Results of a logistic regression indicated that the odds of choosing the sure thing over the gamble were 9.24 times greater under a positive frame compared with a negative frame (p < .001).	chi2(1)=22.261	No	Effect calculated from Table 2, and text (p. 834)
Bellé et al., 2018	2	We expected participants in our experiments to behave in a systematically different manner depending on whether they were exposed to positively or negatively framed information.	factorial experiment	difference-of proportions tests	As expected, we observed risk-averse behaviors more frequently when civil servants were exposed to positively framed (.83) rather than negatively framed (.21) policy options, p < .001	chi2(1)=92.335	No	Effect calculated from Table 2, and text (p. 834)
Bellé et al., 2018	3	We expected participants in our experiments to behave in a systematically different manner depending on whether they were exposed to positively or negatively framed information.	factorial experiment	difference-of proportions tests	As predicted, more public employees preferred the program with the sure outcome over the program with the probabilistic outcome in the positive framing group (.84) compared with the negative framing group (.21), p < .001	chi2(1)=96.706	No	Effect calculated from Table 2, and text (p. 834)
Bellé et al., 2018	4	We expected participants in our experiments to behave in a systematically different manner depending on whether they were exposed to positively or negatively framed information.	factorial experiment	difference-of proportions tests	In experiment 4, civil servants who read that 80 percent of families and students were satisfied with the software (i.e., positive framing, N = 105) tended to report a higher propensity to purchase (81.06) compared with their peers who read that 20 percent of families and students were dissatisfied (i.e., negative framing, N = 99) (45.24), p < .001.	chi2(1)=26.74	No	Effect calculated from Table 2, and text (p. 834)

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Bellé et al., 2018	5	We expected participants in our experiments to behave in a systematically different manner depending on whether they were exposed to positively or negatively framed information.	factorial experiment	difference-of proportions tests	Respondents were more willing to buy the software when they were exposed to the satisfaction rate (i.e., positive framing, N = 189, 79.72) than when they were exposed to the same information expressed in terms of dissatisfaction rate (i.e., negative framing, N = 207, 60.43), p < .001.	chi2(1)=16.486	No	Effect calculated from Table 2, and text (p. 834)
Bellé et al., 2019	1	Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.	one factor between subject	Odds ratios	Compared to the control group, the odds of choosing the worse option increased by 5.21 times (p < .001) under a coercive pressure,	chi2(1)=37.732	No	Effect calculated from text (p. 384)
Bellé et al., 2019	2	Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.	one factor between subject	Odds ratios	Compared to the control group, the odds of choosing the worse option increased by 5.21 times (p < .001) under a coercive pressure,	Z=4.32	No	Effect calculated from Table 2
Bellé et al., 2019	3a	Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.	one factor between subject	Odds ratios	Compared to the control group, the odds of choosing the worse option increased by 5.21 times (p < .001) under a coercive pressure,	Z=2.99	No	Effect calculated from Table 2
Bellé et al., 2019	3b	Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.	one factor between subject	Odds ratios	Compared to the control group, the odds of choosing the worse option increased by 5.21 times (p < .001) under a coercive pressure,	Z=2.89	No	Effect calculated from Table 2
Bellé et al., 2019	4a	Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will	one factor between subject	Odds ratios	Compared to the control group, the odds of choosing the worse option increased by 5.21 times (p < .001) under a coercive pressure,	Z=2.36	No	Effect calculated from Table 2

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		choose A, even if A is inferior to B.						
Bellé et al., 2019	4b	Given options A and B, a coercive isomorphic pressure towards A increases the probability that public employees will choose A, even if A is inferior to B.	one factor between subject	Odds ratios	Compared to the control group, the odds of choosing the worse option increased by 5.21 times (p < .001) under a coercive pressure,	Z=2.22	No	Effect calculated from Table 2
Bellé, 2013		Direct contact with benefi ciaries has a positive eff ect on job performance.	between subject	t-test	Both group B (contact with benefi ciaries) and group C (self-persuasion) signifi cantly outperformed group A (control) in each of the four performance areas. Th ese results provide support to hypotheses 1a and 1b.	t(36.191)=2.997	No	Effect calculated from Table 1a
Bellé, 2015		The visibility of rewards will moderate the effect of monetary rewards on job performance such that monetary rewards will have a stronger eff ect when they are secret and a weaker eff ect when they are visible.	3x2x2 between subject	OLS regression coefficient	Th e signifi cant interaction of PRP and visibility in tables 2a and 2b indicates that, averaging across the two levels of benefi ciary contact (yes/no), monetary rewards were less eff ective among participants whose performance and pay were observable by their colleagues compared with nurses in the secret condition, F(1,199) = 31.72, p < .001. Th is fi nding supports hypothesis 1.	r(191)=-0.190	No	Effect calculated from Table 2b
Bergner et al., 2019		H1: Activists' willingness to act on behalf of a policy will increase with a general reference to science, as compared to no reference to science.	field experiment	Chi-square test	The experimental manipulation did significantly affect click-through rates in both groups, as indicated by statistically significant chi-square test.	chi2(2)=14.551	No	Effect calculated from Table 2 and text
Blom-Hansen et al., 2016		the preferences of politicians for a policy proposal depend on what information is presented to them.	one factor between subject	Ordered logistic regression coefficient	Model 1 reproduces the test of the effect of providing cost information on preferences. As in the t-test, we find strong and statistically significant differences between the treatment and the baseline group.	r(1202)=-0.198	No	Effect size calculated Table 2, Model 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Borry et al., 2018	1	Written rules will be followed at higher rates than unwritten rules.	one factor between subject	difference-of proportions tests	Proportional difference tests between the written and unwritten scenarios are statistically significant (z = 2.78, p < .01)	Z=2.78	No	
Borry et al., 2018	2	Written rules will be followed at higher rates than unwritten rules.	one factor between subject	difference-of proportions tests	Proportional difference tests between the written and unwritten scenarios are highly statistically significant (z = 5.20, p < .01)	Z=5.20	No	
Bouwman et al., 2018		Accountable negotiators will show lower performance at the group level than nonaccountable negotiators.	between subject	t-test	The first hypothesis is supported by our data. Triads of players in the control condition obtained higher group scores (Mcontrol = 464.44, SD = 27.43) than triads in the treatment condition (Mtreatment = 374.75, SD = 61.44).	t(85.000)=9.413	No	Effect calculated from text (p. 42)
Brewer & Brewer, 2011		individuals respond differently when they believe they are working for a government agency rather than a business firm.	one factor between subject	t-test	Together, these results indicate that when participants were informed that the research they were taking part in was funded either by a public or private sector sponsor—either a government agency or a business firm—their performance changed in systematic and reliable ways.	t(38)=2.22	No	Effect calculated from text (p. 357)
Bro & Jensen, 2020		Transformational leadership positively affects user orientation	field experiment	OLS regression coefficient	Consistent with our expectation, model 1 reveals a positive and statistically significant regression coefficient estimate for user orientation three months after the training programme (β = 2.66, p < .05).	r(583)=0.102	No	Effect calculated from Model 1, Column 1
Cantarelli, Bellé, & Belardinelli, 2020	1	The asymmetric dominance effect, also known as decoy or attraction, causes individuals' preferences between a target option and a competing option to shift toward the target when a decoy, similar to the target but	one factor between subject	Logistic regression	A logistic regression showed that the odds of choosing the target option (i.e., the diagnostic instrument providing both paper and electronic records and costing €9,000) were 3.42 times greater among participants presented with the decoy option (i.e., the diagnostic instru-	r(496)=0.1484	Yes	Effect size calculated from p-value in text (p. 61)

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		in no way better, is added to the choice set			ment providing paper records and costing €9,000) as compared to participants who were not presented with the decoy option (p < .001).			
Cantarelli, Bellé, & Longo, 2020	1	Nurses prefer jobs that require fewer overtime hours	discrete choice experiment	Mixed effects regression coefficient	Keeping everything else constant, professionals in the pooled sample preferred jobs that required 2 rather than 8 overtime hours per week	r(3341)=-0.165	No	Effect calculated from Table 1, Row 1
Charbonneau & Van Ryzin, 2015		we would expect that citizens would give the lowest evaluations when their school's reported performance is lower than for similar schools or the state average (comparative benchmarks)	one factor between subject	Chi2	As a whole, the differences in the distribution of responses across experimental branches are statistically significant at the 10 per cent level for evaluations (Table 1) and at the 5 per cent level for verdicts (Table 2).	chi2(15)=26.10	No	Effect calculated from text (p. 295)
Christensen et al., 2018		When facing contradictory performance data, elected officials will use goal reprioritization as a strategy to make their use of performance information fit with governance preferences for public or private service provision.	one factor between subject	Logistic regression	The significant coefficients for the interaction term "pro-public preference × sector visible" provide evidence that is consistent with our hypothesis: respondents' do engage in goal reprioritization:	r(482)=-0.362	No	Effect calculated from Table 1, Model 1
Christensen, 2018	1-2015	Service users find performance indicators more important when their chosen service provider performs well on the indicators (compared to competing providers) and less important when their chosen provider performs less well.	one factor between subject	t-test	Effects are statistically significant in one test (Experiment 1 in 2015 where the effect of Aarhus University performing best is a +0.85 increase in perceived importance, p = .015)	r(200)=0.171	Yes	Effect size calculated from p-value in text (p. 473)
De Fine Licht et al., 2014		transparency will have zero or even negative effects in policy areas that involve trade-off s in which human life and well-being	one factor between subject	ANOVA	transparency has a positive eff ect on perceived legitimacy in the policy area that handles routine trade-off s (culture and leisure) and a tendency for a nega-	F(2,505)=29.503	No	Effect calculated from text (p. 367)

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		are at least implicitly weighed against money			tive eff ect in the policy area that handles trade-off s with elements of taboo (traffi c security).			
de Fine Licht, 2014		There is good reason to believe that transparency can generate legitimacy.	one factor between subject	OLS regression coefficient	it is clear that transparency can have the positive effect on legitimacy, which motivates many political and administrative transparency reforms.	r(159)=0.155	Yes	Effect size calculated from p-value in Figure 1
Demaj, 2017		Performance information helps in making a more pronounced budget request and therefore causes legislators to choose more extreme deviations from the current level of funding.	one factor between subject	ANOVA	Overall, the results presented in table 3 support the notion suggested by hypotheses 1a and 1b.	F(3,119)=30.950	No	
Deslatte, 2019		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).	2x3 between subject design	Anova	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) =3.62, p < .1.	F(1,332)=10.26	No	
Doberstein & Charbonneau, 2020		Does knowing the number of disclosure investigations in jurisdiction affect public servant confidence in the WB regime?		t-test	What we see with this comparative experimental data is that showing respondents data on reported and confirmed wrongdoing is associated with a larger clustering on data in higher levels of confidence in the whistleblowing regime.	t(668)=-2.19	No	Effect calculated from text (p. 651)
Filtenborg et al., 2017	1	Replication: Van Ryzin (2013)	2x2x3 factorial between subject	t-test	All the core relationships in the expectancy disconfirmation model follow the expected direction since expectations are negatively related to satisfaction	r(229)=0.1719	Yes	Effect size calculated from p-value in Table 3 Row 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
					contrary to performance that is positively related to disconfirmation.			
Filtenborg et al., 2017	2	Replication: Van Ryzin (2013)	2x2x3 factorial between subject	t-test	All the core relationships in the expectancy disconfirmation model follow the expected direction since expectations are negatively related to satisfaction contrary to performance that is positively related to disconfirmation.	r(229)=0.1719	Yes	Effect size calculated from p-value in Table 3 Row 1
George et al., 2017		Performance information showing low performance has a positive impact on politicians' preferences for public spending in the same policy area.	one factor between subject	Random-effects regression coefficient	In accordance with our expectations (H1), we find that these respondents are more inclined to support higher spending if they have received the performance information treatment.	r(391)=0.191	No	Effect calculated from Table 5, Column 2
Geys & Sørensen, 2018		Public sector performance affects politicians ' reform preferences, particularly when performance information is explicitly provided.	one factor between subject	odds ratios	we only observe statistically significant effects for three out of eight reform options, such that is it important to account for heterogeneity across alternative types of (school) reforms.	r(3477)=0.0351	No	Effect size calculated from Table 2, Model 1, Panel 2 (plus author response)
Grimmelikhuijsen & Klijn, 2015		Judicial transparency is expected to have an overall positive effect on trust in judges.	one factor between subject	OLS regression coefficient	Looking at the regression analysis on the overall population, we found a main effect of the television series on trust in judges (F=22.54, β =0.254, p<0.001, R2 =0.041).	F(2,1047)=22.54	No	
Grimmelikhuijsen & Porumbescu, 2017	1	Replication: Van Ryzin (2013)	2x2 between subject	ANOVA	A subsequent two-way factorial ANOVA, using expectations and performance as separate factors, indeed showed that performance had a strong and significant effect on citizen satisfaction (F(1,305) = 39.93, p = .000, eta-squared = .116).	F(1,305)=39.93	No	
Grimmelikhuijsen & Porumbescu, 2017	2	Replication: Van Ryzin (2013)	2x2 between subject	ANOVA	A subsequent two-way factorial ANOVA, using expectations and performance as separate factors, showed that performance had a highly significant effect on citizen satisfaction	F(1,169)=362.19	No	

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
					(F(1,169) = 362.19, p = .000, eta-squared = .682)			
Grimmelikhuijsen & Porumbescu, 2017	3	Replication: Van Ryzin (2013)	2x2 between subject	ANOVA	The two-way factorial ANOVA showed that performance nevertheless did have a significant effect on citizen satisfaction (F(1,163) = 32.54, p = .000, eta-squared = .166).	F(1,163)=32.54	No	
Grimmelikhuijsen et al., 2013		Transparency will have a stronger negative/less positive effect on trust in government in national cultures with high power distance.	one factor between subject	t-test	people who were shown complete decision-making information were much more negative about the perceived competence of the government organization (–0.93), which is indicated by the superscripts a and b.	t(41.143)=26.967	No	Effect calculated from Table 4
Grimmelikhuijsen et al., 2018		H1: A press release does not negatively affect the perceived trustworthiness of an independent regulatory agency.	one factor between subject?	ANOVA	Perceived competence is most strongly af-fected by the treatment (F(4,554)=5.09, p < .001, partial eta2=.035)	F(4,554)=5.09	No	
Grimmelikhuijsen et al., 2019		H1: Local authorities will be more responsive to a FOI request than to an informal ask.	field experiment	OLS regression	We find that 89.2 percent of the treated municipalities provide a response, whereas only 64.6 percent of the municipalities in the control group responded; when cross-tabulated against all other responses, this is a significant difference (Chi2=32.99, p<.001). [...] Overall, we find support for the first hypothesis: a FOI request leads to overall higher response rates.	chi2(1)=32.99	No	
Guardino & Mettler, 2020		Such factors suggest that providing citizens with basic information about tax expenditures could influence their opinions, despite the political predispositions and psychological biases that can block or mitigate information effects.	mixed subject design		The basic information treatment caused subjects to express a higher level of favorability toward the home mortgage interest deduction (Figure 1) and retirement savings contribution tax credit (Figure 2):	t(332.342)=-2.349	No	Effect provided by authors

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Hamidullah et al., 2020		we predict that citizens will trust that the government is doing the right thing, and perceive it as fairer, if it makes an effort to narrow the pay gap between women and men.	2x2 between subject	regression coefficient	As Figure 2 shows, respondents in both the relatively low and high women's performance conditions feel like they can trust the ad-hoc committee more when the pay gap between women and men is reduced.	r(1025)=0.250	No	Effect calculated from text (p. 11)
Hattke et al., 2020		Hypothesis 1a: Delay, burden, and the two combined will evoke negative emotional responses	between subject	One-Way Variance Analysis (ANOVA)	The significantly higher mean values for anger, confusion, and frustration caused by the treatments indicate that this is indeed the case, confirming this hypothesis.	t(28.019)=-2.961	No	Effect calculated from text (p. 59)
Hattke & Kalucza, 2019		H1a. Intrinsic motivation positively influences citizens' willingness to coproduce.	between subject	Multilevel mixed effects linear model	Results only partially confirm Hypotheses 1a, 1b, and 1d as only context-specific intrinsic motivation, prosocial motivation, and self-efficacy positively influence the willingness to coproduce.	r(194)=0.192	No	Effect calculated from Table 3
Herian et al., 2012		Information about the use of deliberative public participation procedures will increase perceptions of fairness about the budgeting process.	one factor between subject	ANOVA	the results provide support for hypothesis 1 and suggest that information about government's use of public input can increase perceptions of fairness among the public, particularly when that information emphasizes the deliberative aspect of the public participation techniques employed by the city	F(1,200)=4.94	No	Effect calculated from text (p. 823)
Hjortskov, 2017	1	priming, both in the form of prior information and prior items in a survey, may affect attitude accessibility and therefore influence answers to later questions in a citizen survey.	2x3 between subject	OLS regression coefficient	Two different performance questions were asked with direct relevance for the crime and police treatments: one about police protection and one about police–community relations. Both are strongly affected by the positively framed crime and police questions (1A and 2A).	r(1601)=0.068	No	Effect calculated from Table 3, Model 1
Hjortskov, 2017	2	priming, both in the form of prior information and prior items in a survey, may affect attitude accessibility and therefore	2x3 between subject	OLS regression coefficient	The positive police questions have a positive effect on the police protection,	r(1183)=0.085	No	Effect calculated from Table 7, Model 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		influence answers to later questions in a citizen survey.			police–community, expectations, perceived performance and satisfaction questions.			
Hollibaugh et al., 2020	1	The hypothesis in this study is, therefore, that citizens will also interpret expectations differently.	3x2 between subject	OLS regression	(proportion of responses to most frequent category was used to conduct a one-sample t-test against 100%)	t(1589)=49.467	No	Effect calculated from Figure 1
Holm, 2017		Hypothesis 1: The more a proposed policy is viewed as congruent with public servants' code of ethics, the less likely they are to engage in guerrilla government activity.	Conjoint experiment	OLS regression coefficient	Consistent with some of the literature, the strongest predictors of employing guerrilla government tactics are the extent to which the policy violates an individual's code of ethics and whether the person thinks that the proposal is a good idea	r(6405)=0.037	No	Effect calculated from Table A4, Column 1
Houlberg et al., 2016		Public managers evaluate performance results differently when presented with an aspiration level that is either above or below the absolute result.	one factor between subject	OLS regression coefficient	The first notable result in the figure is that unambiguous negative and positive feedback changes principals' performance evaluations. This result is in line with the expectation in H1, namely that aspiration level changes public managers' evaluations of identical performance level.	r(623)=0.268	No	Effect calculated from Table A1, Model 1.1
Hvidman & Andersen, 2016		In a spending scenario, tractability is the dominant dimension: policy areas with low tractability will attract more spending, while this is less the case for areas with high deservingness combined with high tractability.	2x2x5 factorial between design	t-test	Figure 3 shows that these three policy areas fall in a group where the preferences for spending are among the highest and preferences for cutting relatively low, but the results also show that cutting and spending are not perfectly correlated.	t(778.854)=-14.771	No	Effect calculated from Table 1
Ingrams et al., 2020		H1: Transparency in a public decision-making process increases citizen trust, satisfaction, and perception of fairness.	2x2 between subject	ANOVA	The same is true for trust (F(1,281) = 9.42, p < .01)	F(1,281)=9.42	No	

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Jakobsen & Andersen, 2013		whether improving both efficiency and equity in service outcomes is possible through a coproduction program targeted at increasing service user participation in coproduction.	one factor between subject	OLS regression coefficient	In model II, the program has a statistically signifi cant eff ect of 0.86 (p < .01) for children whose mothers have no registered education.	r(277)=0.155	No	Effect calculated from Table 3, Model 2
Jakobsen et al., 2019		Training can be seen as a management tool that can be used to direct employee behavior	between subject	OLS regression coefficient	In this model, the overall effect of the training treatment is 0.39, which is statistically significant at the 0.01- level.	r(2675)=0.063	No	Effect calculated from Table 2, Column 1
James & Moseley, 2014		Information about low absolute performance will lower perceptions of absolute performance	2x2 between subject	OLS regression coefficient	Support was found for Hypothesis 1a, information about absolute low performance lowered perceptions of the percentage of waste recycled.	r(287)=0.188	No	Effect calculated from Table 2
James & Petersen, 2018	2	H1:The international source (compared to the national government source) increases citizens' belief in the performance measure's truthfulness;	one factor between subject	OLS regression coefficient	The exact/direct replication in the UK similarly revealed a rise in belief in truthfulness by 0.12 points, with the null clearly rejected at the 0.05 level, p = 0.01, 95 per cent CI [0.03 0.20] using the same scale (UK min = 1, max = 6, SD = 0.84).	r(1519)=0.077	No	Effect calculated from Table 2, Model 2
James & Petersen, 2018	3	H1:The international source (compared to the national government source) increases citizens' belief in the performance measure's truthfulness;	one factor between subject	OLS regression coefficient	In the Netherlands, the equivalent figures were an increase of 0.14 points, p = 0.00, 95 per cent CI [0.05 0.24] on the same scale (NL min = 1 max = 5, SD = 0.76).	r(997)=0.088	No	Effect calculated from Table 2, Model 3
James & Van Ryzin, 2017		First, we expect to observe partisan differences in initial beliefs about the Affordable Care Act; and, moreover, we expect such differences to be accentuated by the political prime, in contrast to the health care needs prime.	one factor between subject	OLS regression coefficient	compared to the health care prime, the politics prime appears to widen partisan differences in the evaluation of evidence statements that report good results for the Affordable Care Act (pro statements).	r(302)=0.124	No	Effect calculated from Table A3, Model 1
James & Van Ryzin, 2019	1	H: reporting violent crimes and infant mortality using a larger denominator will lead citizens to perceive more risk and worse	between subject	ANOVA	Figure 1 shows the descriptive results for perceived risk of violent crime across treatment groups in the FBI experiment. A clear pattern can be seen	F(5,825)=4.8	No	

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		conditions in society, compared to reporting the mathematically same statistic using a smaller denominator.			of lower perceived risk when respondents viewed smaller-number ratios (one-way ANOVA F = 4.80, p < .001, see Table C1 in the Supplement).			
James & Van Ryzin, 2019	2	H: reporting violent crimes and infant mortality using a larger denominator will lead citizens to perceive more risk and worse conditions in society, compared to reporting the mathematically same statistic using a smaller denominator.	between subject	ANOVA	The pattern is not as clear (in terms of decreasing perception of risk with smaller number ratios) as it was for the FBI statistics, but still there is a significant overall difference in risk perceptions across groups (ANOVA F = 5.6, p < .001, see Table C3 in the Supplement).	F(5,826)=5.56	No	
James, 2011	1	providing information about relatively good performance will raise citizens'assessments of performance	one factor between subject	t-test	The hypothesis that information about relatively good performance raises citizens' assessments of performance (H1 a) is supported.	t(349)=4.372	No	Effect calculated from Table 2
James, 2011	2	providing information about relatively good performance will raise citizens'assessments of performance	one factor between subject	t-test	The hypothesis that information about relatively good performance raises citizens' assessments of performance (H1 a) is supported.	t(98)=5.599	No	Effect calculated from Table 2
Jensen et al., 2019		Hypothesis 1a: Transformational leadership increases public employees' PSM.	field experiment	OLS regression coefficient	Contrary to the expectation of hypothesis 1a, PSM on average declines among employees in organizations in which the managers were assigned to the transformational leadership group (compared with employees in organizations in which the managers were assigned to the control group).	r(3459)=-0.041	No	Effect calculated from Table 2, Model 1
Jilke & Tummers, 2018		Street-level bureaucrats will more likely intend to help clients who show a high effort (i.e., high earned deservingness), than those who show a low effort (i.e., low earned deservingness).	Conjoint experiment	OLS regression coefficient	Hypothesis 1 states that street-level bureaucrats will more likely intend to help clients who exhibit a high effort. Our results support this expectation; teachers are 6% more likely (p < .05) intending to prioritize hardworking students, when compared to students that are described as "somewhat lazy."	r(2079)=0.050	No	Effect calculated from Table A1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Jilke et al., 2016	1	All things being equal, citizens who experience severe dissatisfaction with a given service will be less likely to switch away from their current service provider when faced with many alternative providers, compared to people who are equally dissatisfied but have a more limited set of providers to choose from.	2x2 between subject	Logistic regression	Moreover, the interaction term between both treatments turns statistically significant and exhibits a negative effect direction. This means that those respondents in the severe dissatisfaction condition who were given many choices were less likely to abandon their default provider. This lends support to the choice-overload hypothesis as outlined in the previous section.	r(1150)=-0.0903	No	Effect calculated from Table 1, Model 2
Jilke et al., 2016	2	All things being equal, citizens who experience severe dissatisfaction with a given service will be less likely to switch away from their current service provider when faced with many alternative providers, compared to people who are equally dissatisfied but have a more limited set of providers to choose from.	2x2 between subject	Logistic regression	Examining the coefficient for the included term, we find that the combined effects of both treatments has a statistically significant effect on subjects' stated choice decision.	r(541)=-0.0853	No	Effect calculated from Table 2, Model 2
John & Blume, 2018	1	simplification: The conclusion to draw is that simplification designs are easy wins for local delivery organizations and provide positive effects that are transferable across many domains.	randomized controlled trial	Chi-square test	The table shows that the simplification raises payment by 3.8 percentage points for the simplification only group.	chi2(1)=6.7314	No	Effect calculated from Table 2
John & Blume, 2018	2	simplification: The conclusion to draw is that simplification designs are easy wins for local delivery organizations and provide positive effects that are transferable across many domains.	randomized controlled trial	Chi-square test	It shows that 41.40% in the treatment group paid in full, whereas 43.57% did so in the control, which indicates the social norm backfired with less people paying in full in the treatment group. This difference is statistically significant at p < .001	chi2(1)=25.312	No	Effect calculated from Table 5
Jurcevic & Fyall, 2019	1	Hypothesis 1: Within the non-profit context, Whites find	2x2 between subject	ANOVA	Consistent with our proposal that an instrumental diversity frame applied to a nonprofit context may deter Whites, White participants felt more dehumanized than Black and Latino participants	F(1,340)=4.62	No	

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		moral diversity frames more favorable than instrumental diversity frames.			by the nonprofit with an instrumental frame, F(1, 340)=4.62, p=.032, ηp2=.013, and marginally more dehumanized following the instrumental, compared to moral, frame, F(1, 340)=3.40, p=.066, ηp2=.010.			
Karens et al., 2016	1	using the EU brand to brand policies will positively affect people's trust in those policies.	one factor between subject	ANOVA	The experimental results in Belgium show a signifi cant difference between the control group and the treatment group.	t(63.844)=-3.451	No	Effect calculated from Table 5
Karens et al., 2016	2	using the EU brand to brand policies will positively affect people's trust in those policies.	one factor between subject	ANOVA	The MANOVA test conducted with data from the experiment in Poland yields a Pillai's trace F value of 9.715 (p = .000)	t(80.092)=-3.882	No	Effect calculated from Table 6
Karens et al., 2016	3	using the EU brand to brand policies will positively affect people's trust in those policies.	one factor between subject	ANOVA	The MANOVA test for the experiment in the Netherlands yields a Pillai's trace F value of 3.373 (p = .042)	t(55.706)=-2.412	No	Effect calculated from Table 7
Kaufmann & Feeney, 2014		Red tape perceptions will be positively related to procedural length.	one factor between subject	OLS regression coefficient	This means that the three vignettes with increasing levels of red tape (Versions B, C, and D) are positively related to experimental red tape perceptions when compared to the lowest red tape vignette (Version A).	r(70)=0.460	No	Effect calculated from Table 4, Model 2
Kaufmann & Tummers, 2017		H1: Red tape has a negative effect on procedural satisfaction	one factor between subject	ANOVA	By contrast, the treatment group (high red tape) rated a far lower level of satisfaction, averaging 2.47 (SD = 1.06). This difference is highly significant (F(1,137) = 39.751, p < .001, partial η2 = .225).	F(1,137)=39.751	No	
Keiser & Miller, 2020		Hypothesis 1: Information about a high level of administrative burden in the application process leads to greater support for TANF and its recipients	one factor between subject	difference-of proportions	Generally, the results provide evidence in support of the hypotheses for high burden; respondents in the high burden condition are generally more supportive of the TANF program and its recipients.	chi2(1)=4.429	No	Effect size calculated from Figure 1, Panel 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		compared with no information about the application process.						
M.-H. Kim et al., 2019	1	Citizens exposed to information detailing high police performance in fighting crime will evaluate police trustworthiness more positively than those exposed to information detailing low police performance.	2x2 between subject	ANOVA	For table 2, we see the coefficients for performance are positive, meaning that among white respondents, those assigned to the high performance treatment (X=1) perceived police as more competent, benevolent, and sincere than those assigned to the low performance treatment (X=0).	t(163.586)=11.644	No	Effect calculated from Table 4, Row 1
M.-H. Kim et al., 2019	2	Citizens exposed to information detailing high police performance in fighting crime will evaluate police trustworthiness more positively than those exposed to information detailing low police performance.	2x2 between subject	ANOVA	For table 2, we see the coefficients for performance are positive, meaning that among white respondents, those assigned to the high performance treatment (X=1) perceived police as more competent, benevolent, and sincere than those assigned to the low performance treatment (X=0).	t(210.644)=13.814	No	Effect calculated from Table 4, Row 5
S. H. Kim & Kim, 2016		Respondents are prone to giving more socially desirable answers instead of expressing their true feelings.	list experiment	t-test	On the item of self-sacrifice (SS), the proportion of affirmative answers to the overt question is 44% but the true support for that item is only 26%, and therefore the magnitude of SDB in the entire sample is 18%. This result indicates that a significant percentage of people were hiding their true feelings on this item of SS.	r(7199)=0.03912	Yes	Effect size calculated from p-value in Table 1
S. H. Kim & Kim, 2017		Hypothesis 1: Respondents in general are prone to giving more socially desirable answers instead of expressing their true feelings.	list experiment	OLS regression coefficient	Although we can find some differences in the distribution of the subgroup SDB, it is apparent that all three measures were heavily influenced by SDB. Thus, this finding supports Hypothesis 1 that respondents in an individualistic society are more likely to give positive answers instead of expressing their true feelings when responding to sensitive items.	r(3999)=0.0524	Yes	Effect size calculated from p-value in Table 2

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Larkin et al., 2019		Given that a local authority is a more proximal and familiar institution to residents compared to national institutions, we predict that a social norm intervention is more effective at the subnational level.	field experiment	OLS regression	The social norm and enforcement salience letters significantly increased a household's propensity to pay their Council Tax, compared to the control letter.	$r(9090)=0.047$	No	Effect calculated from Table 2, Model 1
Lee et al., 2017	1	A high-uncertainty (U+) or high-risk (R+) setting increases the number of rounds required to achieve the Nash equilibrium than a lo uncertainty andlow-risk (U−R−) setting	2x2 between subject	OLS regression coefficient	As shown in Table 2, high uncertainty and equilibrium convergence have strong effects on higher rewards.	$t(70)=-2.99$	No	Effect calculated from Table 2, Model 1
Lee et al., 2017	2	A high-uncertainty (U+) or high-risk (R+) setting increases the number of rounds required to achieve the Nash equilibrium than a lo uncertainty andlow-risk (U−R−) setting	2x2 between subject	OLS regression coefficient	As shown in Table 2, high uncertainty and equilibrium convergence have strong effects on higher rewards.	$t(91)=-4.03$	No	Effect calculated from Table 2, Model 2
Levitats et al., 2019		Hypothesis 1a:The emotional intelligence of civil servants positively affects their social responsibility.	one factor between subject	t-test	A t-test found that this difference was significantly higher on average for those in the experimental group compared with the control group (t = 2.25, df = 50.53, p = .03).	$t(50.53)=2.25$	No	Effect calculated from text (p. 845)
Linos & Riesch, 2020		Specifically, we consider three different types of costs associated with administrative burden: learning costs, compliance costs, and psychological costs. While we, too, hypothesize that reductions in these types of costs should improve compliance, it is not a priori obvious that a reduction in overall administrative burden correlates with a reduction in these costs	one factor between subject	OLS regression coefficient	Our results indicate a positive effect of treatment on PHS submission, online submission, and submission within two weeks of receiving the treatment/control email.	$r(2581)=0.073$	No	Effect calculated from Table 5, Column 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Linos et al., 2017		We hypothesized that since applicants from underrepresented groups should be more likely to feel uncertain about their levels of belonging, they should also be more responsive to cues of belonging than would be their white counterparts.	field experiment	OLS regression coefficient	Non-white applicants in the treatment group gained 12 percentage points in their percentile ranking.	$r(1588)=-0.055$	No	Effect calculated from Table 3, Model 2
Livnat-Lerer et al., 2018	2	(H1) OK-based PA relationship increases agents' risk aversion	between subject	logit regression	Anticipated outcome knowledge had a negative effect on the propensity to choose the risky option (no-OK-based PA: P(risk)=.745 [.630, .860], OK-based PA: P(risk)=.447 [.317, .577], p=.002).	$chi2(1)=9.034$	No	Effect calculated from text
Livnat-Lerer et al., 2018	3	(H1) OK-based PA relationship increases agents' risk aversion	between subject	logit regression	In line with hypothesis 1 and the results of Studies 1 and 2, there was a significant difference between the percentage of risk taking under the non-OKbased PA and the OK-based PA conditions (Chi2= 5.49, p = .019): 68.9% and 44.9%, respectively.	$chi2(1)=5.49$	No	
Lu et al., 2019		Hypothesis 1: Chengguan officers tend to be more lenient toward clients with high help deservingness.	2x2 between subject	OLS regression coefficient	the main effect on Chengguan officers' discretion, the effect of clients' help deservingness on the fines imposed by the officers is negative and statistically significant (p<0.01), which is consistent with our expectation (Hypothesis 1).	$r(418)=-0.411$	No	Effect calculated from Table 4, Column 1
Marvel & Girth, 2016		Citizens will perceive political principals to have more control when the provider of a service is a public agency (versus a private contractor).	2x3 between subject	regression	Consistent with this hypothesis, the topmost coeffi - cient in table 2's fi rst column indicates that perceived mayoral control over provider is, on average, about 1.38 units higher among subjects who read a Department of Public Works (DPW) vignette than among subjects who read a Waste Collection Incorporated (WCI) vignette.	$r(685)=0.402$	No	Effect calculated from Table 2, Column 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Marvel, 2015		Individuals' assessments of public sector performance will exhibit expectation bias.	2x3 between subject	t-test	Subjects' performance guesses reveal that performance expectations are more modest for the Postal Service than for FedEx.	t(506.800)=-4.516	No	Effect calculated from Table S2 and text (p. 218)
Mazepus & van Leeuwen, 2020		a positive personal outcome increases the perceived legitimacy of political authorities (Hypothesis 1).	2x2x2x2 between subject	ANOVA	In support of Hypothesis 1, estimated marginal means showed that the presence of a positive outcome (M = 4.08, SE = 0.03) increased legitimacy compared to the absence of a positive outcome (M = 3.63, SE = 0.03), F(1, 2597) = 89.26, p < .001, partial $\eta2$ = .033.	F(1,2597)=89.26	No	
Meyer-Sahling et al., 2019		Hypothesis 1 (H1): Activating PSM will make respondents more willing to report ethical problems to management.	one factor between subject	OLS regression coefficient	The figure shows estimates from an ordinary least squares (OLS) model regressing our PSM battery treatment on willingness to report	r(4760)=0.078	No	Effect calculated from Table A4.1, Model 1
Migchelbrink & Van de Walle, 2020		Hypothesis 1: The higher the input legitimacy of a participatory process, the more willing public officials are to use its citizen inputs in administrative decision-making.	2x2 within subject	ATEs	Table 1 shows that turnout and participants' representativeness influence officials' willingness to use citizens' inputs in administrative decision-making.	r(1419)=0.330	No	Effect calculated from Table 1, Column 1
Moseley & Stoker, 2015		a default whereby people are automatically assumed to be donors but allowed to opt-out (the 'opt-out condition'), would lead to greater numbers of website visits and registrations than a default where people were not assumed to be donors but allowed to opt in (the 'opt-in condition').	2x3 between subject	difference-of proportions tests	As hypothesised, we found that both opt-out and mandated choice questions generated significantly more visits to the national organ donation website than the opt-in group	Z=5.2	No	Effect calculated from text (p. 255)
Muñoz et al., 2016		support for politicians suspected of corruption increases when they have a good record (e.g., attracting investments, presiding	one factor between subject	t-test	The credibility treatment and the implicit exchange treatment generated a large and statistically significant increase in the reported probability of	r(374)=0.1708	Yes	Effect size calculated from p-value in Table 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		over economic growth and securing well-being for their constituency).			voting for the allegedly corrupt mayor in the next election.			
Nelson & Witko, 2020		H1: Individuals will view jobs more favorably if they are located in states where the governor has been more aggressive in mandating social distancing.	conjoint experiment	AMCE	We find clear support for Hypothesis 1: individuals evaluate jobs more favorably if the governor took more aggressive action regarding social distancing.	Z=2.818	No	Effect calculated from published Code and Data
Ngoye et al., 2019		we hypothesize that specific texts, as environmental cues, can nonconsciously prime institutional logics.	one factor between subject	ANOVA	Examination of the results displayed in table 2 reveals statistically significant differences at p < .05 between group 1 (primed for public logic) and group 2 (primed for market-managerial logic) and between group 1 and group 3 (primed for professional logic) in how they ranked recommendations aligned with the public and market-managerial and the public and professional logics, respectively.	r(48)=0.415	Yes	Effect size calculated from p-value in Table 2 Column 1
Nicholson-Crotty et al., 2019	1	We suggest that the prosocial motivations associated with PSM may influence risk tolerance differently when the payoffs for risk taking are framed as benefits for others, rather than the individual.		t-test	However, the mean level of PSM was significantly higher in the group that chose the risky alternative when outcomes were framed negatively, or as a loss of life (.10 vs. -.18, p<.05)	r(149)=0.164	No	Effect calculated from published Code and Data
Nielsen & Baekgaard, 2015		Performance information showing low performance has a positive impact on politicians' attitudes to spending.	one factor between subject	Ordered logistic regression coefficient	Information showing low performance has a positive effect on spending attitudes. The finding thus conforms to the blame avoidance perspective expressed in Hypothesis 1 according to which it is not politically feasible to punish low performers in a popular and highly salient policy area.	r(839)=0.107	No	Effect calculated from Table 4, Model 2

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Nielsen & Jacobsen, 2018	1	Performance information showing high performance increases employee acceptance of management authority.		OLS regression coefficient	For the high-performance group, we find a statistically significant positive effect on management acceptance of receiving performance information, which provides support for hypothesis 1.	$r(910)=0.989$	No	Effect calculated from text (p. 690)
Nielsen & Moynihan, 2017		When provided with performance data, elected officials' attribution of responsibility to public sector leaders will be most pronounced in cases of low performance scores.	2x3 between subject	OLS regression coefficient	consistent with hypothesis 1, table 2 also shows that the relationship between the provision of performance data and attribution of leadership control is concentrated among local elected officials whose schools are in the lowest category of performance.	$r(227)=0.130$	No	Effect calculated from Table 2, Model 2
Olsen, 2015		exposure to a satisfaction rate should induce more positive evaluations of public services, whereas exposure to a dissatisfaction rate should result in more negative evaluations.	2x3 between subject	OLS regression coefficient	On average, citizens evaluated hospitals under the satisfied frame as significantly better, with an average difference of 20.5 points (p < .01).	$r(3440)=-0.414$	No	Effect calculated from Table 3, Model 1
Olsen, 2017a	1	When making deliberate decisions about what information to engage with, citizens are more likely to select statistical information over episodic information.	one factor between subject	difference-of proportions tests	Across all three conditions, we find that citizens strongly prefer statistical information from the satisfaction survey over the episodic description of individual patients.	$Z=10.4551$	No	Effect calculated from text (p. 411)
Olsen, 2017b	1	Historical reference points (H1): Providing information about the better past performance of an organization will lower citizens' assessment of current performance, and providing information about worse past performance of an organization will raise citizens' assessment of current performance.	one factor between subject	OLS regression coefficient	Moving on to the historical reference point treatment group, we also find a positive effect for the absolute grade average and a negative effect for the historical grade point average	$r(1135)=-0.247$	No	Effect calculated from Table 4, Model 5

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Olsen, 2017b	2	Historical reference points (H1): Providing information about the better past performance of an organization will lower citizens' assessment of current performance, and providing information about worse past performance of an organization will raise citizens' assessment of current performance.	one factor between subject	OLS regression coefficient	A similar pattern is found for the historical comparison treatment, where the subjects were exposed to last year's unemployment rate for the target municipality: citizens respond negatively to higher unemployment in the target municipality and positively to higher levels of past unemployment (see model 5, Table 4).	r(1142)=0.227	No	Effect calculated from Table 3, Model 5
Olsen, 2017c	2	Inactions are evaluated as being better than actions when the outcome is negative, but worse when the outcome is positive	2x2 between subject	OLS regression coefficient	only for the unemployment case do we see a substantial and significant difference: actions are rewarded less (but still better than inactions) for negative outcomes.	r(1001)=-0.124	No	Effect calculated from Table 1, Column 4
Olsen, 2018	1	Across a set experiments we test how precise political-administrative numbers affect citizens' trust in forecasts, confidence in political-administrative decisions, and the likelihood of achieving future performance goals.	between subject	binomial test	For the growth rate in panel A, 51.6% prefer the lower precise estimate of 1.463% over 1.5% (Binomial test (n = 368), p = .56).	chi2(1)=104.12	No	
Pedersen & Nielsen, 2020		Caseworkers are inclined to make decisions favoring citizens of the same gender as themselves	one factor between subject	LPM regression coefficient	Supporting the gender similarity hypothesis, female caseworkers are associated with a lower probability than male caseworker of reducing visitation for the non-resident parent when the non-resident parent is female of roughly 0.50.	r(47)=-0.358	No	Effect calculated from Table 3, Column 1
Pedersen & Stritch, 2018		Emphasis on internal management that relates to setting challenging but feasible goals has a positive effect on an individual's perception of MTW.	one factor between subject	OLS regression coefficient	However, the goal treatment has a positive effect for the sub-dimensions of ability and integrity, while the participation treatment has a positive effect for benevolence.	r(1823)=0.047	No	Effect calculated from Table 2, Model 1

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Pedersen et al., 2017		Citizens will rate the procedural fairness of a hiring situation lower when the situation appears to be influenced by advocacy from an applicant's social contact.	one factor between subject	OLS regression coefficient	In line with hypothesis 1, we see that the treatment has a negative effect on CPPF. Citizens for whom the hiring process appears to be influenced by advocacy from an applicant's social contacts report a lower CPPF in the size of .66 standard deviations.	r(820)=-0.330	No	Effect calculated from Table 2, Model 1
Pedersen et al., 2018		Caseworkers are more likely to recommend sanctions for ethnic minority clients (here, clients of Middle-Eastern origin) than for ethnic majority clients (here, clients of ethnic Danish origin).	one factor between subject	OLS regression coefficient	In support of Hypothesis 1, we see that caseworkers receiving the Mohammed treatment report being more likely to sanction the client than those assigned to the blinded version of the vignette (β = .824, p = .01).	r(493)=0.094	No	Effect calculated from Table A1, Model 1
Pedersen, 2015		The external activation of PSM increases the amount of time that a person is willing to spend completing a task.	one factor between subject	OLS regression coefficient	The findings show that recipients of the PSM treatments (PSMP and PSMC) are willing to spend more time completing a future research survey than those in the control group (Z0)	r(523)=0.162	No	Effect calculated from Table 2, Model 1
Petersen et al., 2019		Frontline employees are more likely to accept a performance indicator when it displays high performance for their organization compared to when it displays low performance.	2x3 between subject	OLS regression coefficient	Figure 1 clearly shows that the provision of performance information has a negative effect on acceptance when the teachers' high schools have low performance scores	r(642)=-0.148	No	Effect calculated from Table S4, Model 1
Petersen, 2020		H1: When performance information is provided by a vertical source, frontline employees are less likely to perceive the information as relevant and useful than if the performance information is provided by a horizontal source.	one factor between subject	fixed effect linear regression model	However, it is also evident from graph (a) and (b) that when the performance information is provided by a vertical source (the management), it has significant negative effects on the teachers' perceptions of the information as a useful tool for their work compared to both the control group and the two other horizontal sources.	r(1929)=-0.073	No	Effect calculated from Table A1, Model 1
Porumbescu & Grimmelikhuijsen, 2018	1	Hypothesis 1: High procedural fairness will increase willingness to accept government's deci-	2x2 between subject	ANOVA	An ANOVA reveals that the effect of the procedural fairness treatment on participants' willingness to accept the	F(3,134)=30.42	No	

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		sion, which is negatively associated with citizens' voice behavior.			proposed policy was strongly significant, F(3, 134) = 30.42, p = .000, and had a large effect size (partial R2 = .190).			
Porumbescu & Grimmelikhuijsen, 2018	2	Hypothesis 1: High procedural fairness will increase willingness to accept government's decision, which is negatively associated with citizens' voice behavior.	2x2 between subject	ANOVA	An ANOVA demonstrates that there was a significant effect of procedural fairness on participants' willingness to accept the government's decision to reduce the Fire Department's budget, F(3, 151) = 11.37, p = .001, partial R2 = .072	F(3,151)=11.37	No	
Porumbescu et al., 2019		Citizens are more likely to perceive police as trustworthy when information about police performance is positive than when performance information is negative.	one factor between subject	t-test	Participants exposed to positive performance information evaluated police as more trustworthy than those exposed to negative performance information: F(1, 250) = 165.101, p < 0.001, partial eta squared: 0.398.	t(242.244)=12.819	No	Effect calculated from text (p. 222)
Riccucci et al., 2014		the police response to domestic violence is an area in which the gender composition of the police unit involved would clearly seem to matter, as most perpetrators of domestic violence are men and most of the victims are women	2x2 between subject	OLS regression coefficient	In other words, the more representative agency (with 6 females and 4 males) was seen to be doing a better job, regardless of its reported performance (arrest rate). This representativeness (gender factor) effect is statistically significant (see table 2).	r(766)=0.0709	Yes	Effect size calculated from p-value in Table 2 Model 1
Teodoro & An, 2018	1	Support for government action is higher (lower) when it is associated with a specific agency's name than when it is not.	one factor between subject	t-test	Together these results offer strong support for hypothesis H1.	F(1,1941)=208.87	No	Effect calculated from Table 3, Column 1 and text (p. 330)
Teodoro & An, 2018	2	Support for government action is higher (lower) when it is associated with a specific agency's name than when it is not.	one factor between subject	t-test	Together these results offer strong support for hypothesis H1.	F(1,1932)=57.86	No	Effect calculated from Table 3, Column 3 and text (p. 330)

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Teodoro & An, 2018	3	Support for government action is higher (lower) when it is associated with a specific agency's name than when it is not.	one factor between subject	t-test	Together these results offer strong support for hypothesis H1.	$F(1,1281)=26.5$	No	Effect calculated from Table 3, Column 7 and text (p. 330)
Thomassen et al., 2017	1	In H1 we hypothesized that compensating customers would lead to more positive evaluations.	2x2x2 between subject	t-test	Results of the ANOVAs show that there are main Offering Compensation effects for all dependent variables. The pattern of means (Comp versus No-Comp) moves as we hypothesized (see table 2), concluding that H1 can be confirmed	$t(153.511)=-7.573$	No	Effect calculated from Table 2
Thomassen et al., 2017	2	In H1 we hypothesized that compensating customers would lead to more positive evaluations.	2x2x2 between subject	t-test	The pattern of means was as expected by the hypothesis (see table 4), concluding that H1 can be confirmed.	$t(883.760)=-20.081$	No	Effect calculated from Table 4
Tummers et al., 2016		Red tape has a negative effect on citizen satisfaction.	one factor between subject	OLS regression coefficient	This is statistically significant ($F=26.989$, $p < .01$), thereby providing support for our first hypothesis: red tape negatively affects citizen satisfaction.	$F(1,178)=26.989$	No	Effect calculated from Table 5, Model 1
Tummers et al., 2018		Leader's use of task communication increases employee vitality.	2x2 between subject	Anova	Results of the ANOVA revealed a significant main effect of leader's task communication on vitality, $F(1, 101) = 21.86$, $p < .01$, partial $\eta2 = .18$, with participants in the high leader task communication condition reporting greater experienced vitality ($M = 3.60$, $SD = .79$) than those in the low leader task communication condition ($M = 3.04$, $SD = .79$). This confirms Hypothesis 1.	$t(99.843)=3.579$	No	Effect calculated from text (p.369)
Van De Walle & Van Ryzin, 2011		whether and how the order of specific and general questions about local public services influences reported levels of satisfaction with these services.	one factor between subject	t-test	In Version A, the overall satisfaction question perhaps reflects some of this positive priming.	$t(1636)=3.34$	No	Effect calculated from text (p. 1444)

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
Van der Voet, 2019	1	H1: Distributive equality is negatively related to resistance to change.	one factor between subject	OLS regression coefficient	The overall conclusion of study 1 is that distributive equality and distributive equity have independent, additive effects on employee resistance to change.	r(2124)=-0.315	No	Effect calculated from Table 3, Column 1
Van der Voet, 2019	2	H1: Distributive equality is negatively related to resistance to change.	one factor between subject	OLS regression coefficient	The overall conclusion of study 1 is that distributive equality and distributive equity have independent, additive effects on employee resistance to change.	r(1559)=-0.096	No	Effect calculated from Table 3, Column 1
van Engen et al., 2019		Hypothesis 1: Policy consistency (as against policy inconsistency) has a positive effect on how frontline workers perceive a policy's (a) meaningfulness and (b) legitimacy.	2x2 between subject	t-test	Overall, the results of our analyses support hypothesis 1: policy consistency, as expected, has a positive effect on teachers' perceptions of meaningfulness and, particularly, of legitimacy.	t(387.765)=5.101	No	Effect calculated from Table 2
Van Puyvelde et al., 2016		By investigating managerial objectives, we test the simultaneous need for both control (agency theory) and collaboration (stewardship theory) in public and non-profit governance.	discrete choice experiment	OLS regression coefficient	The estimated mean coefficients of the attributes are all statistically significant, which implies that nursing home managers gain positive utility from the realization of the six objectives included in the DCE.	r(631)=0.1038	Yes	Effect size calculated from p-value in Table 1
Vashdi, 2013		Teams in public organizations that engage in structured reflexivity will pay more attention to detail compared with other such teams that operate under similar conditions but engage in peer assessment.	one factor between subject	OLS regression coefficient	I found that members of public sector teams that engaged in guided team reflexivity paid more attention to detail, felt psychologically safer, were more cooperative, and completed their tasks more quickly than members of teams which either engaged in peer assessment or received no form of formal feedback.	r(57)=-0.318	No	Effect calculated from Table 2, Model 2
Walker et al., 2018	1	We anticipate that priming people to consider different aspects of performance affects the way	3x2 between subject	t-test	The ANOVA showed a main vignette effect: a statistically significant difference in performance ratings for treatments (bottom-left panel of table 2). The post hoc analysis showed that the	F(2,552)=3.953	No	Effect calculated from Table 2

Article	Study	(1) Quoted text from original paper indicating prediction of interest to researchers	(2) Study design	(3) Key statistical result	(4) Quoted text from original paper with statistical results	(5) Results	Effect based on sig. level	Comment
		information about the effectiveness of a program is interpreted.			participants who read the equity and effectiveness (by themselves) vignettes demonstrated higher satisfaction with the organization than those who read the efficiency vignette. Notably, internal stakeholders rated effectiveness (by itself) above both efficiency and equity, whereas the external stakeholders offered the greatest support for equity.			
Walker et al., 2018	2	We anticipate that priming people to consider different aspects of performance affects the way information about the effectiveness of a program is interpreted.	3x2 between subject	t-test	The results indicate that external-archival data are preferred over the internal ones in the mathematics secondary education vignettes (left panel of table 3).	F(3,551)=9.538	No	Effect calculated from Table 3
Weibel et al., 2010		Pay for performance affects the relation between intrinsic motivation and effort negatively: High performance-contingent rewards reduce the effect of intrinsic motivation on effort.	Conjoint experiment	OLS regression coefficient	In line with Hypotheses H2, the findings show that pay for performance has a strong negative effect on intrinsic motivation	r(435)=-0.104	No	Effect calculated from Table 4, Model 4
Worthy et al., 2016		Local authorities will be more responsive to an FOI request than to an informal ask	one factor between subject	Ordered logistic regression coefficient	parishes contacted through an FOI were more than twice as likely to respond than those that were simply "asked," with 438 FOIs answered against 199 asks: 20.62% against 9.21%.	r(4305)=0.038	No	Effect calculated from Table 4, Model 1
Yackee, 2015	1	Specifically, I hypothesize that people with higher capacity view their participation as more efficacious (internal efficacy).	one factor between subject	Probit	Overall, my expectation—that a perceived business advantage exists among participants active in rulemaking—is well supported in the survey experiment.	r(288)=0.1445	No	Effect size calculated from Table 2, Model 2

References

Alon-Barkat, S. (2020). The emotive effect of government branding on citizens' trust and its boundaries: Does the personal relevance of the policy issue matter? *Public Administration*, *98*(3), 551–569. https://doi.org/10.1111/padm.12647

An, S.-H., Meier, K. J., Bøllingtoft, A., & Andersen, L. B. (2019). Employee Perceived Effect of Leadership Training: Comparing Public and Private Organizations. *International Public Management Journal*, *22*(1), 2–28. https://doi.org/10.1080/10967494.2018.1497739

Andersen, S. C. (2017). From Passive to Active Representation—Experimental Evidence on the Role of Normative Values in Shaping White and Minority Bureaucrats' Policy Attitudes. *Journal of Public Administration Research and Theory*, *27*(3), 400–414. https://doi.org/10.1093/jopart/mux006

Andersen, S. C., & Guul, T. S. (2019). Reducing Minority Discrimination at the Front Line—Combined Survey and Field Experimental Evidence. *Journal of Public Administration Research and Theory*, *29*(3), 429–444. https://doi.org/10.1093/jopart/muy083

Andersen, S. C., & Hjortskov, M. (2016). Cognitive Biases in Performance Evaluations. *Journal of Public Administration Research and Theory*, *26*(4), 647–662. https://doi.org/10.1093/jopart/muv036

Andersen, S. C., & Jakobsen, M. L. (2018). Political Pressure, Conformity Pressure, and Performance Information as Drivers of Public Sector Innovation Adoption. *International Public Management Journal*, *21*(2), 213–242. https://doi.org/10.1080/10967494.2018.1425227

Andersen, S. C., & Moynihan, D. P. (2018). How Do Socially Distinctive Newcomers Fare? Evidence from a Field Experiment: How Do Socially Distinctive Newcomers Fare? Evidence from a Field Experiment. *Public Administration Review*, *78*(6), 874–882. https://doi.org/10.1111/puar.12957

Anderson, D. M., & Stritch, J. M. (2016). Goal Clarity, Task Significance, and Performance: Evidence From a Laboratory Experiment. *Journal of Public Administration Research and Theory*, *26*(2), 211–225. https://doi.org/10.1093/jopart/muv019

Avellaneda, C. N., & Olvera, J. G. (2018). Chief executives' approval of immigrants: Evidence from a survey experiment of 101 Latin American and Caribbean mayors. *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.14

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. https://doi.org/10.1080/10967494.2015.1022676

Baekgaard, M. (2017). Prospect theory and public service outcomes: Examining risk preferences in relation to public sector reforms. *Public Administration*, *95*(4), 927–942. https://doi.org/10.1111/padm.12324

Baekgaard, M., Belle, N., Serritzlew, S., Sicilia, M., & Steccolini, I. (2019). Performance information in politics: How framing, format, and rhetoric matter to politicians' preferences. *Journal of Behavioral Public Administration*, *2*(2). https://doi.org/10.30636/jbpa.22.67

Baekgaard, M., & George, B. (2018). Equal Access to the Top? Representative Bureaucracy and Politicians' Recruitment Preferences for Top Administrative Staff. *Journal of Public Administration Research and Theory*, *28*(4), 535–550. https://doi.org/10.1093/jopart/muy038

Baekgaard, M., James, O., Serritzlew, S., & Ryzin, G. G. V. (2020). Citizens' motivated reasoning about public performance: experimental findings from the US and Denmark. *International Public Management Journal*, *23*(2), 186–204. https://doi.org/10.1080/10967494.2019.1659891

Baekgaard, M., & Kjaergaard, M. (2016). Intergovernmental Grants and Public Expenditures: Evidence from a Survey Experiment. *Local Government Studies*, *42*(2), 189–207. https://doi.org/10.1080/03003930.2015.1110521

Baekgaard, M., & Serritzlew, S. (2016). Interpreting Performance Information: Motivated Reasoning or Unbiased Comprehension. *Public Administration Review*, *76*(1), 73–82. https://doi.org/10.1111/puar.12406

Baekgaard, M., & Serritzlew, S. (2020). Those Who Understand It will not be Persuaded: A Performance Information Paradox. *International Public Management Journal*, *23*(1), 138–160. https://doi.org/10.1080/10967494.2018.1461152

Barnes, T. D., Beaulieu, E., & Saxton, G. W. (2018). Restoring trust in the police: Why female officers reduce suspicions of corruption. *Governance*, *31*(1), 143–161. https://doi.org/10.1111/gove.12281

Barrows, S., Henderson, M., Peterson, P. E., & West, M. R. (2016). Relative Performance Information and Perceptions of Public Service Quality: Evidence From American School Districts. *Journal of Public Administration Research and Theory*, *26*(3), 571–583. https://doi.org/10.1093/jopart/muw028

Belardinelli, P., Bellé, N., Sicilia, M., & Steccolini, I. (2018). Framing Effects under Different Uses of Performance Information: An Experimental Study on Public Managers: Framing Effects under Different Uses of Performance Information: An Experimental Study on Public Managers. *Public Administration Review*, *78*(6), 841–851. https://doi.org/10.1111/puar.12969

Bellé, N. (2013). Experimental Evidence on the Relationship between Public Service Motivation and Job Performance. *Public Administration Review*, *73*(1), 143–153. https://doi.org/10.1111/j.1540-6210.2012.02621.x

Bellé, N. (2015). Performance-Related Pay and the Crowding Out of Motivation in the Public Sector: A Randomized Field Experiment. *Public Administration Review*, *75*(2), 230–241. https://doi.org/10.1111/puar.12313

Bellé, N., Belardinelli, P., Cantarelli, P., & Mele, V. (2019). On iron cages and suboptimal choices: An experimental test of the micro-foundations of isomorphism in the public sector. *International Public Management Journal*, *22*(2), 373–414. https://doi.org/10.1080/10967494.2018.1494066

Bellé, N., & Cantarelli, P. (2018). The Role of Motivation and Leadership in Public Employees' Job Preferences: Evidence from Two Discrete Choice Experiments. *International Public Management Journal*, *21*(2), 191–212. https://doi.org/10.1080/10967494.2018.1425229

Bellé, N., & Cantarelli, P. (2019). Do Ethical Leadership, Visibility, External Regulation, and Prosocial Impact Affect Unethical Behavior? Evidence From a Laboratory and a Field Experiment. *Review of Public Personnel Administration*, *39*(3), 349–371. https://doi.org/10.1177/0734371X17721301

Bellé, N., Cantarelli, P., & Belardinelli, P. (2017). Cognitive Biases in Performance Appraisal: Experimental Evidence on Anchoring and Halo Effects With Public Sector Managers and Employees. *Review of Public Personnel Administration*, *37*(3), 275–294. https://doi.org/10.1177/0734371X17704891

Bellé, N., Cantarelli, P., & Belardinelli, P. (2018). Prospect Theory Goes Public: Experimental Evidence on Cognitive Biases in Public Policy and Management Decisions. *Public Administration Review*, *78*(6), 828–840. https://doi.org/10.1111/puar.12960

Bergner, C., Desmarais, B. A., & Hird, J. (2019). Speaking truth in power: Scientific evidence as motivation for policy activism. *Journal of Behavioral Public Administration*, *2*(1). https://doi.org/10.30636/jbpa.21.27

Blom-Hansen, J., Baekgaard, M., & Serritzlew, S. (2016). Shaping Political Preferences: Information Effects in Political-Administrative Systems. *Local Government Studies*, *42*(1), 119–138. https://doi.org/10.1080/03003930.2015.1084925

Borry, E. L., DeHart-Davis, L., Kaufmann, W., Merritt, C. C., Mohr, Z., & Tummers, L. (2018). Formalization and consistency heighten organizational rule following: Experimental and survey evidence. *Public Administration*, *96*(2), 368–385. https://doi.org/10.1111/padm.12407

Bouwman, R., van Thiel, S., van Deemen, A., & Rouwette, E. (2018). Accountability and Coalitions: Evidence from a Negotiation Experiment. *Public Administration Review*, *78*(1), 37–47. https://doi.org/10.1111/puar.12858

Brewer, G. A., & Brewer, G. A. (2011). Parsing Public/Private Differences in Work Motivation and Performance: An Experimental Study. *Journal of Public Administration Research and Theory*, *21*(Supplement 3), i347–i362. https://doi.org/10.1093/jopart/mur030

Bro, L. L., & Jensen, U. T. (2020). Does transformational leadership stimulate user orientation? Evidence from a field experiment. *Public Administration*, *98*(1), 177–193. https://doi.org/10.1111/padm.12612

Cantarelli, P., Bellé, N., & Belardinelli, P. (2020). Behavioral Public HR: Experimental Evidence on Cognitive Biases and Debiasing Interventions. *Review of Public Personnel Administration*, *40*(1), 56–81. https://doi.org/10.1177/0734371X18778090

Cantarelli, P., Bellé, N., & Longo, F. (2020). Exploring the motivational bases of public mission-driven professions using a sequential-explanatory design. *Public Management Review*, *22*(10), 1535–1559. https://doi.org/10.1080/14719037.2019.1642950

Charbonneau, É., & Van Ryzin, G. G. (2015). Benchmarks and Citizen Judgments of Local Government Performance: Findings from a survey experiment. *Public Management Review*, *17*(2), 288–304. https://doi.org/10.1080/14719037.2013.798027

Christensen, J. (2018). Biased, not blind: An experimental test of self-serving biases in service users' evaluations of performance information. *Public Administration*, *96*(3), 468–480. https://doi.org/10.1111/padm.12520

Christensen, J., Dahlmann, C. M., Mathiasen, A. H., Moynihan, D. P., & Petersen, N. B. G. (2018). How Do Elected Officials Evaluate Performance? Goal Preferences, Governance Preferences, and the

Process of Goal Reprioritization. *Journal of Public Administration Research and Theory*, *28*(2), 197–211. https://doi.org/10.1093/jopart/muy001

de Fine Licht, J. (2014). Policy Area as a Potential Moderator of Transparency Effects: An Experiment. *Public Administration Review*, *74*(3), 361–371. https://doi.org/10.1111/puar.12194

De Fine Licht, J., Naurin, D., Esaiasson, P., & Gilljam, M. (2014). When Does Transparency Generate Legitimacy? Experimenting on a Context-Bound Relationship: When Does Transparency Generate Legitimacy? *Governance*, *27*(1), 111–134. https://doi.org/10.1111/gove.12021

Demaj, L. (2017). What Can Performance Information Do to Legislators? A Budget-Decision Experiment with Legislators. *Public Administration Review*, *77*(3), 366–379. https://doi.org/10.1111/puar.12713

Deslatte, A. (2019). A bayesian approach for behavioral public administration: Citizen assessments of local government sustainability performance. *Journal of Behavioral Public Administration*, *2*(1). https://doi.org/10.30636/jbpa.21.32

Doberstein, C., & Charbonneau, É. (2020). The origins and effects of public servant confidence in whistleblowing protection regimes. *Public Administration*, *98*(3), 643–658. https://doi.org/10.1111/padm.12644

Filtenborg, A. F., Gaardboe, F., & Sigsgaard-Rasmussen, J. (2017). Experimental replication: an experimental test of the expectancy disconfirmation theory of citizen satisfaction. *Public Management Review*, *19*(9), 1235–1250. https://doi.org/10.1080/14719037.2017.1295099

George, B., Desmidt, S., Nielsen, P. A., & Baekgaard, M. (2017). Rational planning and politicians' preferences for spending and reform: replication and extension of a survey experiment. *Public Management Review*, *19*(9), 1251–1271. https://doi.org/10.1080/14719037.2016.1210905

Geys, B., & Sørensen, R. J. (2018). Never Change a Winning Policy? Public Sector Performance and Politicians' Preferences for Reforms: Never Change a Winning Policy? Public Sector Performance and Politicians' Preferences for Reforms. *Public Administration Review*, *78*(2), 206–216. https://doi.org/10.1111/puar.12824

Grimmelikhuijsen, S., John, P., Meijer, A., & Worthy, B. (2019). Do freedom of information laws increase transparency of government? A replication of a field experiment. *Journal of Behavioral Public Administration*, *2*(1). https://doi.org/10.30636/jbpa.12.34

Grimmelikhuijsen, S., & Klijn, A. (2015). The effect of judicial transparency on public trust: Evidence from a field experiment. *Public Administration*, *93*(4), 995–1011. https://doi.org/10.1111/padm.12149

Grimmelikhuijsen, S., & Porumbescu, G. A. (2017). Reconsidering the expectancy disconfirmation model. Three experimental replications. *Public Management Review*, *19*(9), 1272–1292. https://doi.org/10.1080/14719037.2017.1282000

Grimmelikhuijsen, S., Porumbescu, G., Hong, B., & Im, T. (2013). The Effect of Transparency on Trust in Government: A Cross-National Comparative Experiment. *Public Administration Review*, *73*(4), 575–586. https://doi.org/10.1111/puar.12047

Grimmelikhuijsen, S., Vries, F. de, & Zijlstra, W. (2018). Breaking bad news without breaking trust: The effects of a press release and newspaper coverage on perceived trustworthiness. *Journal of Behavioral Public Administration*, *1*(1), Article 1. https://doi.org/10.30636/jbpa.11.16

Guardino, M., & Mettler, S. (2020). Revealing the "Hidden welfare state": How policy information influences public attitudes about tax expenditures. *Journal of Behavioral Public Administration*, *3*(1). https://doi.org/10.30636/jbpa.31.108

Hamidullah, M. F., Riccucci, N. M., & Lee, I. P. (2020). Citizens' perceptions of closing the gender pay gap: an experimental study. *Public Management Review*, 1–24. https://doi.org/10.1080/14719037.2020.1722207

Hattke, F., Hensel, D., & Kalucza, J. (2020). Emotional Responses to Bureaucratic Red Tape. *Public Administration Review*, *80*(1), 53–63. https://doi.org/10.1111/puar.13116

Hattke, F., & Kalucza, J. (2019). What influences the willingness of citizens to coproduce public services? Results from a vignette experiment. *Journal of Behavioral Public Administration*, *2*(1). https://doi.org/10.30636/jbpa.21.60

Herian, M. N., Hamm, J. A., Tomkins, A. J., & Pytlik Zillig, L. M. (2012). Public Participation, Procedural Fairness, and Evaluations of Local Governance: The Moderating Role of Uncertainty. *Journal of Public Administration Research and Theory*, *22*(4), 815–840. https://doi.org/10.1093/jopart/mur064

Hjortskov, M. (2017). Priming and context effects in citizen satisfaction surveys. *Public Administration*, *95*(4), 912–926. https://doi.org/10.1111/padm.12346

Hollibaugh, G. E., Miles, M. R., & Newswander, C. B. (2020). Why Public Employees Rebel: Guerrilla Government in the Public Sector. *Public Administration Review*, *80*(1), 64–74. https://doi.org/10.1111/puar.13118

Holm, J. M. (2017). Double standards? How historical and political aspiration levels guide managerial performance information use. *Public Administration*, *95*(4), 1026–1040. https://doi.org/10.1111/padm.12379

Houlberg, K., Olsen, A. L., & Pedersen, L. H. (2016). Spending and cutting are two different worlds: experimental evidence from Danish local councils. *Local Government Studies*, *42*(5), 821–841. https://doi.org/10.1080/03003930.2016.1189414

Hvidman, U., & Andersen, S. C. (2016). Perceptions of Public and Private Performance: Evidence from a Survey Experiment. *Public Administration Review*, *76*(1), 111–120. https://doi.org/10.1111/puar.12441

Ingrams, A., Kaufmann, W., & Jacobs, D. (2020). Testing the open government recipe: Are vision and voice good governance ingredients? *Journal of Behavioral Public Administration*, *3*(1). https://doi.org/10.30636/jbpa.31.114

Jakobsen, M., & Andersen, S. C. (2013). Coproduction and Equity in Public Service Delivery. *Public Administration Review*, *73*(5), 704–713. https://doi.org/10.1111/puar.12094

Jakobsen, M., Jacobsen, C. B., & Serritzlew, S. (2019). Managing the Behavior of Public Frontline Employees through Change-Oriented Training: Evidence from a Randomized Field Experiment. *Journal of Public Administration Research and Theory*, *29*(4), 556–571. https://doi.org/10.1093/jopart/muy080

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. https://doi.org/10.1093/jopart/muq057

James, O., & Moseley, A. (2014). Does performance information about public services affect citizens' perceptions, satisfaction, and voice behaviour? Field experiments with absolute and relative performance information. *Public Administration*, *92*(2), 493–511. https://doi.org/10.1111/padm.12066

James, O., & Petersen, C. (2018). International rankings of government performance and source credibility for citizens: experiments about e-government rankings in the UK and the Netherlands. *Public Management Review*, *20*(4), 469–484. https://doi.org/10.1080/14719037.2017.1296965

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. https://doi.org/10.1093/jopart/muw049

James, O., & Van Ryzin, G. G. (2019). Rates and the judgment of government performance. *Journal of Behavioral Public Administration*, *2*(2). https://doi.org/10.30636/jbpa.22.41

Jensen, U. T., Andersen, L. B., & Jacobsen, C. B. (2019). Only When We Agree! How Value Congruence Moderates the Impact of Goal-Oriented Leadership on Public Service Motivation. *Public Administration Review*, *79*(1), 12–24. https://doi.org/10.1111/puar.13008

Jilke, S., & Tummers, L. (2018). Which Clients are Deserving of Help? A Theoretical Model and Experimental Test. *Journal of Public Administration Research and Theory*, *28*(2), 226–238. https://doi.org/10.1093/jopart/muy002

Jilke, S., Van Ryzin, G. G., & Van de Walle, S. (2016). Responses to Decline in Marketized Public Services: An Experimental Evaluation of Choice Overload. *Journal of Public Administration Research and Theory*, *26*(3), 421–432. https://doi.org/10.1093/jopart/muv021

John, P., & Blume, T. (2018). How best to nudge taxpayers? The impact of message simplification and descriptive social norms on payment rates in a central London local authority. *Journal of Behavioral Public Administration*, *1*(1). https://doi.org/10.30636/jbpa.11.10

Jurcevic, I., & Fyall, R. (2019). Does a business-like approach to diversity in nonprofit organizations have a chilling effect on stakeholders? *Journal of Behavioral Public Administration*, *3*(1). https://doi.org/10.30636/jbpa.31.100

Karens, R., Eshuis, J., Klijn, E.-H., & Voets, J. (2016). The Impact of Public Branding: An Experimental Study on the Effects of Branding Policy on Citizen Trust. *Public Administration Review*, *76*(3), 486–494. https://doi.org/10.1111/puar.12501

Kaufmann, W., & Feeney, M. K. (2014). Beyond the rules: The effect of outcome favourability on red tape perceptions. *Public Administration*, *92*(1), 178–191. https://doi.org/10.1111/padm.12049

Kaufmann, W., & Tummers, L. (2017). The negative effect of red tape on procedural satisfaction. *Public Management Review*, *19*(9), 1311–1327. https://doi.org/10.1080/14719037.2016.1210907

Keiser, L. R., & Miller, S. M. (2020). Does Administrative Burden Influence Public Support for Government Programs? Evidence from a Survey Experiment. *Public Administration Review*, *80*(1), 137–150. https://doi.org/10.1111/puar.13133

Kim, M.-H., Porumbescu, G. A., & Neshkova, M. I. (2019). How Does Race Affect Perceptions of Police Trustworthiness? *International Public Management Journal*, *22*(2), 343–372. https://doi.org/10.1080/10967494.2018.1486343

Kim, S. H., & Kim, S. (2016). Social Desirability Bias in Measuring Public Service Motivation. *International Public Management Journal*, *19*(3), 293–319. https://doi.org/10.1080/10967494.2015.1021497

Kim, S. H., & Kim, S. (2017). Ethnic Differences in Social Desirability Bias: Effects on the Analysis of Public Service Motivation. *Review of Public Personnel Administration*, *37*(4), 472–491. https://doi.org/10.1177/0734371X15605147

Larkin, C., Sanders, M., Andresen, I., & Algate, F. (2019). Testing local descriptive norms and salience of enforcement action: A field experiment to increase tax collection. *Journal of Behavioral Public Administration*, *2*(1). https://doi.org/10.30636/jbpa.21.54

Lee, M. J., Moon, M. J., & Kim, J. (2017). Insights from experiments with duopoly games: rational incremental decision-making. *Public Management Review*, *19*(9), 1328–1351. https://doi.org/10.1080/14719037.2017.1282002

Levitats, Z., Vigoda-Gadot, E., & Vashdi, D. R. (2019). Engage Them through Emotions: Exploring the Role of Emotional Intelligence in Public Sector Engagement. *Public Administration Review*, *79*(6), 841–852. https://doi.org/10.1111/puar.13113

Linos, E., Reinhard, J., & Ruda, S. (2017). Levelling the playing field in police recruitment: Evidence from a field experiment on test performance. *Public Administration*, *95*(4), 943–956. https://doi.org/10.1111/padm.12344

Linos, E., & Riesch, N. (2020). Thick Red Tape and the Thin Blue Line: A Field Study on Reducing Administrative Burden in Police Recruitment. *Public Administration Review*, *80*(1), 92–103. https://doi.org/10.1111/puar.13115

Livnat-Lerer, M., Sulitzeau-Kenan, R., & Kogut, T. (2018). Foresighted outcome effect: A micro-founda-
tion of agents' risk aversion in principal-agent relations. *Journal of Behavioral Public Administration*,
1(1). https://doi.org/10.30636/jbpa.11.12

Lu, X., Xu, H., & Wang, W. (2019). Clients' Help Deservingness, Crowd Situational Stress And Discre-
tionary Decision-making: An Experimental Study Of Regulatory Street-level Bureaucrats In
China. *International Public Management Journal*, 1–26.
https://doi.org/10.1080/10967494.2019.1661892

Marvel, J. D. (2015). Public Opinion and Public Sector Performance: Are Individuals' Beliefs About Per-
formance Evidence-Based or the Product of Anti–Public Sector Bias? *International Public Manage-
ment Journal*, *18*(2), 209–227. https://doi.org/10.1080/10967494.2014.996627

Marvel, J. D., & Girth, A. M. (2016). Citizen Attributions of Blame in Third-Party Governance. *Public Ad-
ministration Review*, *76*(1), 96–108. https://doi.org/10.1111/puar.12474

Mazepus, H., & van Leeuwen, F. (2020). Fairness matters when responding to disasters: An experimental
study of government legitimacy. *Governance*, *33*(3), 621–637. https://doi.org/10.1111/gove.12440

Meyer-Sahling, J.-H., Mikkelsen, K. S., & Schuster, C. (2019). The Causal Effect of Public Service Motiva-
tion on Ethical Behavior in the Public Sector: Evidence from a Large-Scale Survey Experiment.
Journal of Public Administration Research and Theory, *29*(3), 445–459. https://doi.org/10.1093/jop-
art/muy071

Migchelbrink, K., & Van de Walle, S. (2020). When Will Public Officials Listen? A Vignette Experiment
on the Effects of Input Legitimacy on Public Officials' Willingness to Use Public Participation.
Public Administration Review, *80*(2), 271–280. https://doi.org/10.1111/puar.13138

Moseley, A., & Stoker, G. (2015). Putting Public Policy Defaults to the Test: The Case of Organ Donor
Registration. *International Public Management Journal*, *18*(2), 246–264.
https://doi.org/10.1080/10967494.2015.1012574

Muñoz, J., Anduiza, E., & Gallego, A. (2016). Why do voters forgive corrupt mayors? Implicit exchange,
credibility of information and clean alternatives. *Local Government Studies*, *42*(4), 598–615.
https://doi.org/10.1080/03003930.2016.1154847

Nelson, M. J., & Witko, C. (2020). Government reputational effects of COVID-19 public health actions: A job opportunity evaluation conjoint experiment. *Journal of Behavioral Public Administration*, *3*(1). https://doi.org/10.30636/jbpa.31.174

Ngoye, B., Sierra, V., & Ysa, T. (2019). Different Shades of Gray: A Priming Experimental Study on How Institutional Logics Influence Organizational Actor Judgment. *Public Administration Review*, *79*(2), 256–266. https://doi.org/10.1111/puar.13006

Nicholson-Crotty, S., Nicholson-Crotty, J., & Webeck, S. (2019). Are public managers more risk averse? Framing effects and status quo bias across the sectors. *Journal of Behavioral Public Administration*, *2*(1). https://doi.org/10.30636/jbpa.21.35

Nielsen, P. A., & Baekgaard, M. (2015). Performance Information, Blame Avoidance, and Politicians' Attitudes to Spending and Reform: Evidence from an Experiment. *Journal of Public Administration Research and Theory*, *25*(2), 545–569. https://doi.org/10.1093/jopart/mut051

Nielsen, P. A., & Jacobsen, C. B. (2018). Zone of Acceptance under Performance Measurement: Does Performance Information Affect Employee Acceptance of Management Authority? *Public Administration Review*, *78*(5), 684–693. https://doi.org/10.1111/puar.12947

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. https://doi.org/10.1093/jopart/muw060

Olsen, A. L. (2015). Citizen (Dis)satisfaction: An Experimental Equivalence Framing Study. *Public Administration Review*, *75*(3), 469–478. https://doi.org/10.1111/puar.12337

Olsen, A. L. (2017a). Human Interest or Hard Numbers? Experiments on Citizens' Selection, Exposure, and Recall of Performance Information. *Public Administration Review*, *77*(3), 408–420. https://doi.org/10.1111/puar.12638

Olsen, A. L. (2017b). Compared to What? How Social and Historical Reference Points Affect Citizens' Performance Evaluations. *Journal of Public Administration Research and Theory*, *27*(4), 562–580. https://doi.org/10.1093/jopart/mux023

Olsen, A. L. (2017c). Responding to problems: actions are rewarded, regardless of the outcome. *Public Management Review*, *19*(9), 1352–1364. https://doi.org/10.1080/14719037.2017.1281998

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

Pedersen, M. J. (2015). Activating the Forces of Public Service Motivation: Evidence from a Low-Intensity Randomized Survey Experiment. *Public Administration Review*, *75*(5), 734–746. https://doi.org/10.1111/puar.12325

Pedersen, M. J., & Nielsen, V. L. (2020). Bureaucratic decision-making: A multi-method study of gender similarity bias and gender stereotype beliefs. *Public Administration*, *98*(2), 424–440. https://doi.org/10.1111/padm.12622

Pedersen, M. J., & Stritch, J. M. (2018). Internal Management and Perceived Managerial Trustworthiness: Evidence From a Survey Experiment. *The American Review of Public Administration*, *48*(1), 67–81. https://doi.org/10.1177/0275074016657179

Pedersen, M. J., Stritch, J. M., & Taggart, G. (2017). Citizen perceptions of procedural fairness and the moderating roles of 'belief in a just world' and 'public service motivation' in public hiring. *Public Administration*, *95*(4), 874–894. https://doi.org/10.1111/padm.12353

Pedersen, M. J., Stritch, J. M., & Thuesen, F. (2018). Punishment on the Frontlines of Public Service Delivery: Client Ethnicity and Caseworker Sanctioning Decisions in a Scandinavian Welfare State. *Journal of Public Administration Research and Theory*, *28*(3), 339–354. https://doi.org/10.1093/jopart/muy018

Petersen, N. B. G. (2020). How the source of performance information matters to learning on the frontlines: evidence from a survey experiment. *International Public Management Journal*, *23*(2), 276–291. https://doi.org/10.1080/10967494.2019.1699214

Petersen, N. B. G., Laumann, T. V., & Jakobsen, M. (2019). Acceptance or Disapproval: Performance Information in the Eyes of Public Frontline Employees. *Journal of Public Administration Research and Theory*, *29*(1), 101–117. https://doi.org/10.1093/jopart/muy035

Porumbescu, G. A., & Grimmelikhuijsen, S. (2018). Linking Decision-Making Procedures to Decision Acceptance and Citizen Voice: Evidence From Two Studies. *The American Review of Public Administration*, *48*(8), 902–914. https://doi.org/10.1177/0275074017734642

Porumbescu, G. A., Neshkova, M. I., & Huntoon, M. (2019). The effects of police performance on agency trustworthiness and citizen participation. *Public Management Review*, *21*(2), 212–237. https://doi.org/10.1080/14719037.2018.1473473

Riccucci, N. M., Van Ryzin, G. G., & Lavena, C. F. (2014). Representative Bureaucracy in Policing: Does It Increase Perceived Legitimacy? *Journal of Public Administration Research and Theory*, *24*(3), 537–551. https://doi.org/10.1093/jopart/muu006

Teodoro, M. P., & An, S.-H. (2018). Citizen-Based Brand Equity: A Model and Experimental Evaluation. *Journal of Public Administration Research and Theory*, *28*(3), 321–338. https://doi.org/10.1093/jopart/mux044

Thomassen, J. P., Leliveld, M. C., Van de Walle, S., & Ahaus, K. (2017). Compensating citizens for poor service delivery: Experimental research in public and private settings. *Public Administration*, *95*(4), 895–911. https://doi.org/10.1111/padm.12339

Tummers, L., Steijn, B., Nevicka, B., & Heerema, M. (2018). The Effects of Leadership and Job Autonomy on Vitality: Survey and Experimental Evidence. *Review of Public Personnel Administration*, *38*(3), 355–377. https://doi.org/10.1177/0734371X16671980

Tummers, L., Weske, U., Bouwman, R., & Grimmelikhuijsen, S. (2016). The Impact of Red Tape on Citizen Satisfaction: An Experimental Study. *International Public Management Journal*, *19*(3), 320–341. https://doi.org/10.1080/10967494.2015.1027800

Van De Walle, S., & Van Ryzin, G. G. (2011). The order of questions in a survey on citizen satisfaction with public services: Lessons from a split-ballot experiment. *Public Administration*, *89*(4), 1436–1450. https://doi.org/10.1111/j.1467-9299.2011.01922.x

Van der Voet, J. (2019). Distributive Justice and Resistance to Change in Cutback Management: Evidence from a Survey Experiment and a Longitudinal Field Study. *International Public Management Journal*, 1–26. https://doi.org/10.1080/10967494.2019.1676347

van Engen, N., Steijn, B., & Tummers, L. (2019). Do consistent government policies lead to greater meaningfulness and legitimacy on the front line? *Public Administration*, *97*(1), 97–115. https://doi.org/10.1111/padm.12570

Van Puyvelde, S., Caers, R., Du Bois, C., & Jegers, M. (2016). Managerial Objectives and the Governance of Public and Non-Profit Organizations. *Public Management Review*, *18*(2), 221–237. https://doi.org/10.1080/14719037.2014.969760

Vashdi, D. R. (2013). Teams in Public Administration: A Field Study of Team Feedback and Effectiveness in the Israeli Public Healthcare System. *International Public Management Journal*, *16*(2), 275–306. https://doi.org/10.1080/10967494.2013.817255

Walker, R. M., Lee, M. J., James, O., & Ho, S. M. Y. (2018). Analyzing the Complexity of Performance Information Use: Experiments with Stakeholders to Disaggregate Dimensions of Performance, Data Sources, and Data Types. *Public Administration Review*, *78*(6), 852–863. https://doi.org/10.1111/puar.12920

Weibel, A., Rost, K., & Osterloh, M. (2010). Pay for Performance in the Public Sector--Benefits and (Hidden) Costs. *Journal of Public Administration Research and Theory*, *20*(2), 387–412. https://doi.org/10.1093/jopart/mup009

Worthy, B., John, P., & Vannoni, M. (2016). Transparency at the Parish Pump: A Field Experiment to Measure the Effectiveness of Freedom of Information Requests in England. *Journal of Public Administration Research and Theory*, muw063. https://doi.org/10.1093/jopart/muw063

Yackee, S. W. (2015). Participant Voice in the Bureaucratic Policymaking Process. *Journal of Public Administration Research and Theory*, *25*(2), 427–449. https://doi.org/10.1093/jopart/muu007