


ORIGINAL ARTICLE

# Making the Grade: Policy Design and Effects of Information about Government Performance

Michael Henderson<sup>1</sup> and Belinda C. Davis<sup>2</sup> 

<sup>1</sup>Manship School of Mass Communication, Louisiana State University, Baton Rouge, LA, USA and

<sup>2</sup>Department of Political Science, Louisiana State University, Baton Rouge, LA, USA

**Corresponding author:** Michael Henderson; Email: [mbhende1@lsu.edu](mailto:mbhende1@lsu.edu)

(Received 18 August 2023; revised 08 February 2024; accepted 29 February 2024)

## Abstract

Increasingly, governments report on public service quality, which has the potential to inform evaluations of performance that underlie voters' opinions and behaviors. We argue these policies have important effects that go beyond informing voters. Specifically, we contend that the format in which policymakers choose to report information will steer the direction of opinion by exacerbating or mitigating biases in information processing. Using the case of school accountability systems in the United States and a variety of experimental and observational approaches, we find that letter grade systems for rating public school performance, as opposed to other reporting formats, exacerbate negativity bias. Public opinion proves more responsive to negative information than to positive information in letter grade systems than in alternate formats. Policymakers, then, do not simply inform public opinion; rather, their decisions about how to present information shape the interpretations that voters ultimately draw from the information provided.

**Keywords:** accountability; performance reporting; negativity bias; expectation disconfirmation; education policy

Governments routinely issue reports about the quality of public services, an increasingly widespread practice across the domains of education, social services, health care, waste disposal, and environmental quality (Baekgaard and Serritzlew 2015; James 2011; James and John 2007). The normative appeal of these policies arises from the idea that individuals best hold governments accountable when they can access accurate information about performance. In requiring public reporting about service quality, governments subsidize access to this information. Research on these policies has largely focused on the question of whether reports on service quality have this effect – that is, whether they shape political attitudes and behavior. We turn from the rudimentary question of *whether* such reporting matters to an examination of *the conditions under which* it matters.

We emphasize the importance of policy design in creating conditions that give rise to particular kinds of informational effects. Officials' choices in the design of

policy matter for the kind of effects the policy has on attitudes and behavior (Ingram and Schneider 1993). For example, decisions about whether a policy is means-tested versus universal or about bureaucratic responsiveness shape downstream attitudes about policy and government (Hacker, Mettler, and Pinderhughes 2005). We argue that the specific choices made regarding how to communicate service quality in government reports likewise shape their effects. We root our claim in research on cognitive biases in information processing. These reporting policies vary tremendously in design – specifically in the content and format of reports. Our contention is that policymakers' decisions about design influence the judgments citizens ultimately make about government performance by exacerbating or mitigating these biases in information processing.

To test this theory, we examine the case of school accountability systems in the United States. Using several methodological approaches, we find that states' decisions to use letter grade systems to rate public schools (as opposed to other reporting formats) have an asymmetric effect such that attitudes respond more to negative information than to positive information. Furthermore, we show that this asymmetry reflects, in part, a tendency of these policies to exacerbate *negativity bias* in how people process information.

The analysis highlights the importance of considering policy design when assessing the consequences of reporting policies. In doing so, we go beyond the benign view of these policies as simply information subsidies. Because individuals are prone to biases in information processing, certain formats for government reporting can play to these biases. As such, the direction and scope of these policies' effects on public opinion depend on how policymakers choose to design them. The results draw attention to the power that governments have, not simply to *inform* public opinion, but to *shape* it.

## Performance reporting and information processing

From a normative perspective, the connection between individuals' assessments of government performance and their political response to it fosters accountability only insofar as beliefs reflect actual performance. Citizens must have informed judgments about how well government is performing if they are to hold officials accountable (Iyengar 1987; James and John 2007). Yet, most citizens lack the incentive to seek out information about the quality of services (Downs 1957) and lack this sort of policy knowledge (Delli Carpini and Keeter 1996). Performance reporting offers a way to provide this information to voters (Van de Walle and Roberts 2011). By disseminating information, these reporting regimes subsidize its cost and may improve the accuracy of evaluations.

To investigate this, scholars often turn to American public education – a policy domain rich in government-issued reports on performance (e.g., Berry and Howell 2007; Chingos, Henderson, and West 2012; Clinton and Grissom 2015; Holbein 2016; Kogan, Lavertu, and Peskowitz 2016a, 2016b; Rhodes 2015).<sup>1</sup> Yet, despite evidence that beliefs about school performance shape attitudes about education policy (e.g., Moe 2001; Peterson, Henderson, and West 2014), the evidence is mixed about

<sup>1</sup>The public administration literature has explored the impact of reporting policies in other sectors (Barrows et al. 2016; James 2011; Van Ryzin and Lavena 2013).

whether the information provided through school accountability policies shifts these beliefs. For example, the kind of information typically reported moves opinions about the quality of local public schooling in Florida (Chingos, Henderson, and West 2012) but not in Tennessee (Clinton and Grissom 2015). Similarly, evidence from North Carolina indicates that school accountability information shapes election outcomes (Holbein 2016), but evidence from South Carolina indicates the effect varies across elections (Berry and Howell 2007). Evidence from Ohio does not show any effect on school board elections (Kogan, Lavertu, and Peskowitz 2016b) but does show an effect for school tax elections (Kogan, Lavertu, and Peskowitz 2016a).

We offer a theoretical argument for the heterogeneity of information effects rooted in the relationship between how officials design policies and how people process information. We argue that policymakers' decisions about how to present information affect how people process it. Specifically, we theorize that formats with more easily recognized valence (i.e., positive versus negative) are prone to asymmetric effects in which negative information is more influential than positive information.

Our argument builds upon two theoretical perspectives. The first concerns the interpretive effects of policy design – that is, how structures of policy shape the inferences drawn about the policy and the government behind it (Ingram and Schneider 1993). For example, benefit programs that camouflage the role of government are less likely to produce positive attitudes toward the program among beneficiaries than programs in which the government's role in providing the benefit is clearer (Hacker, Mettler, and Pinderhughes 2005). There is little reason to assume that the interpretive effects of performance reporting would be less sensitive to design. Indeed, information about government performance is often ambiguous, leaving it open to multiple interpretations, which can shape how voters use it to evaluate government. Measures of school performance, especially those based on test scores, the meaning and relevance of which are difficult for most individuals to interpret without guidance, are no exception (Hambleton and Slater 1995). Unsurprisingly, opinions about school quality vary across states in a manner consistent with differences in the structure of accountability systems (Rhodes 2015), and experimental evidence shows the effect of hypothetical information varies across formats (Jacobsen, Snyder, and Saultz 2014).

The second theoretical perspective concerns information processing. Individuals rarely process information in an unbiased way. Rather, information processing is fraught with various distortions that lend some kinds of information greater influence. Although these biases are widely recognized in research on political psychology and behavior generally, political science research on the effects of reporting policies has largely neglected them.<sup>2</sup>

Negativity bias is one form of biased cognition that yields asymmetric effects. The pattern, common in psychological studies, reveals individuals assign more weight to negative information than to positive information (see Kanouse and Hanson 1972 for a review). Similarly, individuals are more responsive to negative information when evaluating the economy, the president, candidates for office, or the

<sup>2</sup>The literature in public administration goes further in incorporating theories of biased information processing (e.g., Baekgaard and Serritzlew 2015; James and Van Ryzin 2017), but it too has yet to examine how policy design shapes these effects.

government as a whole (Fridkin and Kenney 2004; Kernell 1977; Lau 1985; Niven 2000; Soroka 2006).

The expectation disconfirmation model (EDM) describes another bias in which cognition privileges unexpected information. People update their evaluations in accordance with new information only when it contrasts with prior belief (Hjortskov 2019).<sup>3</sup> For example, if a government report provides negative information about service quality, then people would view the service more negatively *if they previously held a relatively positive view*. Similarly, people's assessments would improve if the report offered positive information when people expect negative information. Although the EDM allows for both positive and negative shifts, it would nevertheless yield asymmetric effects when prior beliefs tend to the opposite valence of new information – and asymmetrically negative effects when these prior positive assessments confront less rosy reports.

Implicit in both negativity bias and the EDM is the idea that recipients are able to identify the valence of information. Indeed, individuals are more prone to negativity bias when they are more familiar with the format of the information (Kanouse and Hanson 1972). This is akin to the concept of accessibility, the tendency to attach weight to concepts or considerations that are recognizable (Iyengar 1990). Individuals are better at recognizing the valence of signals and, thus, more prone to biases that privilege information by valence, when they are familiar with features that characterize valence.

For public sector performance reporting, this means designs in which the valence of a signal is more accessible are more prone to invoke biases in response. Policymakers, then, do not simply inform when launching informational policies; rather, decisions about the format of information shape public opinion through the interaction between design and cognition. When evaluating the information effects of reports, the key question is not whether they have a general effect but rather how does the presence and type of effect depend on the design of the policy.

## Design of school accountability systems and accessibility of negative information

The advantages of school accountability for studying the effects of policy design are that performance reporting is both ubiquitous and varied across states. The 2001 No Child Left Behind (NCLB) Act required all states to implement accountability systems in order to receive federal funding. Specifically, NCLB required that states annually assess the performance of public school students and release aggregate test results for schools and districts to the public. However, NCLB left considerable flexibility to states to set academic standards, to choose test instruments, to generate composite measures of school and district performance, and to determine how to report these measures to the public (Rhodes 2012). The 2015 Every Student Succeed Act largely renewed these aspects.

As a result, states vary in how they report the performance of public schools and districts. Most states use an ordinal rating scale to describe the overall quality of a school or district, often based on how well students score on state standardized tests. For example, California rated schools on a ten-point scale, while Texas initially used a four

<sup>3</sup>See Zhang et al. (2022) for a review of literature on the EDM.

point scale with the labels “unacceptable,” “acceptable,” “recognized,” and “exemplary.” Twenty states adopted systems that assign an A–F grade to schools and/or districts – the same scale commonly used to grade students.<sup>4</sup> Florida was the first state to adopt a letter grade system in 1999. Fifteen states followed between 2010 and 2013, and four others have done so since.<sup>5</sup> Supporters of letter grade systems often invoke accessibility as justification for use even as critics argue these grades oversimplify and, therefore, misrepresent school performance. For example, Jeb Bush, the Republican governor of Florida, argued: “They should not have to struggle through confusing mazes of charts and spreadsheets to find out if their children are in a good learning environment. To get there, we begin with a *simple*, comprehensive, actionable score that captures the overall success of a school in advancing academic achievement. The *most intuitive* approach for parents is grading schools on an A–F scale” (Bush, Hough, and Kirst 2017).<sup>6</sup> When Louisiana issued its first letter grades in 2011, the state Department of Education argued the switch would, “provide communities and families with a *clear and meaningful* depiction of school performance.”<sup>7</sup> South Carolina’s Department of Education described the system as “*simple and easy to understand*” when first adopting its use there.<sup>8</sup> In West Virginia, a Democratic governor urging adoption of letter grades argued, “This is a *transparent* education accountability system that rates student progress and performance in every West Virginia school using *language that parents and the community can understand*.”<sup>9</sup> When Michigan lawmakers repealed their letter grade system, defenders described letter grades as “*easily understandable* measures of school quality” (Thiel 2023). Organizations advocating for letter grade systems continue to argue, “Parents *understand* A-F grades” (Gergens 2024).

We argue that accessibility has consequences for how people process information. Public familiarity with the structure and connotation of a letter grade leaves individuals more susceptible to negative information because they more easily attach valence and the significance of that valence to an F grade than they might to similar information conveyed through another rating format. Consequently, there will be an asymmetric impact of this information whereby lower grades have a stronger effect on beliefs about the quality of schools than higher grades.

## Identifying effects of policy design

Evidence in support of our argument that policy design can yield asymmetric effects on evaluations of quality must go beyond simply demonstrating that individuals

<sup>4</sup>These are Alabama, Arizona, Arkansas, Florida, Georgia, Indiana, Louisiana, Maine, Michigan, Mississippi, New Mexico, North Carolina, Ohio, Oklahoma, South Carolina, Tennessee, Texas, Utah, Virginia, and West Virginia. Nine of these states have since repealed their letter grade system. Three others have indefinitely delayed grades for the most recent school year (2022–2023), but the policy remains officially in place.

<sup>5</sup>Sixteen states adopted via legislative approval. In three states, governors directed state departments of education to implement letter grades. In one state, a governor- and legislature-appointed board adopted the system.

<sup>6</sup>Italics added to quotes in this paragraph.

<sup>7</sup>Press release available at: <https://www.louisianabelieves.com/newsroom/news-releases/2011/10/05/historical-release-of-letter-grades-underscore-state-s-education-reform-agenda.0>

<sup>8</sup>In the department’s 2012 power point presentation describing the new system available here: <https://ed.sc.gov/data/report-cards/federal-accountability/esea/2012/>

<sup>9</sup>Press release available at: <https://wvde.state.wv.us/news/2980/>

exposed to information about public schools evaluate them more negatively. There are three additional challenges not yet addressed in extant research. First, demonstrating an asymmetric effect requires evidence that the effect of a negative signal is stronger than the effect of a positive signal. Evidence for an overall negative effect of exposure to information on opinions is insufficient because such an average effect could result from a skew in the supply of valence even with unbiased processing, for example, if there are far more F-rated schools or districts than A-rated schools or districts. Our argument is about the interplay between policy design and information processing – an effect that goes beyond the content of the information itself. Evidence in support of our theoretical argument requires demonstrating the effect is due to processing rather than supply. Second, identifying the effect of policy design requires demonstrating that differences between those exposed to letter grades and those not exposed to letter grades are not just a general effect of exposure to *any* information. The appropriate counterfactual is not “no information” but equivalent information presented under alternate policy designs. Finally, to identify whether asymmetries arise from negativity bias or expectation disconfirmation, we need to demonstrate that negative shifts occur even when reported information does not contradict prior beliefs.

To address these challenges, we use multiple studies each with advantages and limitations. We begin with a survey experiment to test the effect of letter grade ratings in a realistic format in one state. Next, in the same state, we use a difference-in-difference approach with pooled-cross-section surveys collected over several years straddling the transition to a letter grade system to examine how evaluations differ for individuals who live in comparably performing school districts when the state changed from an alternate rating system. Then, using national surveys pooled across time in conjunction with district-level test score data scaled for comparisons across states, we compare individuals exposed to letter grades to individuals exposed to alternate systems who live in school districts with similar average test scores. Finally, we turn to panel data collected within those national surveys to examine individual-level change in evaluations of public schools when states adopt letter grade reporting systems. The panel data also allow us to test for negativity bias versus the EDM by conditioning individual-level shifts on prior beliefs about school quality.

### Study 1: analysis of a survey experiment

We begin with analysis of experimental and observational studies conducted in Louisiana. Louisiana is a useful case for examining the effects of letter grade systems with a survey experiment because, with the exception of five city-based districts, public school districts in Louisiana are county-based,<sup>10</sup> which facilitates linking respondents to actual district-level school accountability information for the purpose of randomizing exposure within the survey. The ability to incorporate performance information directly into survey experiments allows us to identify effects on a key attitudinal outcome: individuals’ evaluations of schools. Using actual performance data tailored to the survey participant’s local school district provides a more direct test of information about local public schools *in the actual form the state provides it* on these evaluations than prior survey-based experiments that rely on

<sup>10</sup>Louisiana uses the term “parish” instead of “county.” We use “county,” the term more familiar outside the state.

state-level performance information (e.g., Clinton and Grissom 2015), hypothetical schools (e.g., Jacobsen, Snyder, and Saultz 2014), or stylized presentation formats not used in practice (e.g., Barrows et al. 2016).

To identify the effect of a letter grade system, we embed an experiment in a telephone survey of adult state residents.<sup>11</sup> We ask participants for their county of residence and, for participants living in one of the three counties that also contain the five city-based school districts, their city of residence. This geographic information identifies in which of the state's 69 public school districts a participant resides. Participants are randomly assigned to one of the two conditions. In one condition, we expose participants to the state's letter grade rating of their local school district. Specifically, participants are told, "As you may know, each year the Louisiana Department of Education grades each local public school district in the state. The state of Louisiana assigned a letter grade of [*insert participant's school district's grade*] to your school district."<sup>12</sup> The treatment information used the most recent district letter grade available at the time. Among participants in the treatment condition, 14.4% were informed the state issued their district an A grade, 44.4% were informed the state issued their district a B grade, 38.4% were informed the state issued their district a C grade, and 2.9% were informed the state issued their district a D grade. The state rated no districts as F for the 2014–2015 academic year.<sup>13</sup> Immediately after exposure to this information, participants are asked, "What grade would you give to the public schools in your local community?"<sup>14</sup> In the other condition, participants are simply asked the evaluation question without first being told the state's rating of their district.

Importantly, because they received the actual district grade issued by the state, individuals in the treatment condition did not all receive the *same* information. Therefore, an average treatment effect of exposure to grades across the grades actually provided cannot demonstrate asymmetry in processing because the direction of the

---

<sup>11</sup>The telephone survey was in February 2016. The survey includes 1,001 participants. The survey used live-interviewers, random digit dialing, and dual-frame samples including both landline and cellular telephones. The sample was weighted to the race, gender, age, household income, and educational attainment demographics for the state available in the *American Community Survey*.

<sup>12</sup>In the 2016 experiment, 25 participants did not provide a county and are not included in the analysis of treatment effect by actual letter grade below.

<sup>13</sup>The distribution of participants across district letter grades in the treatment condition is balanced with the distribution for participants in the no-information condition and similar to the population-weighted distribution of district letter grades for the state. See Table A1 in the [Supplementary Material](#) for the distributions of grades.

<sup>14</sup>Although this question is the standard measure for evaluations of public school quality at the local level, in use since at least 1974 (e.g., Phi Delta Kappan 2017), it creates a potential mismatch in the minds of participants between the object of this question ("public schools in your local community") and the object of the information provided in the exposure condition ("your school district"). In a subsequent similar general population survey of Louisiana residents, we ask participants to grade "public schools in your local school district" but without randomly assigning any to receive the letter grade rating assigned to the district. We find that the average response is slightly lower when using the second phrase (2.13 versus 2.35 on a four-point scale from 0 to 4) but the difference is not statistically significant. Further, such a difference makes for a conservative test of the effect of letter grade information on evaluations of "public schools in your local community" because if participants truly distinguish between public schools in their community and their school district, then it would be easier for them to discount information about the latter when evaluating the former, thus, making it more difficult to identify a causal effect of information based on their district.

effect may reflect the direction of the information provided. We break out the results of the experiment by the value of the grade issued to participants' district. Some live in an A-district, but most (85.6%) live in a district with a lower grade. Even unbiased processing would yield fewer A grade evaluations when the balance of the information supply tilts to lower grades. On the other hand, unbiased processing should also reduce the share of individuals who grade the quality of their local public schools with a D or F grade because most individuals in the treatment condition (97.1%) were exposed to district grades higher than a D. In short, with this distribution of actual district grades, unbiased processing should pull individuals' evaluations toward B and C, but that is not what we find.

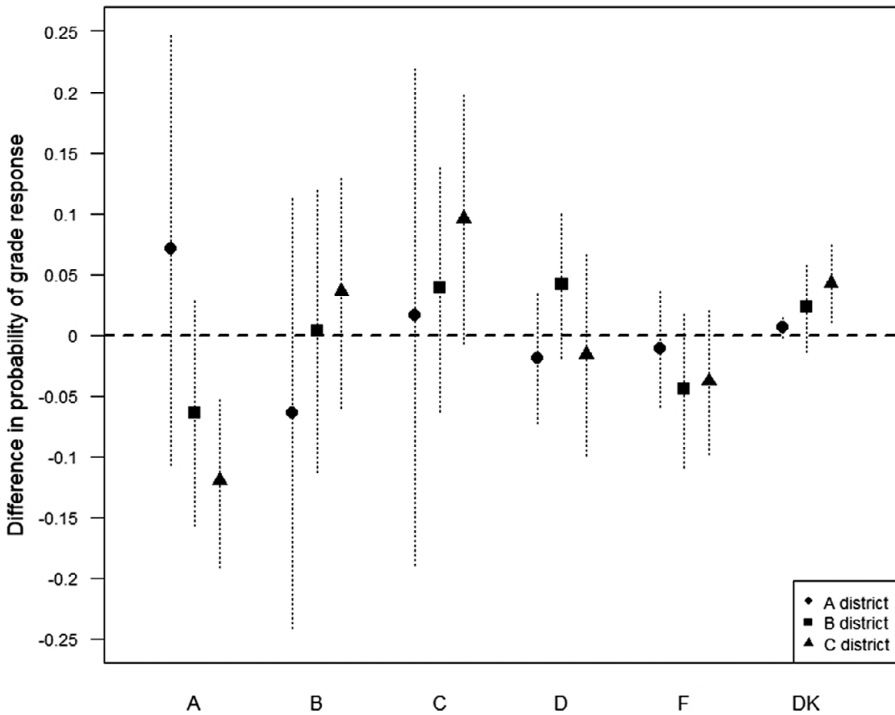
Instead, we find asymmetry – that is, greater responsiveness to negative than positive information. We test for asymmetry by conditioning treatment effects on the actual state-issued grades of districts in which participants reside. We separately test among individuals who live in districts graded as A, districts graded as B, and districts graded as C.<sup>15</sup> This approach allows us to identify the effect of exposure to a *particular* letter grade rather than the effect of exposure to *any* letter grade. More importantly, it allows us to distinguish between the effects of exposure to a district grade of A (positive signal) and exposure to a district grade of C (relatively negative signal).<sup>16</sup> These results appear in [Figure 1](#), which shows the treatment effect of providing individuals with the letter grade their school district actually received on their own evaluations of the quality of local public schools broken out by the value of the letter grade they received. Specifically, the figure displays the effect of exposure to the district's letter grade on the probability that a participant grades the quality of her local public schools as A, B, C, D, or F, as well as the probability she voluntarily indicates she is unsure (an option not explicitly read to participants). These response options appear on the horizontal axis.<sup>17</sup> There is no evidence that an A grade improves evaluations of local public schools. While the estimates of the effect of exposure to a district grade of A on the probability individuals evaluate their local schools with an A or B are noisy, of critical importance is the more precise evidence that a district grade of A has no effect on the likelihood of evaluating local schools with a D or F. In short, telling someone that she lives in a district the state grades with an A does not make her any less likely to evaluate her local schools with a grade of D or F. However, there is evidence that the more negative signal of a C grade does affect evaluations. Exposure to a district grade of C decreases the probability of evaluating local public schools with an A without also reducing the probability of evaluating

<sup>15</sup>We do not estimate treatment effects for individuals who live in districts graded as D because fewer than five percent of participants reside in such districts. Similarly, we do not estimate effects for exposure to an F because no such grades were issued that year.

<sup>16</sup>To validate that a C is perceived as a (relatively) negative signal, we embedded an additional experiment in another survey of state residents administered by YouGov. Individuals were exposed to a vignette about a hypothetical school district from another state including a letter grade from that state's accountability system and asked to rate the valence of the information on a five-point scale from "very positive" to "very negative." The content of the vignette remained fixed across participants except for the grade, which was randomized across values of A, B, C, D, and F. On average, individuals exposed to a grade of C rated the information as more negative than individuals exposed to a grade of A or B but not quite as negative as those exposed to a grade of D or F.

<sup>17</sup>Probability estimates in [Figure 1](#) are based on results of a multinomial logit model of responses. Model estimates appear in [Supplementary Table A2](#), and the distributions of evaluations of local public schools by experimental condition appear in [Supplementary Table A3](#).





**Figure 1.** Effect of exposure to state-issued district letter grade on respondents' own letter grade evaluations of local public schools, by state-issued grade. *Note:* Horizontal axis displays the response options for participants to grade the quality of public schools in their local community. Points mark the difference in the probability of each response by exposure to the grade the state issued to participants' local school districts by residence in districts with various state-issued grades. Dashed lines represent 95% confidence intervals. Full model results are available in [Supplementary Table A2](#).

schools without a D or F.<sup>18</sup> In short, participants do not respond to positive information but do respond to negative information.<sup>19</sup>

<sup>18</sup>These results are not an artifact of ceiling effects (i.e., individuals in school districts with an A grade already evaluating their local public schools with an A). Although we do not observe evaluations of public schools among the treatment condition prior to exposure to district grades, individuals in the control condition reveal the distribution of evaluations within districts receiving each grade from the state. Among residents of school districts graded with an A, 70% of individuals evaluate their local public schools with a grade lower than an A. This leaves significant room for improvement in their evaluations when exposed to state-issued grades. Additionally, 31% of residents in B districts and 29% of residents in C districts evaluate the quality of public schools in their local communities with grades lower than that issued by the state. Yet, there is no evidence that exposure to any of these grades issued by the state improves their evaluations. See Table A4 in the [Supplementary Material](#) for the distribution of evaluations by district grade among the control condition.

<sup>19</sup>In the next section, we use an observational study in Louisiana to demonstrate that this asymmetric effect follows from policy design, that is, from providing letter grades rather than simply from providing *any* information. However, we also conducted a second survey experiment to confirm that the results reported in [Figure 1](#) do not manifest when respondents are provided with alternate reporting formats for school quality (i.e., rankings). See the [Supplementary Material](#) for a discussion of that experiment.

## Study 2: analysis of a series of state surveys

Whereas our first study demonstrates asymmetric effects of letter grades, we now examine the role of policy design in these effects. That is, do these asymmetric effects occur in other reporting systems that do not use letter grades? We turn to an observational analysis of survey data collected under two different information formats, also in Louisiana. The shift between formats is a second advantage of using Louisiana to study the effects of letter grades because this shift occurred amid a series of annual statewide surveys. In the current system, Louisiana assigns a letter grade to school districts based on its District Performance Score (DPS).<sup>20</sup> From the 2002 to 2003 academic year through the 2009 to 2010 academic year, however, the state used an alternate rating system for districts. In that period, a district's DPS determined its rating on a six-point scale: Academic Warning/Academically Unacceptable; One Star; Two Stars; Three Stars; Four Stars; or Five Stars.<sup>21</sup> The annual survey data permit a within-state analysis of opinion under alternate reporting systems, which is more advantageous for identifying the effects of the letter grade system than comparing different states because states may vary on unobserved dimensions that correlate both with selection of a letter grade system and public opinion about school quality.

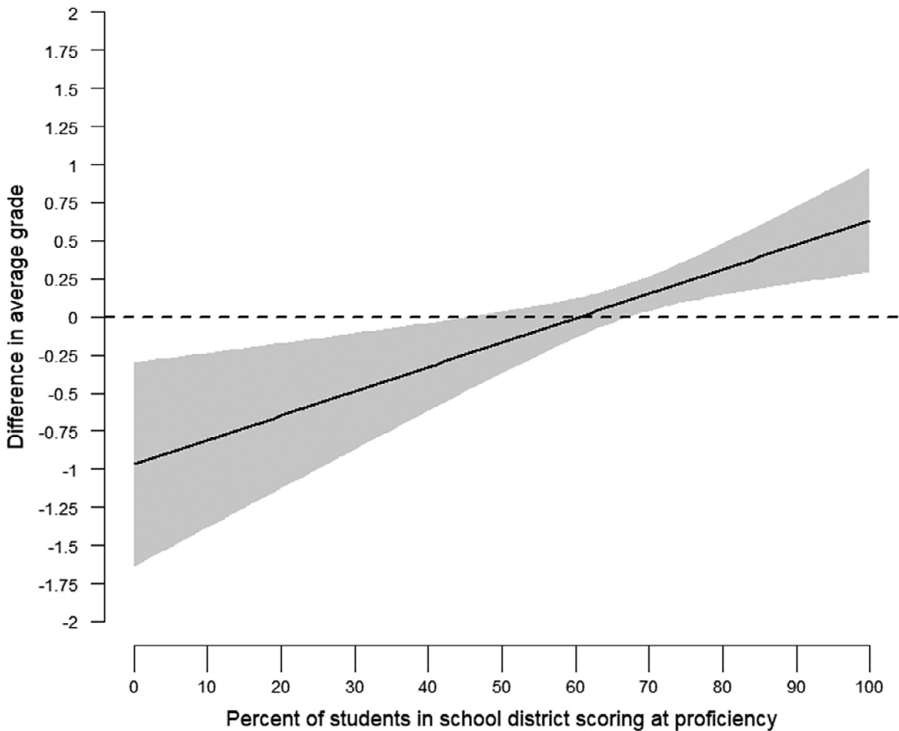
In this section, we report the results of an analysis of survey data collected under the current letter grade system and under the earlier system that assigned star ratings. The data are from an annual telephone survey administered to samples of adult Louisiana residents since 2003. On seven occasions during this period, the survey asked participants to evaluate the quality of their local public schools using the question described above: 2004, 2007, 2008, and 2014 through 2017.<sup>22</sup> The first three occurred during the period when the state used a six-point rating system to evaluate school districts from "academically unacceptable" to "five stars." The latter four surveys occurred during the letter grade system. Unlike the survey experiments, our second approach to examining letter grade systems lacks random assignment of exposure to ratings. Nevertheless, this analysis has the advantage of comparing opinion between two rating systems that Louisiana actually used.

Across these two periods, the valence of the information shifted in a positive direction. In the star-rating era, several districts were rated as "academically unacceptable," but none were rated at the highest two levels. In contrast, during the letter grade era, very few districts received F grades and many received A grades. Again, identifying asymmetric effects requires examining differences in responsiveness to negative information versus positive information. We pool participants across the surveys and model mean response as a function of an indicator for the years during which the state used a letter grade system, the percent of fourth- and eighth-grade students in participants' districts who scored at grade level on the state's standardized reading and mathematics tests in the academic year preceding the survey, and the interaction of these two variables. Additionally, we include school district

<sup>20</sup>More details on the DPS calculation can be found in the [Supplementary Material](#).

<sup>21</sup>The state used the term "Academic Warning" to describe the lowest rating during the first two years of this system (2002–2003 and 2003–2004). After that, until adoption of the letter grade system with the 2010–2011 academic year, the state used "Academically Unacceptable."

<sup>22</sup>The survey experiments described in the previous section were embedded in the 2016 and 2017 iterations of this survey. Therefore, the analysis in this section uses only those participants in the no-information conditions from those years.



**Figure 2.** Difference in average opinions of local school quality between letter grade era and star-rating era, by percent of students in school district scoring at grade level proficiency. *Note:* Solid line represents the difference in the average evaluation of local public schools on a scale of 1 (F) to 5 (A) between participants during the star-rating era and participants during the letter grade era. The difference is displayed by the percent of students in the participant's local school district who scored at grade level or above on standardized exams. Shaded area represents 95% confidence intervals. Full model results are available in [Supplementary Table A6](#).

fixed effects.<sup>23</sup> This approach compares individuals whose districts had similar test scores, which largely drove rating assignment under both systems, but who saw those scores presented in different formats.

Figure 2 displays the difference in average ratings between the letter grade system period and the star rating period across the values of the percent of students testing at grade level. If the two rating systems do not lead to different interpretations of school quality, the line should be flat. Instead, there is a sharply positive slope, indicating people process the same underlying data (i.e., the percent of students scoring at grade level) differently between the two rating regimes. Districts with larger shares of students performing at grade level are evaluated more positively in the letter grade

<sup>23</sup>We also include controls for race, household income, gender, age, educational attainment, and partisanship of the participant. Additionally, we analyzed a similar model that included three school district characteristics rather than district fixed effects: racial demographics, economic demographics, and size of participants' school districts. The alternate specification yields similar results. Estimates for both models appear in [Supplementary Table A6](#).

system than the star rating system, and districts with smaller shares of students performing at grade level are judged more harshly. Importantly, there is an asymmetry in these differences – the gap is larger for districts with lower test scores than for districts with higher test scores. In other words, individuals penalize schools with lower test scores more than they reward schools with higher test scores when provided a letter grade rather than a star rating. For example, among residents of school districts where approximately 40% of students score at grade level across both periods, resulting in an “academically unacceptable” rating under the earlier period but a D grade in the later period, the estimated difference in mean evaluations on the five-point response scale to the survey item is  $-0.41$  (34.5% of a standard deviation). For residents of school districts where 75% of students scored at grade level, which generally earned a rating of three stars under the earlier system but an A grade under the later system, the estimated difference in opinion is  $+0.23$  (19.3% of a standard deviation). Together, the experimental and observational studies from Louisiana demonstrate that asymmetric effects, whereby negative signals have a stronger effect than positive signals, are not automatic or ubiquitous; rather, they depend upon how states choose to present information about school district performance.

### Study 3: analysis of national cross sections

Next, we turn to national survey data to examine whether perceptions of schools vary between accountability systems when comparing across states that use letter grades for schools or do not use grades. We use the series of *Education Next* surveys conducted annually since 2007.<sup>24</sup> The survey has asked respondents to evaluate the quality of local schools annually, with the exception of 2009 and 2010, with the same question analyzed in the previous sections of this article.<sup>25</sup> Annual samples frequently reach more than 4,000 participants, which we pool to estimate differences across states with different accountability policies. Of course, states differ in many ways that are related to how residents evaluate public schools, such as the academic performance of students or the political culture of the state, which may also relate to the decision to adopt letter grade systems. For instance, 14 of the 20 states that adopted letter grade systems did so under unified Republican control of state government (i.e., holding the governor’s office and majorities in both legislative chambers). Three more states adopted letter grades under Republican governors with Democratic-controlled legislatures. Only three states adopted letter grades under a Democratic governor, including just one that did so under unified Democratic control of state government.<sup>26</sup> We take two approaches to account for these differences. First, to account for characteristics that may correlate with the likelihood a state adopts a

<sup>24</sup>The annual online survey is designed by researchers at Harvard University’s Program for Educational Policy and Governance and conducted using a probability-based online panel (Ipsos KnowledgePanel©).

<sup>25</sup>The survey asks: “Students are often given the grades A, B, C, D, and Fail to denote the quality of their work. Suppose the public schools themselves were graded in the same way. What grade would you give the public schools in your community?”

<sup>26</sup>Party differences are less clear when it comes to repealing letter grade systems. Michigan and New Mexico repealed their plans after Democrats won the governor’s office and majorities in the legislature. When a Democratic governor won office in Virginia, he convinced the Republican-controlled legislature to drop the system. West Virginia also dropped its letter grade system under a Democratic governor after just one year; but he was the same governor who had originally directed the state department of education to adopt the

letter grade system, we limit our sample to only respondents who live in states that eventually adopted a letter grade system. Because we have multiple surveys over time, we can compare responses in states that have implemented a letter grade system to responses in states that would ultimately adopt a letter grade system but had not done so at the time the of the survey in which the responses appear. Second, we condition responses on a measure of student test scores in respondents' school districts. This approach requires a common metric for scores across school districts. We use the Stanford Education Data Archive (SEDA) of the Center for Education Policy Analysis at Stanford University (Reardon et al. 2017). SEDA includes a measure of district-level scores that adjusts state test results to NAEP test results, essentially providing average NAEP score for each district. For nine years of the *Education Next* survey, SEDA data are available in the academic years immediately preceding the survey: 2011 through 2019.

We model responses to the school quality question as a function of an indicator for whether the participant's state used letter grade ratings for the academic year immediately preceding the survey year while controlling for test scores.<sup>27</sup> This approach amounts to comparing participants' evaluations of their local schools in states with letter grade systems to participants' evaluations in states without a letter grade system (but would eventually adopt one) and *who also live in districts with the same level of student performance on standardized tests*. During this period, states emphasized the *levels* of scores on standardized tests (e.g., average score or percent of students who achieved a certain score) in determining what rating a school or district received in the accountability system. Because of this emphasis, our approach amounts to comparing respondents whose state accountability systems take a similar "input" from the respondents' local schools (i.e., similar test scores) but "output" the rating in different formats (i.e., a letter grade versus something else).<sup>28</sup> In a sense, the underlying information these two individuals receive is similar, but the package – a letter grade versus something else – is different.

Results appear in Figure 3. Individuals in states with letter grade rating systems have worse opinions of their local public schools than individuals in other states. These individuals are less likely to think of their local public schools as deserving of an A or B, and more likely to think they merit a lower grade.<sup>29</sup>

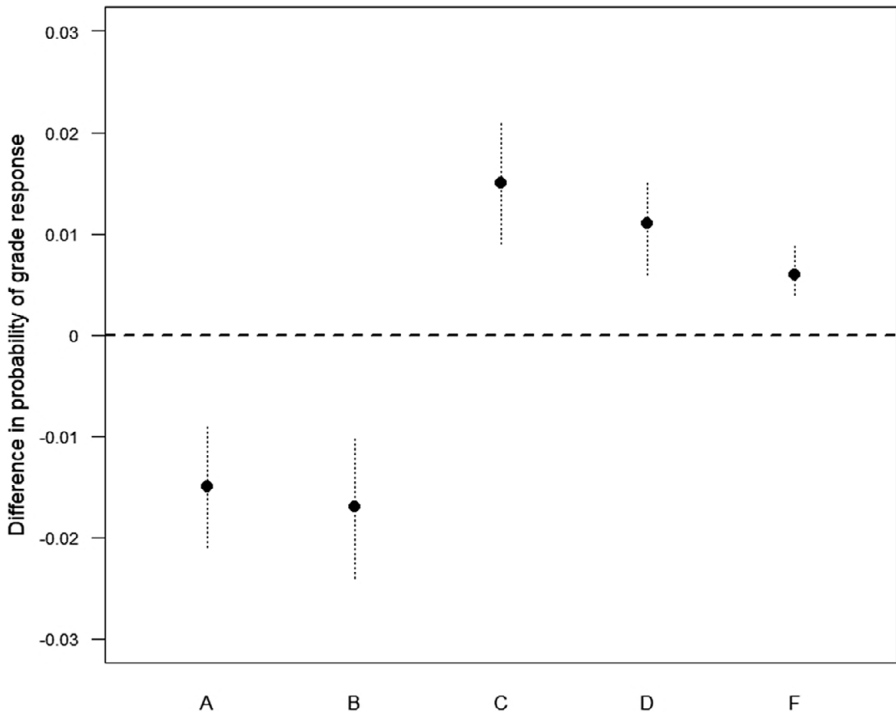
---

system. Five states dropped their letter grade systems under unified Republican control of state government: Indiana, Maine, Ohio, South Carolina, and Utah.

<sup>27</sup>We also control for participant race/ethnicity, gender, household income, age, education, party identification, an indicator for being a parent of school age children, and year of interview. Additionally, in 2013 and 2014, the surveys included an experiment providing some participants (randomly selected) with information about student performance in their local districts (Barrows et al. 2016). We include a control for assignment to an information condition.

<sup>28</sup>To be clear, we do not argue that by comparing across districts with similar test scores we control for school or district quality. Nor, do we argue that test scores measure school or district quality.

<sup>29</sup>The full results of the models underlying Figures 3 and 4 are available in Supplementary Table A7. Of note, the other covariates in these models follow patterns common in the literature (e.g., Moe 2001; Peterson, Henderson, and West 2014). For example, individuals with higher socioeconomic status, measured as household income or educational attainment, evaluate local schools more positively, likely reflecting geographic sorting. Parents with school-age children in the home evaluate schools more positively than other adults do. Democrats rate their local public school more positively than Republicans do. This last result contrasts with earlier work that found no partisan gap in opinions about local school quality using survey data



**Figure 3.** Differences in public evaluations of local schools between individuals in states with letter grade systems and individuals in states with other systems. *Note:* Horizontal axis displays the response options for participants to grade the quality of public schools in their local community. Points mark the difference in the probability of each response between individuals living in states with letter grade rating systems and individuals living in states with other systems. Dashed lines represent 95% confidence intervals. Full model results are available in [Supplementary Table A7](#).

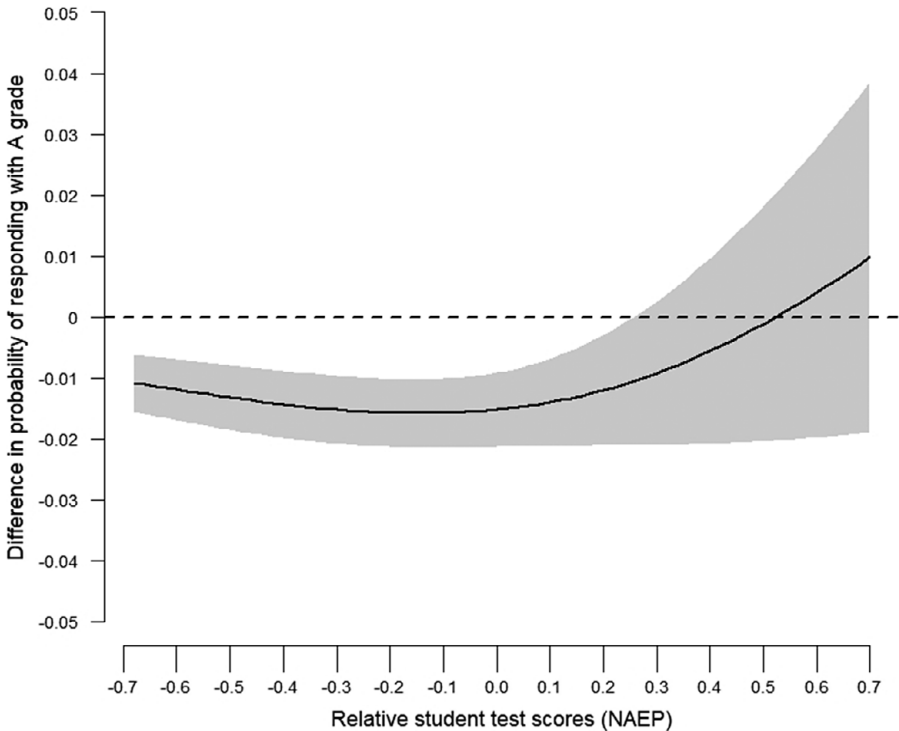
Additionally, in [Figure 4](#), we plot the difference in the probability that individuals grade their local public schools with an A between those in states with and without letter grade rating systems by the average relative NAEP score in their local public school district. Once again, the results demonstrate mostly negative shifts in evaluations. Individuals living in states with letter grade systems rate their local schools worse than individuals in states without these systems but who live in districts with comparable test scores; this pattern holds for respondents up through the eightieth percentile of district-level NAEP scores in the sample. Only in the 81st percentile of scores do estimates become too noisy to distinguish from the null; and the point estimate for the difference does not reach zero until respondents are in the 96th percentile of scores.

#### Study 4: analysis of national panel data

From 2013 through 2018, the *Education Next* surveys included a subset of individuals from the previous year's sample producing a panel of participants observed on

---

from 2008 to 2011 (Holbrook and Weinschenk 2020). However, it is consistent with recent evidence showing a widening gap between partisans (Houston, Peterson, and West 2023; Peterson, Henderson, and West 2014).



**Figure 4.** Difference in probability of grading local schools as A, between individuals in states with letter grade systems and individuals in states with other systems by mean NAEP score in district. *Note:* Solid curves represent the difference in the probability that participants rate public schools in their local community with a grade of A between individuals living in states with letter grade rating systems and individuals living in states with other systems, by relative NAEP score in the local school district (i.e., how the district compares to districts across the country). Shaded area represents 95% confidence intervals. Full model results are available in [Supplementary Table A7](#).

multiple occasions across years. The panel sample allows us to estimate models of opinion with individual-level fixed effects to control for all unobserved time-invariant characteristics. The approach requires use of a linear model to estimate average response on the five-point scale from F to A rather than the ordinal approach used above (Angrist and Pischke 2009). In all, we analyze responses of 1,776 participants interviewed at least twice from 2013 to 2018.<sup>30</sup> This approach estimates within-individual change in opinion when a respondent's state transitions from a non-letter grade system to a letter grade system.<sup>31</sup> The results in the first column of [Table 1](#) indicate respondents' evaluations of their local public schools sour by about  $-0.10$  (on a five-point scale) when their state switches to letter grades.

The evidence presented so far indicates people rate schools lower when those schools receive letter grades, but this evidence is consistent with two different ways

<sup>30</sup>We exclude a handful of panelists who moved between states.

<sup>31</sup>We control for year and whether the respondent was randomly assigned to exposure to information about local district performance in 2013 or 2014.

**Table 1.** Fixed effects model of letter grade rating system effect on evaluations of public schools

	Response (1–5)	
Letter grade system	–0.102*** [0.032]	
Letter grade <i>below</i> prior perception		–0.161*** [0.036]
Letter grade <i>same</i> as prior perception		–0.120** [0.053]
Letter grade <i>above</i> prior perception		0.124** [0.062]
2014	0.031 [0.025]	0.051* [0.031]
2015	0.239*** [0.035]	0.272*** [0.042]
2016	0.206*** [0.037]	0.247*** [0.044]
2017	0.226*** [0.037]	0.277*** [0.044]
2018	0.205*** [0.039]	0.255*** [0.046]
Exposure to information	0.002 [0.029]	0.013 [0.034]
Average NAEP score in district	0.198*** [0.076]	0.180** [0.084]
Constant	3.318*** [0.032]	3.280*** [0.037]
Individual fixed effects?	Yes	Yes
Observations	8,100	6,917
Number of panelists	2,404	2,311
R-squared	0.057	0.019

Note: Estimates from OLS model of evaluations of public schools on scale from 1 (F) to 5 (A) including individual-level fixed effects. Baseline year is 2013. In 2013 and 2014, surveys included experiments in which randomly selected participants were exposed to information comparing test scores in their school district to other areas. “Exposure to information” is an indicator for assignment to one of these conditions in those two years of the survey. Standard errors in brackets.

\*\*\* $p < 0.01$ .

\*\* $p < 0.05$ .

\* $p < 0.1$ .

people may process school accountability information. One possibility is negativity bias – that is, that people respond to negative information more so than positive information and letter grade formats make it easier for people to identify negative signals. Another possibility is the EDM – that is, that people respond to information that contradicts their prior beliefs and letter grade formats make it easier for people to identify positive and negative signals. Because Americans tend to rate their local schools relatively positively, the negative effects of letter grades may simply reflect the distribution of these prior beliefs (relatively positive compared to the information received in the letter grade). In either case, letter grade systems would enhance a form of information processing that yields asymmetric negative effects on beliefs about school quality.

Nevertheless, it is worthwhile to attempt to disentangle these explanations. To do so, we consider additional expectations of the EDM that would not manifest if negativity bias alone was at work: (1) people would be just as responsive to positive grades if they had lower prior beliefs about their local schools’ quality; and (2) people would not respond to information consistent with their prior beliefs about quality. To test these, we leverage another feature of the panel data – respondents’ responses to



the previous year's survey. We collected the actual letter grade assigned to each district in each state that issued letter grades from 2013 through 2018.<sup>32</sup> We compared those grades to the respondents' own evaluation of their local school from the prior year and sorted them into three groups: Respondents whose local schools received a grade below their own prior perception (a negative signal), respondents whose local schools received the same grade as their own prior perception (an equivalent signal), and respondents whose local schools received a grade above their own prior perception (a positive signal).<sup>33</sup> If the EDM holds, we should see a negative effect of the negative signal, a positive effect of the positive signal, and no effect for the equivalent signal. If negativity bias is at work, we should see negative effects even among those who did not receive a negative signal.

We test this by replicating our fixed effects model but replacing the indicator for a letter grade system with three indicators for the signal type. Again, because this model includes individual fixed effects, we are estimating individual-level change in perceptions as someone goes from not having a letter grade system to receiving one of these three signals. The results appear in the final column of Table 1. Unsurprisingly, negative signals produce a negative shift in perceptions of quality. The results also reveal evidence both for the EDM and for negativity bias, suggesting both processes may be at work. Letter grades that are higher than respondents' prior beliefs about school quality are associated with a positive shift in beliefs about quality, consistent with the claim that people respond to information that disconfirms expectations. However, there is also a negative shift in opinions even for people who receive a grade equivalent to their prior beliefs. This latter result is consistent with negativity bias and inconsistent with expectation disconfirmation. It appears that merely packaging the signal in the form of a letter grade exerts a negative influence on perceptions of quality. These results are not entirely definitive and should not be interpreted to mean that the EDM fails to apply at all. Nevertheless, the results do indicate that EDM does not on its own account for the negative shifts in perceptions of quality.

## Discussion

Many Americans take a dim view of government, generally believing it to perform poorly. Although there is substantial evidence that the actual quality of public services shapes beliefs about government performance in specific sectors, there is compelling evidence that underlying partisan identities and ideological orientations bias people's beliefs about government performance (Holbrook & Weinschenk 2020; Lerman 2019). The former evidence – that people evaluate government in accordance with how well it actually performs – is a boon for theories of democratic accountability. The latter evidence – that people have built in inclinations to view government negatively independent of actual performance – challenges those theories. We

<sup>32</sup>We collected these grades from state departments of education. We used the most recent letter grades issued to districts at the time of the survey. For states that issued letter grades only to schools, we generated a district-level grade by averaging school-level grades in the district.

<sup>33</sup>The *Education Next* surveys do not include measures of information exposure, so we cannot directly examine whether respondents know the grade their local schools receive. However, the experiment in Louisiana includes a question asking respondents if they know the grade the state rated their district. This question was posed prior to the experiment and only to respondents in the control condition. While 70% said they did not know, 11% correctly identified the grade and another 17% knew within one grade.

contribute to this discussion by highlighting another factor beyond actual government performance and individuals' political predispositions – how governments describe their performance.

Our evidence indicates that letter grade systems have asymmetric effects, souring opinions of public schools relative to alternative systems. That is, using evidence from one state, we demonstrate a stronger impact of negative information than positive information under letter grade systems than under alternate presentation formats. Then, we generalize this effect across the United States by providing evidence that individuals in states with letter grade systems evaluate their local public schools more poorly, even after controlling for test scores and unobserved characteristics of individuals and states. Finally, we show that the negative effects are not confined only to those people whose local schools receive a grade worse than what they had thought – casting doubt on the EDM as an explanation for this pattern and indicating negativity bias.

To be clear, we do not claim that letter grade systems are the only means of invoking biases in people's judgments about public services. Nor, do we claim that this evidence indicates policymakers necessarily choose these systems for this reason. Nevertheless, the case of letter grade systems demonstrates the importance of decisions about policy design for the political consequences of information policies on opinion. Indeed, even as several states abandoned letter grades in recent years they are replacing them with other formats for public reporting on school quality. How might these new designs interact with cognitive processes to push opinions in one direction or another? There is no reason to think these findings are exclusive to letter grades or even to education.

The findings have two important implications for the study and practice of democracy. First, these results demonstrate that scholars should attend to features of policy design when assessing the impact of performance reporting. Prior empirical analysis of the effects of these policies has focused on a simplistic question: Do these policies influence knowledge, opinion, and behavior? Scholarship on these policies must go further to consider not only whether these policies matter but also the conditions under which they yield different effects – especially conditions of policy design.

Second, the results speak to the role of these policies in democratic practice. A naïve view holds that these policies have a benign effect, simply filling in gaps in knowledge so voters can make informed decisions. Our evidence indicates that these policies do not simply provide information; rather, they provide *packaged* information. People tend to process information with cognitive biases, and how information is presented can exacerbate these tendencies. Our evidence does not indicate that the evaluations people make of public schools under letter grade systems are any more correct or incorrect than those made under other systems, but the fact that opinions vary by policy design does indicate that policymakers have the power to shape opinions. Information provided by these policies is never just information; it is information designed and presented in a certain way.

Through their decisions about how to design informational policies, policymakers have the power not simply to inform public opinion but to steer it, and the political consequences can be quite widespread. Evaluations for public sector performance are related to confidence in government action (Wichowsky and Moynihan 2008), support for continued spending on programs (Peterson, Henderson, and West 2014), and, in the case of public schools, attitudes toward policy interventions to

overhaul the system through market-based initiatives (Moe 2001). Measuring policy performance is a complicated matter fraught with decision points that can steer voters' inferences about government performance. Our research suggests that the process is even more complicated. Accurately measuring government performance is not enough. The process of communicating information about performance plays a critical role in democratic accountability.

**Supplementary material.** The supplementary material for this article can be found at <http://doi.org/10.1017/spq.2024.19>.

**Data availability statement.** Replication materials are available on SPPQ Dataverse at <https://doi.org/10.15139/S3/NLDIQK> (Henderson and Davis 2024).

**Acknowledgments.** The authors thank Marty West, Matt Chingos, Charles Barrilleaux as well as participants in the Louisiana State University Department of Political Science research workshop for valuable comments. The authors also thank staff at the Louisiana Department of Education – especially Jessica Baghian, John White, and Jill Zimmerman – for assistance with state data on Louisiana public schools. Finally, the authors thank the Program on Education Policy and Governance at Harvard University and the Reilly Center for Media and Public Affairs at Louisiana State University for generously allowing them access to their survey data.

**Funding statement.** The authors received no financial support for the research, authorship, and/or publication of this article.

**Competing interest.** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## References

- Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Baekgaard, Martin, and Soren Serritzlew. 2015. "Interpreting Performance Information: Motivated Reasoning or Unbiased Comprehension." *Public Administration Review* 76(1): 73–82.
- Barrows, Samuel, Michael Henderson, Paul E. Peterson, and Martin R. West. 2016. "Relative Performance Information and Perceptions of Public Service Quality: Evidence from American School Districts." *Journal of Public Administration and Theory* 26(3): 571–83.
- Berry, Christopher R., and William G. Howell. 2007. "Accountability and Local Elections: Rethinking Retrospective Voting." *Journal of Politics* 69(3): 844–58.
- Bush, Jeb, Heather J. Hough, and Michael W. Kirst. 2017. "How Should States Design Their Accountability Systems?" *Education Next* 17(1): 54–62.
- Chingos, Matthew M., Michael Henderson, and Martin R. West. 2012. "Citizen Perceptions of Government Service Quality: Evidence from Public Schools" *Quarterly Journal of Political Science* 7(4): 411–45.
- Clinton, Joshua D., and Jason A. Grissom. 2015. "Public Information, Public Learning and Public Opinion: Democratic Accountability in Education Policy." *Journal of Public Policy* 35(3): 355–85.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Brothers.
- Fridkin, Kim Leslie, and Patrick J. Kenney. 2004. "Do Negative Messages Work? The Impact of Negativity on Citizens' Evaluations of Candidates." *American Politics Research* 32(5): 570–605.
- Gergens, Austin. 2024. "What's Behind the Fight over A-F School Grading?" *ChalkboardNews*. [https://www.chalkboardnews.com/issues/accountability/article\\_0147a0f4-a98d-11ee-9652-cb7b89dc891f.html](https://www.chalkboardnews.com/issues/accountability/article_0147a0f4-a98d-11ee-9652-cb7b89dc891f.html) (Accessed October 15, 2024).

- Hacker, Jacob S., Suzanne B. Mettler, and Dianne Pinderhughes. 2005. "Inequality and Public Policy." In *Inequality and American Democracy: What We Know and What We Need to Learn*, eds. Lawrence R. Jacobs and Theda Skocpol. New York: Russell Sage Foundation. 156–213.
- Hambleton, Ronald K., and Sharon C. Slater. 1995. *Are NAEP Executive Summary Reports Understandable to Policy Makers and Educators?* National Center for Research on Evaluation, Standards, and Student Testing. Technical Review Panel for Assessing the Validity of National Assessment of Educational Progress. Los Angeles, CA: Center for the Study of Evaluation.
- Henderson, Michael, and Davis, Belinda C. 2024. "Replication Data for: Making the Grade: Policy Design and Effects of Information about Government Performance." <https://doi.org/10.15139/S3/NLDIQK>, UNC Dataverse, V1, UNF:6:HqPG2qE6VHAinQc3MNTVmA== [fileUNF]
- Hjortskov, Morten. 2019. "Citizen Expectations and Satisfaction over Time: Findings from a Large Sample Panel Survey of Public School Parents in Denmark." *American Review of Public Administration* 49(3): 353–71.
- Holbein, John. 2016. "Left Behind? Citizen Responsiveness to Government Performance Information." *American Political Science Review* 110(2): 353–68.
- Holbrook, Thomas M., and Aaron C. Weinschenck. 2020. "Information, Political Bias, and Public Perceptions of Local Conditions in U.S. Cities." *Political Research Quarterly* 73(1): 221–36.
- Houston, David M., Paul E. Peterson, and Martin R. West. 2023. "Partisan Rifts Widen, Perceptions of School Quality Decline: Results from the 2022 Education Next Survey of Public Opinion." *Education Next* 23(1): 8–19.
- Ingram, Helen, and Anne Schneider. 1993. "Constructing Citizenship: The Subtle Messages of Policy Design." In *Public Policy for Democracy*, eds. Helen Ingram and Steven Rathgeb Smith. Washington, DC: Brookings Institution.
- Iyengar, Shanto. 1990. "Shortcuts to Political Knowledge: The Role of Selective Attention and Accessibility." In *Information and Democratic Processes*, eds. John A. Ferejohn and James H. Kuklinski. Urbana, IL: University of Illinois.
- Iyengar, Shanto. 1987. "Television News and Citizens' Expectations of National Affairs." *American Political Science Review* 81(3):815–32.
- Jacobsen, Rebecca, Jeffrey W. Snyder, and Andrew Saultz. 2014. "Informing or Shaping Public Opinion? The Influence of School Accountability Data Format on Public Perceptions of School Quality." *American Journal of Education* 121(1): 1–27.
- James, Oliver. 2011. "Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments." *Journal of Public Administration Research and Theory* 21(3): 399–418.
- James, Oliver, and Peter John. 2007. "Public Management at the Ballot Box: Performance Information and Electoral Support for Incumbent English Local Governments." *Journal of Public Administration Research and Theory* 17(4): 567–580.
- James, Oliver, and Gregg G. Van Ryzin. 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.
- Kanouse, David E., and L. Reid Hanson Jr. 1972. "Negativity in Evaluations." In *Attribution: Perceiving the Causes of Behavior*, eds. Edward E. Jones, David E. Kanouse, Harold H. Kelley, Richard E. Nisbett, Stuart Valins, and Bernard Weiner. Morristown, NJ: General Learning.
- Kernell, Samuel. 1977. "Presidential Popularity and Negative Voting." *American Political Science Review* 71(1): 44–66.
- Kogan, Vladimer, Stephane Lavertu, and Zachary Peskowitz. 2016a. "Performance Federalism and Local Democracy: Theory and Evidence from School Tax Referenda." *American Journal of Political Science* 60(2): 418–35.
- Kogan, Vladimer, Stephane Lavertu, and Zachary Peskowitz. 2016b. "Do School Report Cards Produce Accountability through the Ballot Box?" *Journal of Policy Analysis and Management* 35(3): 639–61.
- Lau, Richard R. 1985. "Two Explanations for Negativity Effects in Political Behavior." *American Journal of Political Science* 29(1): 119–38.
- Lerman, Amy E. 2019. *Good Enough for Government Work: The Public Reputation Crisis in America (And What We Can Do to Fix It)*. Chicago, IL: University of Chicago.
- Moe, Terry M. 2001. *Schools, Vouchers, and the American Public*. Washington, DC: Brookings Institution Press.

- Niven, David. 2000. "The Other Side of Optimism: High Expectations and the Rejection of Status Quo." *Political Behavior* 22(1): 71–88.
- Peterson, Paul E., Michael Henderson, and Martin R. West. 2014. *Teachers versus the Public: What Americans Think about Schools and How to Fix Them*. Washington, DC: Brookings Institution Press.
- Phi Delta Kappan. 2017. *The 49th Annual PDK Poll of the Public's Attitudes toward the Public Schools*. Arlington, VA: PDK International.
- Reardon, Sean F., Andrew D. Ho, Benjamin R. Shear, Erin M. Fahle, Demetra Kalogrides, and Richard DiSalvo. 2017. "Stanford Education Data Archive (Version 2.0)." <http://purl.stanford.edu/db586ns4974>. (accessed July 24, 2024).
- Rhodes, Jesse H. 2012. *An Education in Politics: The Origins and Evolution of No Child Left Behind*. Ithaca, NY: Cornell University Press.
- Rhodes, Jesse H. 2015. "Learning Citizenship? How State Education Reforms Affect Parents' Political Attitudes and Behavior." *Political Behavior* 37(1): 181–220.
- Soroka, Stuart N. 2006. "Good News and Bad News: Asymmetric Responses to Economic Information." *Journal of Politics* 68(2): 372–85.
- Thiel, Craig. 2023. "The School Accountability System Merry-Go-Round." Citizens Research Council of Michigan. Lansing, MI.
- Van de Walle, Steven, and Alasdair Roberts. 2011. "Publishing Performance Information: An Illusion of Control?" In *Performance Information in the Public Sector: How It Is Used*, eds. Wouter Van Dooren and Steven Van de Walle. Basingstroke, UK: Palgrave Macmillan.
- Van Ryzin, Gregg G. and Cecilia F. Lavena. 2013. "The Credibility of Government Performance Reporting: An Experimental Test." *Public Performance and Management Review* 37(1): 87–103.
- Wichowsky, Amber, and Donald P. Moynihan. 2008. "Measuring How Administration Shapes Citizenship: A Policy Feedback Perspective on Performance Management." *Public Administration Review* 68(5): 908–20.
- Zhang, Jiasheng, Wenna Chen, Nicolai Petrovsky, and Richard M. Walker. 2022. "The Expectancy - Disconfirmation Model and Citizen Satisfaction with Public Services: A Meta-analysis and an Agenda for Best Practice." *Public Administration Review* 82(1): 147–59.

**Author biographies.** Michael Henderson is an associate professor of Political Science and Mass Communication at the Louisiana State University.

Belinda C. Davis is an associate professor of Political Science at the Louisiana State University.

---

**Cite this article:** Henderson, Michael, and Belinda C. Davis. 2024. Making the Grade: Policy Design and Effects of Information about Government Performance. *State Politics & Policy Quarterly* 1–21, doi:10.1017/spq.2024.19