CITIZENS' MOTIVATED REASONING ABOUT PUBLIC PERFORMANCE: EXPERIMENTAL FINDINGS FROM THE US AND DENMARK

MARTIN BAEKGAARD, AARHUS UNIVERSITY

OLIVER JAMES, UNIVERSITY OF EXETER

SØREN SERRITZLEW, AARHUS UNIVERSITY

GREGG G. VAN RYZIN, RUTGERS UNIVERSITY

ABSTRACT: Recent studies find motivated reasoning in citizens' processing of information about public performance. Using experiments in the US and Denmark, we examine effects on an accuracy-based task of two forms of motivated reasoning: partisan identity-based reasoning and reasoning from ideology-based governance preferences (favoring either the public or the private sector). The experiments incorporate a political prime, a health care needs prime (to reduce politicization), and a neutral, no-prime, condition. We find that priming citizens to think politically accentuates the influence of partisan identities and governance preferences on reasoning. In contrast, priming about the need for a service reduces these biases. These findings extend knowledge of motivated reasoning in an accuracy-based task and priming with a noprime benchmark, and confirm some findings of previous studies. Reducing the salience of partisan identities or governance preferences in the presentation of information may help stimulate more accuracy-based reasoning about public performance.

INTRODUCTION

The performance movement in government rests in part on the view that providing

citizens with factual information about the performance of public programs

and services will facilitate an informed citizenry, enhance democratic accountability,

and build public trust in government (Boyne et al. 2003; Heinrich 2003; Hatry

2006). However, recent studies from Denmark suggest that citizens often infer false

conclusions from unambiguous information about the performance of public and

private agencies, depending on their prior preferences for public over private service

provision (Baekgaard and Serritzlew 2016; Baekgaard et al. 2019). This is an

example of motivated reasoning, which refers to the role of motivations or goals in

shaping how people attend to and process information. In a US study James and

Van Ryzin (2017b) found that citizens, when primed to think politically rather

than about their need for health care services, exhibited more partisan bias in their

choice of performance indicators and ratings of the strength of evidence about the

performance of the Affordable Care Act (ACA), also known as Obamacare. Taken

together, these studies from different country contexts suggest that citizens are

prone to selective use, differential rating of evidence strength, and errors in processing

performance information in the presence of ideological or political cues. The

previous studies are valuable in establishing evidence of these two forms of motivated

reasoning but employed different designs and thus leave important questions

about the generalizability of the observed effects unanswered.

Specifically, the effects of priming people to think politically or about their need for

a service when undertaking an accuracy-based use of performance information are

unknown. The James and Van Ryzin (2017b) study established that priming people to

think nonpolitically, by making them focus on their need for a public service (health

care), reduced the importance of motivated reasoning as measured by respondents'

choice of performance indicators and their ratings of evidence strength. The current

2

study asks if this effect extends further to an accuracy-based task where there is clearly a

"correct" answer to a question using provided information. This task was used in the

Baekgaard and Serritzlew (2016) study and is an important measure of motivated reasoning

because it unambiguously shows whether people arrive at a correct decision or

not (Kahan 2013). Additionally, the current study uses a neutral, no-prime condition in

addition to the political prime and health care needs primes used by James and Van

Ryzin (2017b). Including a no-prime condition helps clarify whether the political prime

triggers directional motives (bias) or, in contrast, the health care needs prime promotes

accuracy motives, extending knowledge through comparisons with a neutral benchmark

condition. Finally, in addition to the accuracy-based task used by Baekgaard and

Serritzlew (2016), the current study asks participants to assess the strength of evidence

after undertaking the accuracy-based task. This measure of evidence strength provides

insight into participants' views of the information presented to them and, as such, represents

an indicator of how motivated reasoning operates as a cognitive process.

We begin with an overview of the theory and current evidence about motivated

reasoning, in particular with respect to performance information. We then describe

a new set of experiments in the US and Denmark that combine and extend elements

of the prior studies and that assess the role of motivated reasoning in citizens'

judgments of performance information in very different country contexts. We

next present our findings that motivated reasoning of both kinds (based on partisan

identities and governance preferences) influences accuracy judgments and associated

assessments of evidence strength, although the results are mixed and not

entirely consistent with the prior studies. Finally, we draw some conclusions and

set out the implications for research and practice.

MOTIVATED REASONING ABOUT PERFORMANCE INFORMATION

A widespread assumption in research and contemporary policy is that performance

measurement and reporting corrects information deficits among citizens and

consumers. This feature of performance information has encouraged reformers to create systems of public performance reporting that have become integral and important in many developed democracies (Boyne et al. 2003; Hatry 2006). In this way, performance information is assumed to contribute to better knowledge about public services. However, recent evidence reveals several biases in citizens' search, reception, and use of performance information (Olsen 2013; 2015; Andersen and Hjortskov 2015; Baekgaard and Serritzlew 2016, 2019; James and Van Ryzin 2017b; Baekgaard et al. 2019; James, Jilke, and Van Ryzin 2017). An important strand of this research focuses on the role of political biases, especially forms of politically or ideologically motivated reasoning about performance information (Baekgaard and Serritzlew 2016; James and Van Ryzin 2017b).

Motivated reasoning recognizes the fact that reasoning is not always rational but often driven by desires or preferences that, in the words of Kunda (1990), "affect reasoning through reliance on a biased set of cognitive processes—that is, strategies for accessing, constructing, and evaluating beliefs." In line with prior research (e.g., Taber and Lodge 2006; Bolsen, Druckman, and Cook 2014; Baekgaard and Serritzlew 2016; James and Van Ryzin 2017b), we distinguish between two broad categories of goals. Accuracy goals motivate individuals to "seek out and carefully consider relevant evidence so as to reach a correct or otherwise best conclusion". Directional goals, in contrast, motivate people to "apply their reasoning powers in defense of a prior, specific conclusion" (Taber and Lodge 2006, p. 756). Thus, directional goals refer to situations in which a "person is motivated to arrive at a particular conclusion" (Kunda 1999, p. 236). For example, motivated reasoning can entail people engaging in confirmation bias to seek out new evidence that is consistent with their pre-existing views or disconfirmation bias in trying to find fault with evidence inconsistent with their prior views (Lord, Ross, and Lepper 1979; Taber and Lodge 2006).

4

We consider two sources of motivated reasoning that are especially relevant to the provision of information about government performance. The first is the motivation to remain consistent with one's party identification, following an approach in political psychology applied by Taber and Lodge (2006). Citizens' identification with political parties exerts a powerful influence on how they perceive and reason about the social and political world (e.g., Druckman, Peterson, and Slothuus 2013; Bolsen et al. 2014). Thus, partisanship serves as a potent perceptual screen (Campbell et al. 1960, p. 133), leading partisans to perceive the same reality differently and to reject new information which is inconsistent with their desires and beliefs (Lodge and Taber 2013). In the context of government performance, which is characterized by often controversial debates over the content, shape, and scope of major public services (James and Van Ryzin 2017b), partisan identity becomes especially important in public opinion formation. Thus, the drive to remain loyal to one's own party and to articulate differences with opposing parties is likely to activate directional motivated reasoning (Bolsen et al. 2014). Consistent with this proposition, a study by James (2011) found that citizens who voted for the party controlling a local government were more influenced by positive information about the public services it provided than were citizens who did not vote for the party in power. Relatedly, Van Ryzin (2013) found that political conservatives were especially sensitive to being primed about their expectations concerning local government performance compared to moderates and especially liberals, which may be attributable to the more negative baseline attitudes of conservatives towards government. Likewise, a related stream of literature shows that partisans attribute responsibility for economic crisis selectively depending on whether they support the government or not (Bisgaard 2015; Tilley and Hobolt 2011).

A clear piece of evidence for partisan-motivated reasoning about performance

5

information comes from a study by James and Van Ryzin (2017b) of how citizens

process performance information about the ACA. The ACA remains the object of

intense disagreement between Democrats, who generally support the ACA, and

Republicans, who have worked to overturn the act (Jacobs and Skocpol 2012).

This partisan divide was evident when citizens were asked to choose indicators,

or evaluate statements, about state-level indicators of the performance of the

US health care system, and the divide was amplified under a political

prime compared to a health care needs prime. Since President Trump's election, the

debate over the ACA has become even more politicized as Republicans have made

various attempts to "repeal and replace" the ACA along strictly partisan lines.

The second form of motivated reasoning relates to people's governance preferences,

specifically their views regarding the roles of the public and private sectors

in service delivery. Baekgaard and Serritzlew (2016) found evidence of this form of

motivated reasoning by citizens in the context of Danish health care. Baekgaard

et al. (2019) replicated and extended these findings to politicians. Preferences about

government (vs. private) service delivery may be linked to deeply held beliefs and

identities rooted in people's experiences with the public or private sectors, thus

influencing their views about the different goals of public and private organizations

and, in turn, their performance (Christensen et al. 2018). Indeed, in a study focused

on hospitals and schools, Baekgaard and Serritzlew (2016) found that governance

preferences about public versus private service provision affected whether participants

made correct interpretations of unambiguous performance information.

Specifically, they found that participants in their experiments tended to perceive

their preferred type of organization (public or private) as performing better, independent

of what the unambiguous information suggested.

PRIMING AND MOTIVATED REASONING

Our current study extends the examination of priming effects to an accuracybased task involving performance information. James and Van Ryzin (2017b, with reference to Higgins (1996) and Kay and Ross (2003):199), previously used priming in their experiment, defined as selective activation of "the cognitive accessibility of some concepts above others." In the context of government performance, a prime aimed at activating directional reasoning may create selective attention on partisan conflicts and reinforce bias in interpretations and processing of performance information. In contrast, a prime aimed at activating accuracy goals may attenuate bias by focusing people's minds on accurate and neutral processing. James and Van Ryzin's (2017b) study of the ACA found that citizens who were primed to think about politics engaged in stronger partisan-biased processing of information from real reports about the performance of the ACA. In contrast, when primed to think about their own health care needs, partisan-biased processing of information appeared to be reduced. Thus, partisan motivated reasoning was stimulated when citizens were primed to think politically, while accuracy goals were more prevalent when citizens were primed to think about their own health care needs. The political prime widened the gap between Republicans' and Democrats' judgments of the strength of evidence favorable to the ACA, with Republicans rating it as less strong. Additionally, Republicans' choice of less favorable performance indicators from an array, relative to those chosen by Democrats, was intensified under the political prime. In contrast, the political prime did not increase partisan differences in rating the strength of evidence of unfavorable statements about the Act. Motivated reasoning by Republicans seemed primarily related to their being relatively unwilling to view favorable evidence about the ACA as strong evidence under the political prime.

In the current study, we extend the study of priming to examine priming effects relative to a neutral, no-prime condition. This enables us to examine outcomes

from political priming and health priming compared not only to each other but also to the outcomes of decision making that occurs more naturally when people are not primed in either way. Moreover, we assess effects for both of the two main existing versions of motivated reasoning in the literature, namely partisan identities and governance preferences. To do so, we conduct new experiments in the US and Denmark that extend knowledge beyond findings that have already established the relevance of partisan identities and governance preferences as forms of motivated reasoning in these contexts. Specifically, the experiment in the US focuses on partisan-motivated reasoning, while the experiment in Denmark focuses on governance preferences as the basis for motivated reasoning. In line with James and Van Ryzin (2017b), we use a political prime consisting of a series of survey questions about political ideology and the role of government that aims to induce reasoning along political lines. The health care needs prime consists of a series of questions about participants' health and health care needs that aims to stimulate accuracy-based motivations stemming from their real-life needs for health services.

AIMS AND EXPECTATIONS

In the US experiment, we examine partisan-motivated reasoning in an accuracybased task about the performance of the ACA. We expect a relationship for partisan identity and assessments of health plans labeled as being part of the ACA, 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).

As mentioned, in the experiment in Denmark, we examine motivated reasoning based on governance preferences. We contrast a neutral frame with a politicized frame to present information about a comparison between public and private hospitals. Our expectation is that the politicized frame activates ideologically based motivated reasoning based on citizens' governance preferences, which reduces the

accuracy of assessments of performance.

The current literature on motivated reasoning only gives limited insight into the processes involved when people think about the performance information given to them. In order to examine how people think when undertaking the accuracy-based task, we asked participants to rate the information provided in terms of its strength as evidence about the performance of the organizations. Our expectation is that information not consistent with partisan identities or governance preferences will be rated as less strong evidence by experimental participants.

EXPERIMENTAL DESIGN AND METHOD1

Figure 1 shows the design of our two experiments to test these expectations. The Denmark experiment focused on motivated reasoning associated with governance preferences for public- versus private-sector service provision. Based on the study by Baekgaard and Serritzlew (2016), we measure governance preferences at the outset, prior to randomization. Governance preferences are measured by the following questions (using a five-point agree-disagree scale):

Do you disagree or agree with the following statements?

1. Many public activities could be produced both better and cheaper by private providers.

2. We should to a larger degree contract out public services (such as child care, elderly care, and hospital treatments).

3. The public sector is best at providing public services.

The index of governance preferences runs from 0 to 100, with higher values denoting a preference for public service provision. We included a series of washout questions (on music preferences) after the governance preferences questions so as to minimize potential priming effects of these questions.

The US study instead focused on partisan-motivated reasoning, asking respondents for their party-political affiliations using the same method as James and Van

9

Ryzin (2017b). Using standard questions asked by the Pew Center (2014), this method classified participants as either Democrats or Republicans based on their self-identification as well as how they voted in the last US presidential election.

Figure 1. Experimental design.

In both Denmark and the US, we included the politics and health care primes from the experiment by James and Van Ryzin (2017b), along with a no-prime control condition. The political prime consisted of a 10-point, left-right scale of political views followed by a series of forced choice questions on the role of government from a Pew Center (2014) poll about political polarization in the US. Specifically, the questions asked respondents to choose the statement that comes closest to their views:

_ (a) government often does a better job than people give it credit for; or (b) government is almost always wasteful and inefficient;

_ (a) government regulation of business is necessary to protect the public interest; (b) government regulation of business usually does more harm than good; and

_ (a) poor people have hard lives because government benefits don't go far enough to help them live decently; or (b) poor people today have it easy because they can get government benefits without doing anything in return.

Respondents were also asked the following question: Do you think it is the responsibility of the federal government to make sure all Americans have health care coverage, or is that not the responsibility of the federal government? The purpose of these five political priming questions was to activate associations related to political ideology and the role of government, immediately in advance of their consideration of the performance information.

For the Danish respondents, the political priming questions were modified to fit the Danish political and health care contexts in order to make them relevant and

meaningful to participants. The same 10-point, left-right political ideology question as in the US experiment was asked. Next, the following statements and questions were used in Denmark:

_ (a) Government often does a better job than people give it credit for; (b) government is almost always wasteful and inefficient.

_ (a) Government regulation of business is necessary; (b) government regulation of business usually does more harm than good

_ (a) Social welfare beneficiaries have easy lives because they can receive public benefits without doing anything in return; (b) Social welfare beneficiaries have hard lives because the public benefits are too small to support a proper standard of living.

_ (a) Do you think that the public sector should pay for treatments of citizens at private hospitals? (Yes, the public sector should pay for treatments at private hospitals; No, the public sector should not pay for treatments at private hospitals).

The health care prime, for US respondents, involved a set of questions from a Kaiser Family Foundation/NBC News (2014) poll about people's concerns about access to and affordability of health care. Specifically, they were asked how worried they are (from 1¼not at all worried, to 4¼very worried) about the following:

_ having to pay more for health care or health insurance,

_ not being able to afford needed health services,

_ not being able to afford needed prescription drugs,

_ being locked in a job for fear of losing health benefits, and

_ losing health insurance coverage.

Participants were also asked to rate your own health these days on a five-point scale from 1¼poor to 5¼excellent. The aim of these six health care priming questions was to activate associations related to personal concerns about health and the

need for health care, again in advance of considering the performance information. For the Danish respondents, these items were modified to make them relevant to the Danish health care context. The following statements were used:

_ having a good hospital close to my home,

_ having access to good surgeons,

_ having access to the newest medical equipment

_ not being able to afford the medicine I need

_ that error may occur during hospitalization

The main outcome measured in both experiments was an accuracy task based on the study by Baekgaard and Serritzlew (2016), sometimes referred to as a covariance detection task (Kahan et al. 2017). Figure 2 presents the tasks, which involved the interpretation of a basic 2_2 table of frequencies. For the Danish experiment, as in the study by Baekgaard and Serritzlew (2016), participants were asked to interpret information about performance outcomes for surgical operations. The table enables participants to compare the ratio of hospital patients who experienced a positive outcome (operation with no complications) to those who experienced a negative one (operation with complications). In one treatment arm, these ratios are labeled to present the information as being about a public hospital compared to a private hospital. This enables participants to calculate ratios of operations with/without complications for the private hospital and to compare it with the ratio of operations with/without complications for the public hospital. Comparing these ratios allows participants to decide which hospital, public or private, is performing better. In each case in Figure 2, the correct answer is to choose the bottom row. In the hospital case, the odds of an operation without complications are over 5 to 1 in the private hospital, in contrast to only 3 to 1 odds of a successful outcome of this kind in the public hospital in the top row.

In the other treatment arm for the Danish hospital experiment, as shown in

Figure 2, identical information about the outcomes of operations with and without

complications was presented. However, the hospitals in the rows were "unlabeled",

Figure 2. Performance information (as presented to participants).

in the sense of being neutrally described as Hospital A compared to Hospital B, in

order not to invoke public/private sector governance preferences. This experimental

design enables the accuracy of participants in the labeled (public/private hospital)

task to be compared with the accuracy of participants in the unlabeled (hospital

A/B) task. The difference in accuracy provides a measure of the effect of the

public/private labeling compared to the unlabeled benchmark.

If motivated by accuracy goals, we would expect respondents to use cognitive

effort and attention to arrive at a correct answer. While the ratios can be calculated

and compared by participants, without cognitive reflection, inaccuracy may creep

in, or the implications of the ratio-based comparison may not be fully processed.

For example, participants may use the raw figure in the upper-left cell, which indicates

203 operations without complications (Figure 2). This number suggests more

successes than failures, looking either row-wise or column-wise: This may lead

people to choose the top-row hospital as performing best. Instead, arriving at the

correct decision about comparative performance requires cognitive effort. The ratios

are not intuitively evident and thus need to be calculated (or estimated) mentally. In

other words, respondents must detect the covariance in the table, not just compare

numbers row-wise or column-wise. If motivated by directional goals rather than a

search for accuracy, we would expect respondents to more readily rely on cues (particularly

the row labels) in ways that confirm their existing orientation towards the

issue of performance. In the case of the Danish hospitals, the governance preferences

about public/private sector are invoked by their description as being public or

private hospitals, and these are expected to affect accuracy, with pro-public sector

preferences being more inaccurate in this case because of the information pointing

against the public sector hospital performance compared to the private hospital.

An analogous logic was used in the experiment with US participants but with

the task presented for insurance policies and the ACA (following the focus on this

issue by James and Van Ryzin 2017b). The task involved comparison between two

health insurance plans, the ACA plan in the top row and an employer-provided plan

(which is the typical private-sector option for most Americans) in the bottom row

(as shown in Figure 2). As mentioned, the ACA is a public health care policy that

has been controversial in the US, with evidence of partisan-motivated reasoning in

consideration of evidence about it (James and Van Ryzin 2017b). The new design

in this experiment enables partisan-motivated reasoning to be assessed for a task in

which performance information about outcomes is unambiguous, using the same

design as for the Danish hospitals but adapted to health insurance in the US. In

the same way as the Danish hospital experiment (in which the public hospital was

presented as performing relatively less well than the private), the ACA plan was

suggested as performing relatively less well than the employer plan. In the US

experiment, the other "unlabeled" treatment arm of the experiment refers simply to

insurance plan A in the top row and insurance plan B in the bottom row.

In the analyses that follow, the main dependent variable of interest is a dummy

variable taking the value 1 for respondents giving the correct answer and 0 for

respondents giving the wrong answer about the best-performing hospital or health

insurance plan. Some studies using the covariance detection task randomize the

rows of the table as well to vary which row represents the correct answer (Kahan

et al. 2017). However, previous research suggests that randomizing the rows in this

way has little impact on the overall results in similar contexts (Baekgaard et al.

2019; Baekgaard and Serritzlew 2016). Given our 3_2 factorial design

(priming_labeling) in each experiment, and the resulting sample size concerns, we

14

did not randomize the rows in this way.

In addition to gauging a correct or incorrect choice, we also asked

participants an additional question about the perceived strength of evidence.

This measure was newly developed for the current experiments but drew on

the approach to evidence assessment in the study by James and Van Ryzin

(2017b). We asked participants the following question: How would you rate the

information in the table you just saw as evidence about the performance of these

two health plans [hospitals]? The response scaled ranged from 0¼very weak

evidence to 10¼very strong evidence. This variable allows us to gauge people's

perceptions of the strength of evidence, independent of whether they interpreted

the information correctly or not.

For the US study, we used the Qualtrics research panel and gathered data in

March 2017 from n¼1051 respondents, who were balanced with the US population

in terms of gender, age, race, income, and region of country. For the Danish

study, we collected data in January 2017 from n¼3,228 respondents from the

Danish YouGov panel (balanced with the Danish population in terms of gender,

age, and municipal residence).

RESULTS, PART I: THE US STUDY

Regarding the US study, the analysis focuses on party identification and motivated

reasoning. The partisan form of motivated reasoning is especially relevant in

the current context of health care reform and partisan debate in the US. Indeed,

the US study was conducted online in late March, 2017, when President Trump

and the Republican-controlled Congress were involved in a concerted effort to

repeal the ACA, which was especially alarming to Democrats. We run our analysis

of the effect of labeling the health care performance information (ACA Plan/

Employer plan versus Plan A/Plan B) under the three priming conditions (no

prime, politics prime, and health prime) separately for political party identification.

15

Figures 3 and 4 graphically present the average correct interpretations and the

rating of evidence strength, by priming condition and party identification, for the

US study. Table 1 presents the corresponding regressions.

Beginning with Figure 3, we see that the Democrats under both the no-prime

control condition and the politics prime are less likely to come to a correct interpretation

when the performance information is labeled ACA/employer plan. As

already mentioned, it is interesting to note that the no-prime condition should be

interpreted against a background of intense political debate and media coverage of

Republican efforts to repeal the ACA. This perhaps helps explain the similarity in

results between the no-prime condition and the politics priming condition. Across

all priming conditions, 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). However,

when primed to think about their need for and access to health care, the

Democrats judge the labeled health performance information with greater

The results for the Republicans, shown in the right panel of Figure 3, show

that they judge performance information similarly whether labeled as an ACA/

employer plan or simply Plan A/B. In other words, the Republicans seem less

sensitive to the labeling of the performance information, which is perhaps to be

expected given that their partisan-motivated choice (the employer-provided plan)

is the correct choice in our experimental design. Interestingly, the Republicans

appear somewhat more likely to make a correct choice, regardless of labeling,

when primed to think about their own need for and access to health care.

Indeed, the health care prime is associated with a statistically significant gain

of almost 15 percentage points in the rate of correct answers (Table 1, Model 3).

This finding is also consistent with the notion that a person's awareness of the

need for a service enhances accuracy goals.

16

Figure 4 presents the US results for ratings of evidence strength, with the corresponding

regressions in Table 1. Overall, the experimental effects of both priming

and labeling on ratings of evidence strength are small and not statistically significant

(Table 1, Models 5-8). For Democrats in the study, the health prime seems

to increase their ratings of evidence strength of the labeled information, but

again this interaction effect is not statistically significant (Table 1, Model 6). For

Republicans, ratings of evidence strength are largely independent of labeling and

priming, although the evidence under the no-prime condition and the politics

priming condition appears to be judged somewhat stronger when labeled as ACA/

employer. Again, this could be due to the fact that the evidence favors the

employer-provided plan over the ACA plan.

RESULTS, PART II: THE DANISH STUDY

Denmark is a parliamentary system with multiple parties that are not so easily

classified into two main parties (as in the US) or clear rival categories (Baekgaard

2011). Following the original study by Baekgaard and Serritzlew (2016), therefore,

the analysis focuses on governance preferences as measured by the authors' threeitem

scale described earlier. In the figures, we make a simple split at the midpoint

of the scale, as respondents scoring below 50 mostly favor private service provision,

while respondents scoring above 50 mostly favor public service provision.

Figures 5 and 6 show the experimental results graphically, while Table 2 presents

the regressions.

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). The

labeling effects are significant within the politics prime and the health-care prime,

while we do not identify significantly different effects of labeling across priming

17

conditions (Table 2, Model 2). When labeled public/private, participants who

favor the public sector make less correct interpretations. For the neutral prime,

the ability to interpret correctly is reduced by almost six percentage points; for

political and health care prime, by 10-12 percentage points. For those who favor

the private sector, the labeling of the information has little effect on the correctness

of interpretations. This finding is consistent with the original finding in

Baekgaard and Serritzlew (2016) that labeling matters differently depending on

governance preferences.

Although the present experiment is not a direct replication of the 2016 study, it is

possible to compare the results for the neutral prime of the present study with the

results of Treatment 1 in the Baekgaard and Serritzlew (2016) study. In the online

appendix, Tables A1–A3, we compare the results of the present study with the 2016

study, broken down by prime. In Table A2, Model 1, we find a negative, marginally

significant effect of a pro-public governance preference on the likelihood of correct

interpretation. In the 2016 study (Table 3, Model 1 in Baekgaard and Serritzlew,

2016), a negative effect (p<0.01) is found. Hence, in both studies, respondents with a

more favorable view of the public sector make a more biased interpretation when presented

with information indicating that the private hospital performs best. The results

are weaker in the present study (an effect of _0.007 compared with _0.019 in the

2016-study). In the placebo group (Table 3, Model 3 in the 2016 study; Table A2,

Model 2 in the online appendix), both studies indicate a null effect. The results of

both studies show that the biased interpretation disappears as expected when neutral

labeling is used. Finally, the 2016 study shows in an interaction model that the difference

of the estimate of the effect of pro-public attitudes between the treatment and

control groups is statistically significant (Table 3, Model 5 in Baekgaard and

Serritzlew, 2016). In the present study, the difference is smaller and insignificant.

(Table A2,Model 3 in the online appendix). This reflects the weaker effect of pro-public

18

governance preferences for the treated respondents. In sum, the results of the present

study are consistent with Baekgaard and Serritzlew (2016), although the effects

are weaker and less securely estimated. Moreover, in accordance with the general

expectation that political priming increases ideologically motivated reasoning, while

health priming has a negative impact, Tables A1 and A3 provide evidence of even

weaker effects in the health care prime condition, while the results are stronger in the

political prime condition.

Turning to Figure 6, the Danish results show that participants who favor the

public sector overall rate the evidence as less strong than those who favor the private

sector (see also Table 2, Models 5–8). This overall difference in levels appears

to be independent of the labeling of the performance information. Within the

group of participants who tend to favor the public sector, the priming effect is significant,

driven mainly by higher evidence ratings in the health priming condition

than in the other conditions. Within the group favoring the private sector, both the

main labeling effect and the political priming effect are statistically significant

(Table 2, Model 7). When labeled public/private hospital, those who favor the private

sector consistently judge the performance information as providing stronger

evidence. As before, this may be due to the fact that the correct interpretation of

the information supports the conclusion that the private hospital is performing

best. Figure 6 is helpful in terms of providing an explanation of the priming effects.

For all respondents, regardless of whether they favor the public or the private

sector, the politics prime tends to produce less strong evaluations than the other

conditions. Thus, there is some evidence that political priming—albeit weakly so—

tends to reduce evidence ratings.

DISCUSSION AND IMPLICATIONS

The findings of our two experiments extend knowledge about motivated

reasoning concerning performance information. In accordance with previous

19

studies, we find that prior beliefs matter for how new information is processed

and understood. However, we extend these findings to an accuracy-based task

not previously examined for partisan motivated reasoning and extend findings

by examining priming of governance preference-based motivated reasoning.

As a general finding, political priming tends to reduce accuracy and leads to

ratings of the information as weaker evidence. In contrast, a health care prime

tends to increase accuracy and leads to ratings of the information as stronger

evidence. However, the results of the US and Danish experiments also have

country-specific nuances, which is to be expected since they are undertaken in

different political contexts.

In the US context, especially given the highly political period during which the

study was conducted (as Republicans were on the verge of repealing the ACA),

Democrats were the participants most clearly influenced by the labeling of the

performance information, and their biases were only partially mitigated by the

health care prime. Republicans were less sensitive to the labeling of the performance

information, and the health care prime enhanced their accuracy in both the

labeled and unlabeled conditions. The Danish study, however, showed that priming

matters for the evaluation of evidence strength but had less of an effect on accuracy.

This result may, at first sight, seem surprising. If respondents find the evidence

weaker in the political prime compared to the health care prime, why are the

effects on accuracy limited? One reasonable explanation is that respondents were

presented with one piece of information, which had only one factually correct

interpretation. In the political prime, Danish respondents found the evidence

weaker, but they were not provided with other information and thus had to base

their evaluation on this information alone. Even if the information may have

seemed less credible, there was nothing else to base the evaluation on, and, hence,

there were limited effects for accuracy. This is relevant to further studies of the

cognitive process of evaluating information in a political context. It could be particularly

interesting to study whether a lower evaluation of information strength

more generally translates into accuracy if there are more pieces of information to

choose from than just one.

Adding the neutral no-prime condition provides insight about the question of

whether the effect of priming is due to politicization or to de-politicization. In the

US, especially for Democrats at the time of the study, the neutral (no-prime) condition

appears to be akin to the political priming condition, with clear evidence in

both cases of partisan-motivated reasoning. Thus, the main change from the status

quo appears to be the de-politicization associated with the health care prime. In

the Danish study, we find evidence of both: Compared to the neutral prime, the

political prime leads to a lower evaluation of evidence strength while the health

care prime lead to a higher evaluation. However, the effects for accuracy are too

weak to show a different effect of the two types of primes compared to the neutral

(no-prime) condition.

Our findings are also consistent with previous research. First, higher partisanmotivated

reasoning was evident for politically primed Democrats in the US

experiment. This is consistent with the earlier findings of James and Van Ryzin

(2017b) and extends evidence about this form of partisan-political motivated reasoning

to the domain of the accuracy task of judging unambiguous performance

information. The influence of prior governance preferences (for public or private

service provision) was also consistent with the previous study of Baekgaard and

Serritzlew (2016). Prior governance preferences exerted a negative impact on

accuracy in the treatment in which public/private labels were used compared to the

neutral labeling of the organizations being compared.

The findings show that priming matters to the way performance information is

interpreted, and motivated reasoning was evident in both experiments, despite important

differences in how people in the US and Denmark view health care as a public

policy. While health care is generally considered a universal right in the Danish political

system, it is an object of intense political disagreement in the US. An important

line of future research is to look at citizens being exposed to conflicting statements

about performance from different political sources. Multiple experiments of this kind

could be performed on MTurk, given recent research showing the usefulness of this

source of data to experiments in public management research (Stritch, Pedersen, and

Taggart 2017).

Our findings offer insights into potential ways to reduce bias by discouraging

citizens from thinking in terms of their partisan identities or governance preferences.

There is a need for more research on the mechanisms at work in these processes.

Motivated reasoning is probably strongest when the issue is related to

identity. This insight is consistent with emerging findings that point to the need for

strategies to protect science communication from antagonistic cultural meanings

that inhibit rational processing of evidence (Kahan et al. 2017). Additionally,

future research should consider the forms of cognition that occur during reasoning

and whether they differ between kinds of motivated reasoning. For example, a

hypothesis could be that partisan labels are an obvious cue to participants. If

someone supports party X, they may disregard the achievements of party Y. In

contrast, ideological reasoning based on thinking about public sector organizations

may be more engaged and along the lines of avoiding cognitive dissonance.

However, for some participants, particularly in contexts where there has been

much discussion of public and private provision of services because of public

debate about the merits and costs of privatization, the labels may also serve as

cues. Future empirical studies should seek to tease apart these differences in mechanisms

and their outcomes for perceptions, attitudes, and behavior.

Strategies to provide citizens with performance information that they will process

accurately can include the use of credible sources of information that users

trust. In other contexts, an independent source of performance information has

been found to be an important way of overcoming potential skepticism about the

information. For example, citizens may be skeptical when public organizations

report their own good performance, and thus, independent non-governmental sources

partially counter this problem (James and Van Ryzin 2017a). However, such

independent sources may not be so effective in cases of deep-seated identity-based

reasoning that resists this influence. The ACA, in the current highly politically

charged context in the US, represents an important example in which divergence

from identity-based reasoning about information could alienate someone from

their identity-defining group. If instead of relying mainly on sources of information,

different identities among recipients could alternatively, or additionally, be

invoked, such a strategy may be able to directly tackle the roots of motivated reasoning.

The experiments set out here used priming of citizens' need for health care

services to make that part of their identity more salient to them prior to considering

information. In public debate, making some of these shared identities more

salient may be more effective than providing independent sources of information in

reducing bias of this kind. In this way, while groups may not be able to agree on

the ends of public policy, they may be able to move towards more consensus about

what counts as evidence to inform policy decisions.

Supplemental data can be accessed at https://doi.org/10.1080/10967494.2019.1659891.

NOTE

1. Since our study combines the design in Baekgaard and Serritzlew (2016) with

the design in James and Van Ryzin (2017b), this section draws heavily and

sometimes verbatim on these articles in its description of survey questions and

measures used in this study.

REFERENCES

Andersen, S. C., and M. Hjortskov. 2015. "Cognitive Biases in Performance Evaluations."
Journal of Public Administration Research and Theory 26(4):647–662. doi: 10.1093/
jopart/muv036.

Baekgaard, M. 2011. "Committee Bias in Legislatures with a High Degree of Party
Cohesion: Evidence from Danish Municipalities." European Journal of Political
Research 50(3):315–335. doi: 10.1111/j.1475-6765.2010.01927.x.

Baekgaard, M., J. Christensen, C. M. Dahlmann, A. Mathiasen., and N. B. G. Petersen.
2019. "The Role of Evidence in Politics: Motivated Reasoning and Persuasion among
Politicians." British Journal of Political Science 49(3):1117–1140. doi: 10.1017/
S0007123417000084.

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

Baekgaard, M., and S. Serritzlew. 2019. "Those Who Understand It Will Not Be Persuaded:
A Performance Information Paradox." International Public Management Journal 1–23.
doi: 10.1080/10967494.2018.1461152.

Bisgaard, M. 2015. "Bias Will Find a Way: Economic Perceptions, Attributions of Blame,
and Partisan-Motivated Reasoning during Crisis." The Journal of Politics 77(3):
849–860. doi: 10.1086/681591.

Bolsen, T., J. N. Druckman, and F. L. Cook. 2014. "The Influence of Partisan Motivated
Reasoning on Public Opinion." Political Behavior 36(2):235–262. doi: 10.1007/s11109-
013-9238-0.

Boyne, G. A., C. Farrell, J. Law, M. Powell, and R. Walker. 2003. Evaluating Public
Management Reforms: Principles and Practice. Buckingham, UK: Open University Press.

Campbell, A., P. E. Converse, W. E. Miller, and D. E. Stokes. 1960. The American Voter.
New York: Wiley.

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

22 International Public Management Journal Vol. 0, No. 0, 2019

Governance Preferences, and the Process of Goal Reprioritization." Journal of Public Administration Research and Theory 28(2):197–211. doi: 10.1093/jopart/muy001.

Druckman, J. N., E. Peterson, and R. Slothuus. 2013. "How Elite Partisan Polarization Affects Public Opinion Formation." American Political Science Review 107(1):57–79. doi: 10.1017/S0003055412000500.

Hatry, H. P. 2006. Performance Measurement: Getting Results, 2nd ed. Washington, DC: The Urban Institute Press.

Heinrich, C. J. 2003. "Measuring Public Sector Performance." Pp. 25–37 in Handbook of Public Administration, edited by B. G. Peters and J. Pierre. London: Sage.

Higgins, E. T. 1996. "Knowledge Activation: Accessibility, Applicability, and Salience." Pp. 133–168 in Social Psychology: Handbook of Basic Principles, edited by A. W. Kruglanski and E. T. Higgins. New York: Guilford.

Jacobs, L. R., and T. Skocpol. 2012. Health Care Reform and American Politics: What Everyone Needs to Know. New York: Oxford University Press.

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. doi: 10.1093/jopart/muq057.

James, O., and G. G. Van Ryzin. 2017a. "Incredibly Good Performance: An Experimental Study of Source and Level Effects on the Credibility of Government." The American Review of Public Administration 47(1):23–35. doi: 10.1177/0275074015580390.

James, O., and G. G. Van Ryzin. 2017b. "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. doi: 10.1093/jopart/muw049.

James, O., S. Jilke, and G. G. Van Ryzin. 2017. Experiments in Public Management

25

Research: Challenges and Contributions. Cambridge: Cambridge University Press.

Kahan, D. M. 2013. "Ideology, Motivated Reasoning, and Cognitive Reflection." Judgment and Decision Making 8(4):407–424. doi: 10.2139/ssrn.2182588.

Kahan, D. M., E. Peters, E. C. Dawson, and P. Slovic. 2017. "Motivated Numeracy and Enlightened Self-Government." Behavioural Public Policy 1(1):54–86. doi: 10.1017/bpp. 2016.2.

Kaiser Family Foundation. 2014. "Public Opinion of the ACA at the End of the First Open Enrollment Period." JAMA 311(19):1957. doi: 10.1001/jama.2014.5269.

Kay, A. C., and L. Ross. 2003. "The Perceptual Push: The Interplay of Implicit Cues and Explicit Situational Construals on Behavioral Intentions in the Prisoner's Dilemma." Journal of Experimental Social Psychology 39(6):634–643. doi: 10.1016/S0022-1031(03)00057-X.

Kunda, Z. 1990. "The Case for Motivated Reasoning." Psychological Bulletin 108(3): 480–498. doi: 10.1037/0033-2909.108.3.480.

Kunda, Z. 1999. Social Cognition Cambridge: MIT Press.

Lodge, M., and C. S. Taber. 2013. The Rationalizing Voter. Cambridge: Cambridge University Press.

Lord, C. G., L. Ross, and M. R. Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." Journal of Personality and Social Psychology 37(11):2098–2109. doi: 10.1037/0022-3514.37.11.2098.

Olsen, A. L. 2013. "Leftmost-Digit-Bias in an Enumerated Public Sector? An Experiment on Citizens' Judgment of Performance Information." Judgment and Decision Making 8(3):365–371.

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

Pew Center. 2014. Political polarization in American public. Retrieved September 2, 2019 (http://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/).

Stritch, J. M., M. J. Pedersen, and G. Taggart. 2017. "The Opportunities and Limitations of

Using Mechanical Turk (Mturk) in Public Administration and Management

Scholarship." International Public Management Journal 20(3):489–511. doi: 10.1080/

10967494.2016.1276493.

Taber, C. S., and M. Lodge. 2006. "Motivated Skepticism in the Evaluation of Political

Beliefs." American Journal of Political Science 50(3):755–769.

Tilley, J., and S. B. Hobolt. 2011. "Is the Government to Blame? An Experimental Test of

How Partisanship Shapes Perceptions of Performance and Responsibility." Journal of

Politics 73(2):316–330. doi: 10.1017/S0022381611000168.

Van Ryzin, G. G. 2013. "An experimental test of the expectancy-disconfirmation theory of

citizen satisfaction". Journal of Policy Analysis and Management 32(3):597–614.

ABOUT THE AUTHORS

Martin Baekgaard (martinb@ps.au.dk) is a professor of political science at Aarhus

University. His research interests include governance regimes, biases in interpretation

of information of information, and administrative burden.

Oliver James (o.james@exeter.ac.uk), PhD, Professor, Department of Politics,

University of Exeter, UK. His research focuses on performance information design

and use, citizen-public service provider interaction, public management leadership

and performance outcomes, and using experiments as a method. Recent publications

include Experiments in Public Management Research (Cambridge, 2017,

edited with Sebastian Jilke and Gregg G. Van Ryzin) and James, O., Jilke, S.,

Petersen, C., & Van de Walle, S. 2016. "Citizens' blame of politicians for public

service failure." Public Administration Review, 76(1), 83–93.

Søren Serritzlew (soren@ps.au.dk), PhD, is a professor at the Department of

Political Science, Aarhus University, Denmark. His research interests include

effects of public sector reform, use of economic incentives in the public sector,

and democracy.

Gregg G. Van Ryzin (vanryzin@rutgers.edu), PhD, is a professor in the School of

Public Affairs and Administration (SPAA), Rutgers University, Newark, USA. His

work applies experimental and behavioral methods to various issues in public management,

including citizen satisfaction, coproduction, performance measurement,

representative bureaucracy, and organizational behavior. He is author (with

Dahlia Remler) of Research Methods in Practice (SAGE) and editor (with Oliver

James and Sebastian Jilke) of Experiments in Public Management
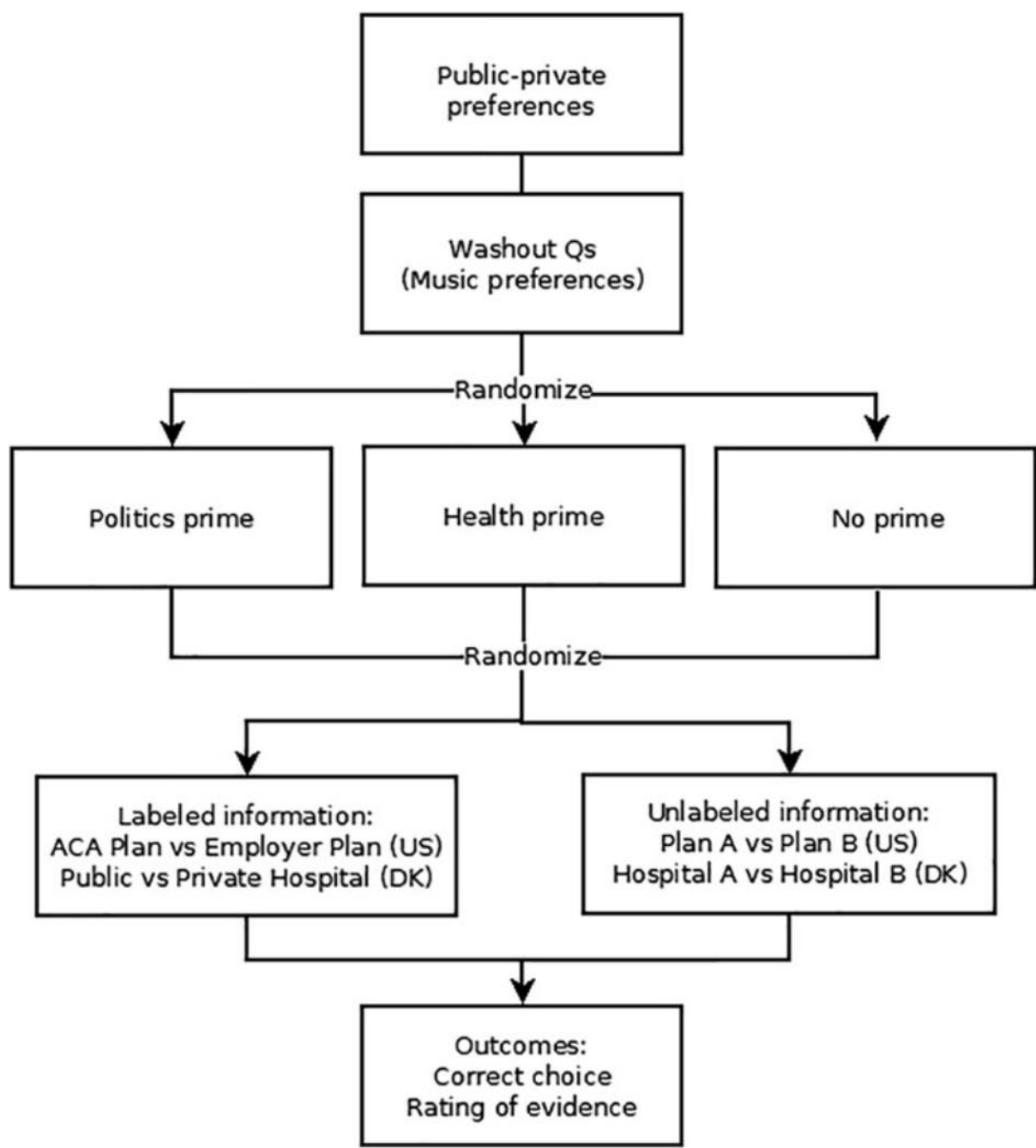
Research (Cambridge

Figure 1

Figure 2

US Experiment

(Labeled): Below is a table with information on appendectomies (the surgical removal of the appendix) under two insurance plans, one an employer-provided insurance plan and the other a plan provided through the Affordable Care Act (ACA). The types of people insured by the two plans are comparable. The table shows how often the operations were performed with, and without, complications.

	Operations without complications	Operations with complications
Affordable Care Act plan	203	68
Employer-provided plan	47	9

(Unlabeled): Below is a table with information on appendectomies (the surgical removal of the appendix) under two insurance plans, Plan A and Plan B. The types of people insured by the two plans are comparable. The table shows how often the operations were performed with, and without, complications.

	Operations without complications	Operations with complications
Insurance plan A	203	68
Insurance plan B	47	9

DK Experiment

(Labeled): Below is a table with information on appendectomies (the surgical removal of the appendix) at two hospitals, one a public hospital and the other a private hospital. The types of people treated at the two hospitals are comparable. The table shows how often the operations were performed with, and without, complications.

	Operations without complications	Operations with complications
Public hospital	203	68
Private hospital	47	9

Figure 2 (cont)

(Unlabeled): Below is a table with information on appendectomies (the surgical removed of the appendix) at two hospitals, Hospital A and Hospital B. The types of people treated at the two hospitals are comparable. The table shows how often the operations were performed with, and without, complications.

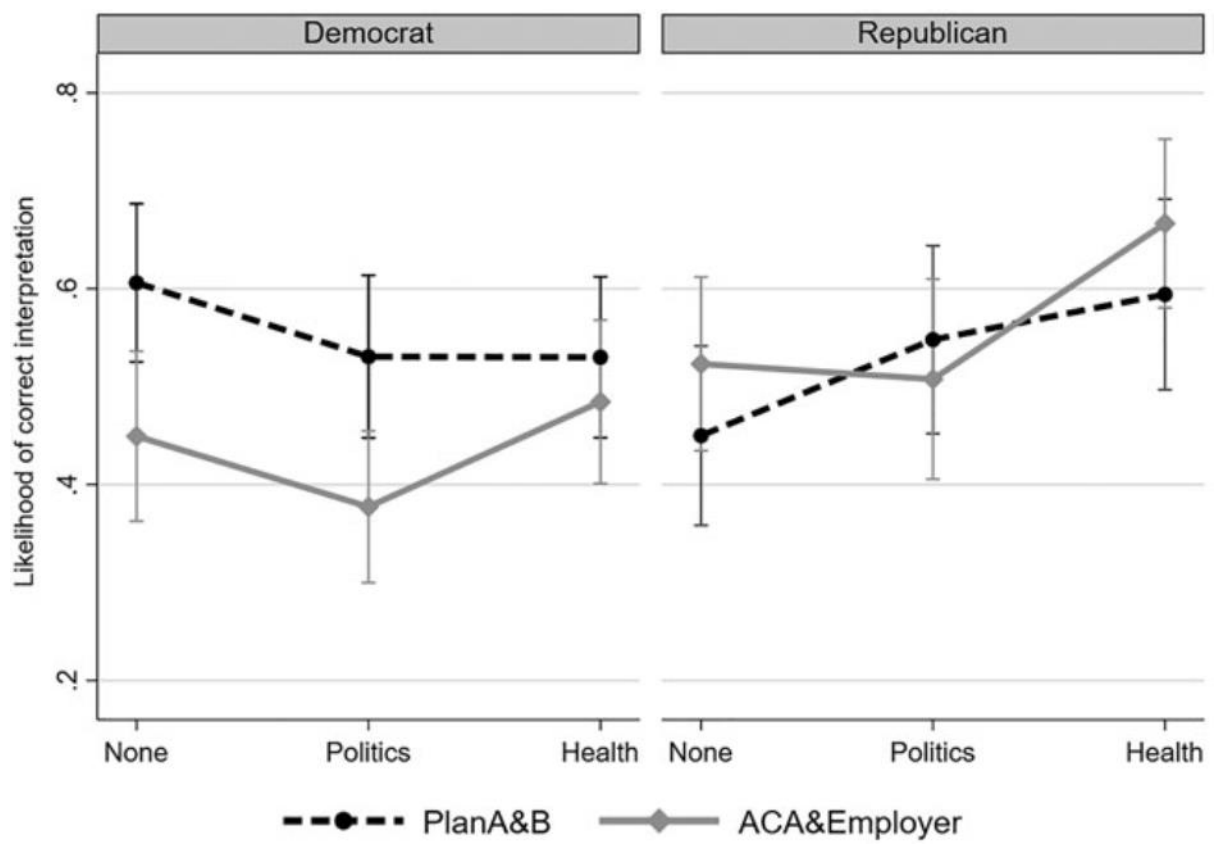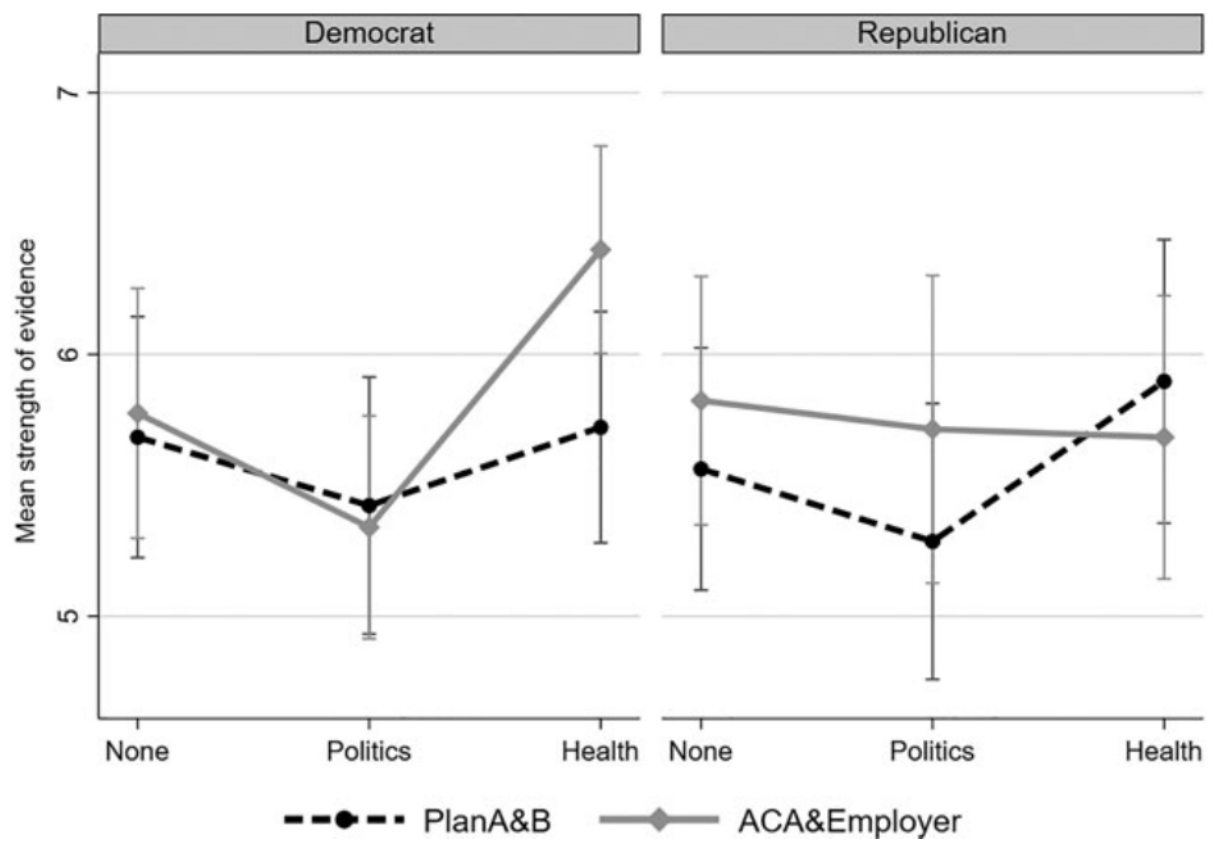|	Operations without complications	Operations with complications
Hospital A	203	68
Hospital B	47	9

Figure 3

Figure 4

TABLE 1
Regression Analysis of Likelihood of Correct Interpretation and Rating of Evidence Strength by Party ID—USA

	Likelihood of a correct answer (OLS)				Rating of evidence strength (OLS)			
	Democrats		Republicans		Democrats		Republicans	
	Direct effects model (1)	Interaction model (2)	Direct effects model (3)	Interaction model (4)	Direct effects model (5)	Interaction model (6)	Direct effects model (7)	Interaction model (8)
Priming (ref = no prime)								
Politics prime	−.075	−.075	.043	.098	−.356	−.261	−.202	−.277
	(.050)	(.071)	(.057)	(.080)	(.273)	(.384)	(.317)	(.445)
Health care prime	−.022	−.076	.145 **	.144 *	.326	.038	.082	.335
	(.051)	(.070)	(.056)	(.081)	(.275)	(.384)	(.308)	(.449)
Labeling (ACA/Employer)	−.118 ***	−.157 **	.038	.073	.226	.092	.155	.261
	(.041)	(.073)	(.047)	(.077)	(.223)	(.392)	(.258)	(.424)
Interactions								
Politics prime × Labeling		.003		−.114		−.174		.168
		(.101)		(.114)		(.545)		(.634)
Health prime × Labeling		.111		−.001		.587		−.475
		(.101)		(.112)		(.551)		(.618)
Constant	.588 **	.606 **	.468 ***	.450	5.620 ***	5.684 ***	5.617 ***	5.563 ***
	(.041)	(.050)	(.045)	(.055)	(.223)	(.271)	(.250)	(.304)
R^2	.012	.021	.017	.020	.013	.016	.003	.005
N	589	589	454	454	579	579	445	445

Note: Table reports OLS regression coefficients; $^*p < .10$, $^{**}p < .05$, $^{***}p < .01$ (two-tailed t tests).

Figure 5

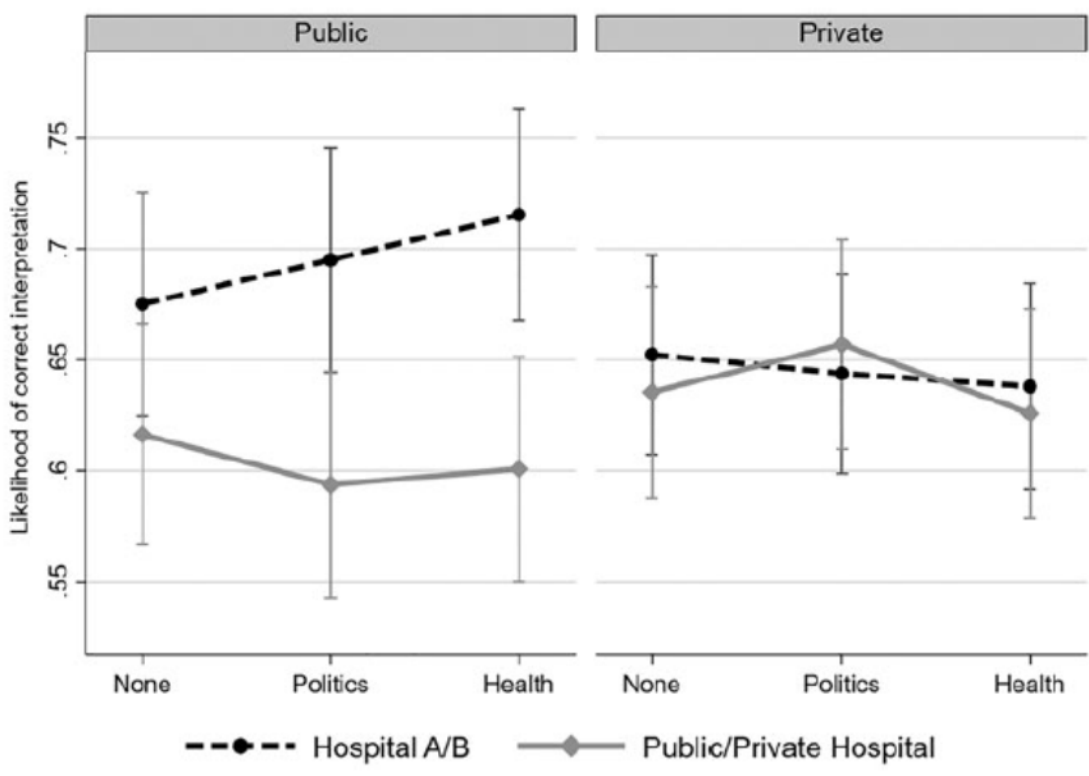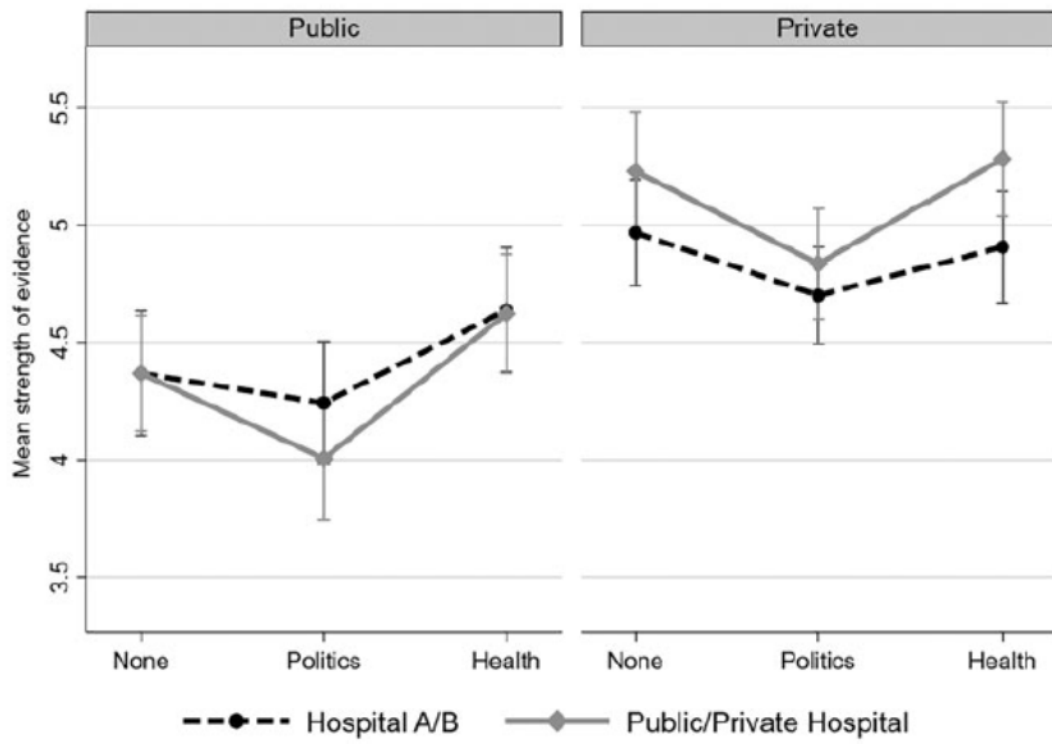Figure 6

TABLE 2
Regression Analysis of Likelihood of Correct Interpretation and Rating of Evidence Strength by Governance Preference—Denmark

	Likelihood of correct answer (OLS)				Rating of evidence strength (OLS)			
	Public		Private		Public		Private	
	Direct effects model (1)	Interaction model (2)	Direct effects model (3)	Interaction model (4)	Direct effects model (5)	Interaction model (6)	Direct effects model (7)	Interaction model (8)
Priming (ref = no prime)								
Politics prime	−.002	.020	.006	−.009	−.251	−.126	−.327**	−.268
	(.031)	(.045)	(.028)	(.039)	(.158)	(.229)	(.141)	(.195)
Health care prime	.011	.040	−.012	−.014	.261*	.271	−.005	−.061
	(.030)	(.044)	(.028)	(.039)	(.156)	(.225)	(.142)	(.197)
Labeling (Public/private)	−.092***	−.059	−.005	−.017	−.083	−.000	.257**	.262
	(.025)	(.043)	(.023)	(.040)	(.128)	(.221)	(.116)	(.200)
Interactions								
Politics prime × Labeling		−.043		.030		−.237		−.127
		(.061)		(.056)		(.316)		(.283)
Health prime × Labeling		−.056		.005		−.017		.114
		(.061)		(.057)		(.312)		(.283)
Constant	.692***	.675***	.647***	.652***	4.411***	4.368***	4.970***	4.968***
	(.025)	(.031)	(.023)	(.028)	(.129)	(.160)	(.114)	(.138)
R^2	.001	.001	.000	.001	.008	.009	.008	.008
N	1,458	1,458	1,735	1,735	1,373	1,373	1,580	1,580

Note: Table reports OLS regression coefficients; standard errors in parentheses; *$p < .10$, **$p < .05$, ***$p < .01$.

38